

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE
SANTA CATARINA – CÂMPUS FLORIANÓPOLIS
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
MESTRADO PROFISSIONAL EM SISTEMAS DE ENERGIA ELÉTRICA**

NEIMAR DE ALMEIDA OLIVA

**PREVISÃO DE GERAÇÃO FOTOVOLTAICA POR *PERCEPTRON*
MULTICAMADAS COM BASE EM GEOPOSICIONAMENTO E DADOS
METEOROLÓGICOS**

FLORIANÓPOLIS, 2022.

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE
SANTA CATARINA – CÂMPUS FLORIANÓPOLIS
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
MESTRADO PROFISSIONAL EM SISTEMAS DE ENERGIA ELÉTRICA**

NEIMAR DE ALMEIDA OLIVA

**PREVISÃO DE GERAÇÃO FOTOVOLTAICA POR *PERCEPTRON*
MULTICAMADAS COM BASE EM GEOPOSICIONAMENTO E DADOS
METEOROLÓGICOS**

Dissertação de mestrado submetida ao Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica.

Professor Orientador: Márcio Silveira Ortmann, Prof. Dr. Eng.

Coorientador: Jackson Lago, Prof. Dr. Eng.

FLORIANÓPOLIS, 2022.

CDD 621.3

O481p

Oliva, Neimar de Almeida

Previsão de geração fotovoltaica por perceptron multicamadas com base em geoposicionamento e dados meteorológicos - [DIS] / Neimar de Almeida Oliva, orientação de Márcio Silveira Ortmann; coorientação de Jackson Lago – Florianópolis, 2022.

1 v.: il.

Dissertação de Mestrado (Sistemas de Energia) – Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina.

Inclui referências.

1. Previsão de geração solar fotovoltaica. 2. Redes neurais artificiais. 3. Perceptron multicamadas. 4. Plantas virtuais de energia. 5. Microrredes I. Ortmann, Márcio Silveira. II. Lago, Jackson. III. Título.

PREVISÃO DE GERAÇÃO FOTOVOLTAICA POR *PERCEPTRON* MULTICAMADAS COM BASE EM GEOPOSICIONAMENTO E DADOS METEOROLÓGICOS

NEIMAR DE ALMEIDA OLIVA

Este trabalho foi julgado adequado para obtenção do título de Mestre em Engenharia Elétrica e aprovado na sua forma final pela banca examinadora do Mestrado Profissional em Sistemas de Energia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina.

Florianópolis, 16 de dezembro, 2022.

Banca Examinadora:

Márcio Silveira Ortmann, Prof. Dr. Eng.
(Orientador)

Jackson Lago, Prof. Dr. Eng.
(Coorientador)

Sérgio Luciano Ávila, Prof. Dr. Eng.
(IFSC)

Cesare Quinteiro Pica, Dr. Eng.
(movE Eletromobilidade)

Rubipiara Cavalcante Fernandes, Prof. Dr. Eng.
(IFSC)

AGRADECIMENTOS

Aos colegas da Fundação Certi e aos parceiros da AES Brasil pela experiência compartilhada, oportunidade de desenvolvimento e inspiração para o tema.

À AES Brasil pelos dados disponibilizados. Ao National Renewable Energy Laboratory (NREL) do Departamento de Energia dos Estados Unidos e à distribuidora de energia Ausgrid da Austrália pela publicação voluntária de informações para o desenvolvimento de pesquisas na área das fontes de energia renováveis.

Ao Instituto Federal de Santa Catarina (IFSC) pela busca permanente de um ambiente de incentivo ao desenvolvimento pessoal e profissional da comunidade catarinense.

Aos idealizadores do Programa de Pós-Graduação em Engenharia Elétrica (PPGEEL) do Departamento de Eletrotécnica (DAE) do IFSC Campus Florianópolis, ao seu coordenador, Prof. Bruno Scortegagna Dupczak, e seu antecessor, Prof. Fabricio Yutaka Kuwabata Takigawa, e a todos os professores pelo empenho para o sucesso.

Aos colegas de turma pelas trocas de experiência e o companheirismo.

Ao meu orientador, Prof. Márcio Silveira Ortmann, e ao meu coorientador, Prof. Jackson Lago, pelo acompanhamento permanente, incentivo e direcionamentos assertivos ao longo de todo o trabalho, mesmo durante a pandemia.

Aos membros da banca, pelos comentários e contribuições essenciais para a materialização da versão final desse trabalho.

Muito obrigado!

*"Esse exemplo ilustra claramente o limite da
inteligência artificial: ela não cria nada,
não gera entendimento,
não produz generalização."*

Miguel Nicolelis

RESUMO

A geração fotovoltaica representa uma porta de entrada para os pequenos consumidores como prossumidores no mercado de energia, pois sua modularidade permite ajustar o investimento de acordo com a escala. Os esforços regulatórios no sentido de aumentar a participação do consumidor no mercado de energia brasileiro tendem a incentivar o agrupamento de cargas e geradores em Microrredes e, principalmente, em Plantas Virtuais de Energia, operadas por agentes varejistas, para ganhos de escala e simplificação operacional. Com sua lógica de agregação de recursos remotos este ente deve alavancar ainda mais esta fonte de energia e criar novas oportunidades de negócios aos seus participantes, as quais demandam algoritmos de planejamento e operação especializados para a atuação competitiva. Esses algoritmos dependem de previsões de consumo e geração adaptados a esta nova realidade através da utilização de dados possíveis de se obter e modelos preparados para a incorporação de novas usinas. Este trabalho propõe uma abordagem para previsão intradiária e do dia seguinte com a aplicação de *perceptron* multicamadas a partir de dados físicos calculados do geoposicionamento solar, fontes de previsões meteorológicas disponíveis a qualquer interessado e dados de geração das usinas fotovoltaicas. Para orientar o desenvolvimento e os testes são utilizados os históricos de três usinas do Brasil, Estados Unidos e Austrália. Os experimentos realizados sugerem, considerados os prós e contras, que é possível selecionar uma estrutura de rede neural artificial comum às três usinas. Os resultados são verificados de acordo com a forma como a previsão é operacionalizada em uma microrrede ou planta virtual, ou seja, com o crescimento gradativo do histórico de dados de geração, algo raro na bibliografia. A agregação das previsões dessas usinas confirma o poder intrínseco de redução de variabilidade proporcionado pelas plantas virtuais. A comparação com um algoritmo de floresta aleatória, a segmentação dos erros por atributo de entrada e a análise dos maiores erros contribuem para a compreensão do potencial do *perceptron* multicamadas nesse tipo de aplicação e suas limitações, bem como para o aperfeiçoamento futuro do método.

Palavras-chave: Previsão de geração solar fotovoltaica. Redes neurais artificiais. *Perceptron* multicamadas. Plantas virtuais de energia. Microrredes.

ABSTRACT

Photovoltaic generation represents a gateway for small consumers as prosumers in the energy market, as its modularity allows fitting the investment to scale. Regulatory efforts to increase consumer participation in the Brazilian energy market tend to encourage the grouping of loads and generators in Microgrids and Virtual Power Plants operated by retail agents for gains in scale and operational simplification. With its logic of aggregating remote resources, this entity should further leverage this energy source and create for its participants new business opportunities, which demand specialized planning and operation algorithms for competitive performance. These algorithms depend on consumption and generation forecasts adapted to the small-scale reality through the appliance of available data and models prepared for the incorporation of new plants. This work proposes an intraday and day-ahead forecasting approach with the application of a multilayer perceptron based on solar ge positioning calculated data, weather forecast sources available to any interested party, and generation data from the photovoltaic plants. Historical data from three plants in Brazil, the United States, and Australia guides the developing process. Considering the pros and cons, experiments suggest that it is possible to select a single structure for the artificial neural network of the three plants. The verification of results takes the gradual growth of the historical generation data, somehow rare in the bibliography, although expected in Microgrids or Virtual Power Plants scenarios. The aggregation of forecasts from these plants confirms the intrinsic power of the variability reduction provided by Virtual Power Plants. The comparison with the random forest algorithm results added to the evaluation of errors by input attribute all contributes to understanding the multilayer perceptron potential in this type of application and its limitations, as well as the future improvement of the method.

Keywords: Photovoltaic generation forecasting. Artificial neural networks. Multilayer perceptron. Virtual power plants. Microgrids.

LISTA DE FIGURAS

Figura 1 – Potência de geração FV para três diferentes condições do tempo . . .	25
Figura 2 – Movimento de translação da Terra	37
Figura 3 – Solstício de verão no hemisfério Norte	38
Figura 4 – Posicionamento solar visto da Terra	39
Figura 5 – Fusos horários convencionados	40
Figura 6 – Relação entre ângulo zenital e altitude com o coeficiente de massa de ar	40
Figura 7 – Características da radiação solar para as massas de ar AM0 e AM1,5	41
Figura 8 – Espectro de luz solar	42
Figura 9 – Inclinação dos painéis solares	43
Figura 10 – Telhado solar comercial	44
Figura 11 – Exemplo do particionamento da amostra para a validação cruzada <i>5-fold</i>	53
Figura 12 – Exemplo de uma árvore de decisão	56
Figura 13 – Estrutura básica de rede neural artificial com TLU – o <i>perceptron</i> . .	57
Figura 14 – Exemplos de classes linearmente separáveis e inseparáveis	58
Figura 15 – Principais variáveis do problema da geração solar fotovoltaica . . .	63
Figura 16 – Regiões de clima subtropical úmido	65
Figura 17 – Cronograma da amostra de desenvolvimento frente à amostra completa	66
Figura 18 – Transformação da escala do azimute por hemisfério	69
Figura 19 – Fluxo de triagem das horas claras	71
Figura 20 – Matriz de colinearidade entre os atributos da usina BRA (2019) . . .	74
Figura 21 – Histograma do atributo Temperatura Ambiente (°C)	75
Figura 22 – Histograma da Geração de Energia (p.u.)	76
Figura 23 – Diagrama de fluxo de dados da rotina de atualização	77
Figura 24 – Fluxograma da rotina de treinamento	78
Figura 25 – Esquema de atualização dos dados meteorológicos – Usina AUS . .	79
Figura 26 – Esquema de atualização dos dados meteorológicos – Usina EUA . .	79
Figura 27 – Fluxograma da rotina de previsão	80
Figura 28 – Esquema de atualização dos dados meteorológicos	81
Figura 29 – Árvore das decisões tomadas para construção do modelo MLP . . .	84
Figura 30 – Variação do tamanho do lote de treinamento por tamanho da amostra	89
Figura 31 – Resultados de validação cruzada da Usina EUA por alternativa . . .	90
Figura 32 – Resultados de validação cruzada da aplicação de dados aumentados	92
Figura 33 – Resultados de validação cruzada por tipo de regularização	94

Figura 34 – RMSE de validação cruzada por número de camadas ocultas e de parâmetros – Usina EUA	95
Figura 35 – R^2 de validação cruzada por número de camadas ocultas e de parâmetros – Usina EUA	95
Figura 36 – Tempo de execução do treinamento em função do número de camadas ocultas e de parâmetros (todas usinas)	96
Figura 37 – Histograma dos erros de validação cruzada da Usina AUS	97
Figura 38 – Linha do tempo dos dados de desenvolvimento e teste	101
Figura 39 – Ciclo de treinamento	102
Figura 40 – Resultado dos testes visto pela métrica da RMSE – Usina AUS	103
Figura 41 – Resultado dos testes visto pela métrica da RMSE	104
Figura 42 – Resultado dos testes visto pela métrica do MAE – Usina AUS	104
Figura 43 – Resultado dos testes visto pela métrica do MAE	105
Figura 44 – Resultado dos testes visto pela métrica do R^2 – Usina AUS	105
Figura 45 – Resultado dos testes visto pela métrica do R^2	106
Figura 46 – Curvas selecionadas de geração real vs. prevista – Usina AUS	108
Figura 47 – Curvas selecionadas de geração real vs. prevista – Usina BRA	109
Figura 48 – Curvas selecionadas de geração real vs. prevista – Usina EUA	110
Figura 49 – Resultado dos testes por ordem do erro absoluto	112
Figura 50 – Descolamento da geração real vs. cobertura de nuvens - Usina EUA	113
Figura 51 – Dados selecionados de treinamento e previsão – Usina EUA	114
Figura 52 – Caso de sombreamento pela queda de neve – Usina EUA	115
Figura 53 – Resultado da previsão por estação do ano e hora do dia	116
Figura 54 – Resultado da previsão por estação do ano e temperatura ambiente	117
Figura 55 – Resultado da previsão por estação do ano e cobertura de nuvens	118
Figura 56 – Resultado da previsão por estação do ano e índice de chuva	118
Figura 57 – Resultado da previsão por estação do ano e índice de neve – Usina EUA	119
Figura 58 – Resultado da previsão por estação do ano e umidade relativa do ar	119
Figura 59 – Resultado da previsão por estação do ano e pressão atmosférica	120
Figura 60 – Chance de céu aberto a parcialmente nublado	121
Figura 61 – Comparativo RNA vs. RF da RMSE	123
Figura 62 – Comparativo RNA vs. RF da RMSE da Usina EUA	124
Figura 63 – Comparativo das curvas R^2 (MM12) entre RNA e RF	124
Figura 64 – Comparativo das curvas R^2 (MM12) entre RNA e RF – Usina EUA	125
Figura 65 – Comparativo do MAE acumulado entre RNA e RF – Usina BRA	125
Figura 66 – Comparativo do R^2 acumulado entre RNA e RF – Usina EUA	126

LISTA DE TABELAS

Tabela 1 – Horizontes de previsão de geração de energia	46
Tabela 2 – Síntese comparativa dos métodos de aprendizado de máquina supervisionado utilizados na previsão de geração de energia solar fotovoltaica	55
Tabela 3 – Usinas de referência do trabalho	64
Tabela 4 – Base do sistema por p.u. de cada usina	67
Tabela 5 – Percentual de horas sem geração de energia na amostra de desenvolvimento	67
Tabela 6 – Lista dos atributos de entrada do modelo	70
Tabela 7 – Saída do modelo	70
Tabela 8 – Estatística dos atributos das usinas para o desenvolvimento	73
Tabela 9 – Estatística de geração das usinas na amostra de desenvolvimento	76
Tabela 10 – Rotinas para operacionalização da previsão de geração	77
Tabela 11 – Principais ferramentas computacionais utilizadas no desenvolvimento	83
Tabela 12 – Especificação da rede neural artificial	85
Tabela 13 – Otimizadores da biblioteca Keras do TensorFlow	87
Tabela 14 – Resultados com Ângulo Diário Solar vs. suas componentes trigonométricas	99
Tabela 15 – Especificação da rede neural artificial de teste	102
Tabela 16 – Resultados consolidados dos testes	107
Tabela 17 – Resultados do <i>pool</i> de usinas para 12 meses de previsão	111
Tabela 18 – Configuração padrão da Floresta Aleatória do pacote Tensorflow	123
Tabela 19 – Estações c/ melhor resultado para a RNA por usina – a partir do 2º ano	126

LISTA DE ABREVIATURAS E SIGLAS

ACL	Ambiente de Contratação Livre
AD	Árvore de decisão (<i>decision tree</i>)
AE	Efeito altitude (<i>altitude effect</i>)
AM	Aprendizado de máquina
ANEEL	Agência Nacional de Energia Elétrica
AP	Aprendizado profundo
API	Interface de programação de aplicativos (<i>application programming interface</i>)
CCEE	Câmara de Comercialização de Energia Elétrica
CNN	Rede neural convolucional (<i>convolutional neural network</i>)
COGE	Centro de Operações da Geração de Energia
CV	Comercializador Varejista
CVU	Custo variável unitário
DHI	Irradiação difusa horizontal (<i>diffuse horizontal irradiation</i>)
DNI	Irradiação direta normal (<i>direct normal irradiation</i>)
DOE	<i>Department of Energy</i>
DP	Desvio-padrão
ELU	Unidade linear exponencial (<i>exponential linear unit</i>)
ETR	Radiação extraterrestre (<i>extraterrestrial radiation</i>)
FV	Fotovoltaico
GBM	Máquina de aumento de gradiente (<i>gradient boosting machine</i>)
GHI	Irradiação global horizontal (<i>global horizontal irradiation</i>)
GLD	Gerenciamento pelo lado da demanda
GLM	Modelo linear generalizado (<i>generalized linear model</i>)
GMG	Grupo motor gerador
IA	Inteligência artificial
IoT	Internet das coisas (<i>internet of things</i>)
LPS	<i>Lab Partnering Service</i>
LSTM	Rede de memória de curto prazo longa (<i>long short-term memory</i>)
MAE	Erro absoluto médio (<i>mean absolute error</i>)

MBE	Erro médio (<i>mean bias error</i>)
MLP	<i>Perceptron</i> multicamadas (<i>multilayer perceptron</i>)
MLPE	Eletrônica de potência no nível do módulo (<i>module-level power electronics</i>)
MPPT	Rastreador de ponto de máxima potência (<i>maximum power point tracker</i>)
MSE	Erro quadrático médio (<i>mean square error</i>)
nMAE	MAE normalizado (<i>normalized MAE</i>)
NREL	<i>National Renewable Energy Laboratory</i>
NSRDB	<i>National Solar Radiation Database</i>
ONS	Operador Nacional do Sistema Elétrico
OOB	Validação <i>out-of-bag</i> (<i>out-of-bag validation</i>)
PDF	Função de distribuição de probabilidade
PL	Projeto de Lei
PLD	Preço de liquidação das diferenças
R²	Coefficiente de determinação
RED	Recurso energético distribuído
RF	Floresta aleatória (<i>random forest</i>)
ReLU	Unidade linear retificada (<i>rectified linear unit</i>)
RMSE	Raiz do erro quadrático médio (<i>root of mean square error</i>)
RNA	Rede neural artificial
RNN	Rede neural recorrente (<i>recurrent neural network</i>)
SGD	Gradiente descendente estocástico (<i>stochastic gradient descent</i>)
SGDM	SGD com momento (<i>SGD with momentum</i>)
SIN	Sistema Interligado Nacional
SMF	Sistema de Medição para Faturamento
SOLPOS	<i>Solar Position and Intensity Calculator</i>
STP	Condições padrão de teste (<i>standard test conditions</i>)
SVM	Máquina de vetores de suporte (<i>support vector machine</i>)
TLU	Unidade lógica de limiar (<i>threshold logic unit</i>)
UFV	Usina de geração fotovoltaica
UTC	Tempo universal coordenado (<i>universal time coordinated</i>)
VPP	Planta virtual de energia (<i>virtual power plant</i>)

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Justificativa	27
1.2	Definição do problema	28
1.3	Objetivos	28
1.3.1	Objetivo Geral	28
1.3.2	Objetivos Específicos	28
1.4	Estrutura do Trabalho	28
2	REVISÃO DA LITERATURA	31
2.1	Contexto das previsões em Plantas Virtuais e Microrredes	31
2.2	Aspectos físicos que interferem na geração fotovoltaica	36
2.2.1	Radiação solar	37
2.2.2	Posicionamento dos painéis solares FV	43
2.2.3	Limpeza dos painéis solares FV	44
2.2.4	Degradação dos painéis solares FV	45
2.3	Métodos de previsão de geração fotovoltaica	45
2.3.1	Características do modelo de previsão	48
2.3.2	Métricas de avaliação de desempenho	49
2.3.3	Critério de avaliação da capacidade de generalização do modelo	52
2.3.4	Aprendizado de Máquina	54
2.3.5	Redes neurais artificiais	56
2.4	Conclusão parcial	60
3	PREVISÃO DE GERAÇÃO SOLAR FOTOVOLTAICA	61
3.1	Bases do método proposto de previsão fotovoltaica	62
3.2	Prospecção de dados para o desenvolvimento da previsão de geração	64
3.2.1	Sincronização dos dados de diferentes fontes	69
3.2.2	Pré-processamento de entradas	71
3.2.3	Estatística dos dados de desenvolvimento	71
3.3	Operacionalização do método de previsão proposto	76
3.4	Justificativa da opção pelo <i>perceptron</i> multicamadas	81
3.5	Construção do <i>perceptron</i> multicamadas de previsão de geração fotovoltaica	82
3.5.1	Especificação completa do <i>perceptron</i> multicamadas (MLP) proposto	83
3.5.2	Seleção do otimizador	86

3.5.3	Normalização das entradas, função de ativação e estratégia de inicialização dos parâmetros	89
3.5.4	Estrutura da rede neural artificial	93
3.5.5	Seleção da função de custo	96
3.5.6	Definição dos atributos de entrada	98
3.5.7	Estratégia de parada do treinamento	99
3.6	Conclusão parcial	99
4	ESTUDO DE CASOS	101
4.1	Apresentação dos resultados de teste	103
4.2	Análise dos resultados de teste	111
4.3	Considerações sobre outros métodos	120
4.3.1	Comparação com Floresta Aleatória	122
4.4	Conclusão parcial	126
5	CONCLUSÃO	129
5.1	Sugestões para trabalhos futuros	131
	REFERÊNCIAS	135
	APÊNDICES	147
	APÊNDICE A – PERCEPTRON MULTICAMADAS	149
A.1	Método de otimização do gradiente descendente	152
A.2	Funções de custo	155
A.3	Funções de ativação	156
A.4	Critérios de parada do treinamento	158
A.5	Métodos de regularização em caso de sobreajuste	160
A.6	Normalização de entradas para estabilidade do treinamento	161
	APÊNDICE B – DADOS DO NREL DE USINAS DOS EUA	165
	APÊNDICE C – DADOS DO API DA OPENWEATHER	171
	APÊNDICE D – USINAS DE REFERÊNCIA	173
D.1	AUS: Usina residencial	173
D.2	BRA: COGE AES Bauru	175
D.3	EUA: 5 Executive Campus	176
	APÊNDICE E – HISTOGRAMAS DOS DADOS DE DESENVOLVIMENTO	181

APÊNDICE F – FORMAÇÃO DAS PARCELAS DE VALIDAÇÃO CRUZADA	187
APÊNDICE G – EXPERIMENTO 1	189
APÊNDICE H – EXPERIMENTO 2	195
APÊNDICE I – EXPERIMENTO 3	201
APÊNDICE J – EXPERIMENTO 4	209
APÊNDICE K – EXPERIMENTO 5	213
APÊNDICE L – EXPERIMENTO 6	217
APÊNDICE M – EXPERIMENTO 7	221
ANEXOS	227
ANEXO A – DIAGRAMAS DE TIPOS DE REDES NEURAIS	229
ANEXO B – DADOS DA AUSGRID DE USINAS DA AUSTRÁLIA .	231

1 INTRODUÇÃO

Por muitos anos os sistemas de potência foram operados de forma verticalizada, ou seja, uma mesma empresa podia controlar a produção, transmissão e distribuição de energia a um conjunto de consumidores. Os projetos de geração de energia tendiam a ser majoritariamente de grande porte e localizados nas proximidades das fontes de recursos de produção (combustíveis, desníveis de rios, água para resfriamento etc.), separados dos consumidores por extensas linhas de transmissão e distribuição, utilizando diferentes níveis de alta tensão, num esquema voltado para os fatores econômicos, segurança e qualidade de fornecimento. Esse tipo de estrutura de produção de energia é operado por uma hierarquia de centros de controle que facilitam o monitoramento e controle constantes do sistema. Portanto, a geração é instantaneamente ajustada à demanda de consumo (pelo monitoramento da frequência e usando modelos de previsão estabelecidos) e a tensão é controlada para estar dentro dos limites especificados por meio de dispositivos apropriados e coordenados (geradores, comutadores sob carga, compensadores de reativos, etc.). No entanto, este modo de operação foi sendo alterado devido a fatores técnicos, econômicos, sociais, ambientais e políticos decorrentes de (RIKOS *et al.*, 2008):

- a) saturação dos sistemas de transmissão e redução das margens de segurança;
- b) restrições ecológicas e geográficas;
- c) problemas de segurança e estabilidade;
- d) crescimento contínuo da demanda;
- e) necessidade de investimento para sustentar o crescimento da demanda;
- f) privatização, desregulamentação e ambiente competitivo do mercado de energia;
- g) emergência de novas tecnologias eficientes, com vantagens ecológicas, e possibilidade de combinação com a geração de calor.

O novo paradigma é caracterizado pela presença da geração distribuída e com controle autônomo, conectada à rede de distribuição, próxima dos centros consumidores, de pequeno e médio porte, combinada com a geração de calor ou sistemas de refrigeração, e contando com um marco regulatório mais favorável. As fontes renováveis de energia têm sido a principal aposta deste modelo, colaborando para a substituição do uso das fontes de energia tradicionais, principalmente térmicas a carvão, diesel ou gás, emissoras de CO² (AKHTER *et al.*, 2019). Essas fontes são

baseadas nas seguintes características (JEBLI; BELOUADHA; KABBAJ, 2020):

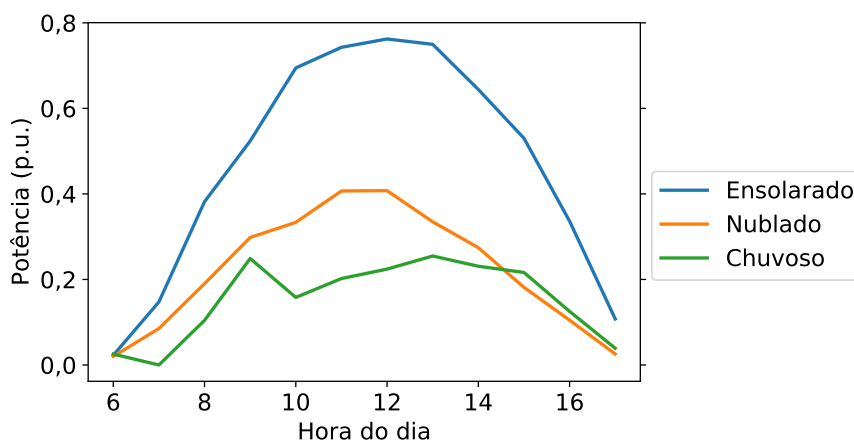
- a) são inesgotáveis, seja por suas grandes reservas ou porque são regeneradas naturalmente ou pelo homem;
- b) seu uso contribui para limitar a dependência dos combustíveis fósseis; e
- c) contribuem para a segurança do abastecimento de energia e a preservação do meio ambiente.

Dentre as fontes renováveis, a energia solar é uma das mais promissoras, e notadamente para geração de eletricidade por meio de sistemas fotovoltaicos (FVs), e mais recentemente, quase a totalidade conectada aos sistemas de transmissão e distribuição, onde desempenham um papel líquido positivo (GANDHI *et al.*, 2020; SHIVASHANKAR *et al.*, 2016). A capacidade global das instalações FVs excederam 100 GW em 2018 (ZHANG *et al.*, 2020). No Brasil, a potência fiscalizada pela Agência Nacional de Energia Elétrica (ANEEL) já ultrapassa 5 GW em quase 10 mil usinas (ANEEL, 2022). Na Alemanha, a geração FV cobriu cerca de 8,7% da demanda em 2018 (LEE; KO; LEE, 2020). Nos Estados Unidos, o Sunshot Vision Study reportou que em 2030 quase 14% da energia consumida no país será proveniente exclusivamente de fonte solar, e que até 2050 esses números devem quase dobrar para 27% (OSPINA; NEWAZ; FARUQUE, 2019).

Com o crescimento da penetração dos sistemas FV vêm alguns impactos negativos às redes. O perfil da energia gerada por este meio é altamente sensível aos fatores climáticos e sazonais, dada a variabilidade da disponibilidade da radiação solar, sujeita às condições atmosféricas, com impacto desfavorável à estabilidade, confiabilidade e operação da rede (AKHTER *et al.*, 2019). A Figura 1 apresenta uma curva de potência FV para três condições de tempo diferentes, representando um dia ensolarado, um de céu nublado e outro chuvoso. Em um dia ensolarado a irradiação solar é elevada e estável, com alta produção energética. Em dias chuvosos ou de céu nublado, a potência FV é baixa e instável (YANG *et al.*, 2014).

Essa característica intermitente por natureza da geração fotovoltaica introduz mais volatilidade ao sistema elétrico, por vezes afetando o nível de aceitação e adoção da tecnologia pelos agentes do setor. Se não há planejamento adequado, a variação de oferta e demanda de energia de acordo com o horário do dia pode provocar desequilíbrio na rede elétrica, e a injeção de energia superior à demanda pode ocasionar uma inversão no fluxo de potência, o que acarreta perdas excessivas e sobrecarga dos alimentadores (DANTAS; POMPERMAYER, 2018). Quando a produção energética é incerta os autoprodutores tendem a adquirir um volume excedente de energia no mercado ou a produzir em excesso por meios alternativos (despacháveis), e a participação dos produtores nos mercados de energia elétrica acaba prejudicada

Figura 1 – Potência de geração FV para três diferentes condições do tempo



Fonte: elaboração própria (2022).

com a possibilidade de prejuízos econômicos relevantes (KUMAR *et al.*, 2020; OSPINA; NEWAZ; FARUQUE, 2019). Outro ponto negativo é que parcela significativa da geração solar FV é considerada invisível, ou seja, não possui monitoramento por parte das distribuidoras de energia ou do operador do sistema elétrico, embora afete significativamente o perfil de carga sistêmico (SHAKER; ZAREIPOUR; WOOD, 2016).

Para lidar com essa variabilidade dos sistemas de geração fotovoltaica, várias iniciativas utilizam a combinação da geração fotovoltaica com a geração eólica ou com os sistemas de armazenamento, com reflexos positivos à confiabilidade do sistema conjugado (LIMA *et al.*, 2017; SANJARI; GOOI; NAIR, 2020). Shivashankar *et al.* (2016) recomendam a dispersão das usinas geograficamente, o uso de tecnologias de armazenamento, a integração com geradores diesel ou com células de combustível e até mesmo o corte de pico de potência ativa através do controlador rastreador de ponto de máxima potência (MPPT) dos inversores solares, com despejo de carga.

No Brasil, a Consulta Pública MME no. 21/2016 (CP21) iniciou uma série de discussões voltadas para a expansão do mercado de energia elétrica e o estudo dos benefícios e riscos envolvidos, de forma a subsidiar e definir visões institucionais capazes de estimular a eficiência e a inovação no setor elétrico. Atualmente, esses pontos são agregados pelo Projeto de Lei (PL) nº 414/2021 (BRASIL, 2021c) em discussão na Câmara dos Deputados após tramitar no Senado Federal (MONTENEGRO, 2022). Dentre essas propostas construídas a partir das discussões ao longo desses anos pode-se destacar:

- a) redução gradativa do limite para o enquadramento do consumidor livre de energia até a universalização do acesso;
- b) representação compulsória de pequenos consumidores por agente varejista, cujo enquadramento é facilitado, podendo este inclusive ser apenas

- agregador, não comercializando a energia;
- c) separação de lastro e energia;
- d) relaxamento da obrigatoriedade de contratação de energia;
- e) novo modelo de formação do preço de liquidação das diferenças (PLD) (CCEE, 2022b) por ofertas de quantidades e preços;
- f) estudo para implementação de bolsas de energia.

Essas propostas caminham no sentido da competitividade do mercado de energia e dão aos pequenos consumidores mais instrumentos de interferência no balanço entre a oferta e a demanda de energia. Através da representação por um agente varejista, suas cargas e seus geradores são agrupados com os de outros pequenos consumidores, saindo da invisibilidade e ganhando a escala necessária para negociar energia em condições melhores ou participar de programas de resposta da demanda e outros serviços ancilares. Os agentes varejistas têm sua atuação simplificada e podem, inclusive, ser apenas agregadores de carga, sem atuação direta na comercialização de energia. Em todos os casos, somente os agentes bem preparados terão condição de competir nesse hipotético¹ cenário do mercado de energia. Para isso, terão de lançar mão de algoritmos especializados de otimização dos recursos energéticos incluídos em suas carteiras.

As plantas virtuais de energia (VPPs) terão um papel importante para a viabilização desse cenário de inclusão e agregação de pequenos consumidores e prossumidores, ao dotar o agregador de ferramentas de gerenciamento ótimo dos recursos energéticos distribuídos (REDs). No escopo local, e para alguns consumidores, as microrredes farão a sincronização das usinas FVs com os demais recursos *behind-the-meter*. Os algoritmos das VPPs e das microrredes atuam a partir de dados históricos e prognósticos de diversos aspectos relacionados ao consumo energético, produção de energia, custos de combustíveis e operacionais, além dos preços de mercado. E dentre esses REDs, as pequenas usinas fotovoltaicas devem ganhar destaque, por serem uma das tecnologias de produção de energia mais modulares, escaláveis e acessíveis aos pequenos consumidores (MELLIT; KALOGIROU, 2017; DANTAS; POMPERMAYER, 2018).

Nesse contexto de aumento da importância das pequenas usinas no mercado de energia elétrica e sua participação nos mercados de energia, a previsão de geração solar fotovoltaica para esse nicho, geralmente desprovido de estações meteorológicas ou de capacidade de integração, ganha importância.

Este trabalho desenvolve um método de previsão de geração solar fotovoltaica factível, voltado para pequenas usinas desprovidas de estação meteorológica,

¹Ainda é um projeto de lei.

participantes de VPPs ou integradas a outros REDs através de uma microrrede, com a aplicação de *perceptron* multicamadas (MLP), uma rede neural artificial (RNA) clássica de aprendizado profundo (AP).

1.1 Justificativa

As fontes renováveis de energia, bem como os sistemas de armazenamento de energia – em função do crescimento da escala das redes inteligentes e das microrredes – terão um papel importante na modernização do setor elétrico e na mudança de paradigma da relação do consumidor com a geração de energia elétrica, em sintonia com os objetivos internacionais de redução de gases de efeito estufa (BRASIL, 2010; CASARIN, 2021; BRASIL, 2021b).

A geração solar fotovoltaica é uma das tecnologias mais modulares e escaláveis, e a que tem apresentado as maiores reduções de preço por unidade de potência instalada (DANTAS; POMPERMAYER, 2018). No entanto, possui uma característica intermitente, causadora de volatilidade ao sistema elétrico e às redes de distribuição, além de impactos econômicos aos seus proprietários – futuros fornecedores de energia ao Mercado Livre, através do Agente Varejista, à medida que o arcabouço legal se desenvolver.

As microrredes, que agregam recursos energéticos locais, e as plantas virtuais de energia, que agregam (virtualmente) recursos energéticos remotos, possuem algoritmos operacionais inteligentes capazes de integrar a geração solar fotovoltaica a outros recursos energéticos como os sistemas de armazenamento, geradores despacháveis ou cargas flexíveis ao desligamento, com efeito técnico e econômico benéfico. Porém, esses resultados são dependentes da qualidade das previsões de carga e geração de energia elétrica, uma vez que os despachos operacionais precisam ser decididos com antecedência.

Diferentemente das grandes usinas, que obrigatoriamente precisam possuir estações meteorológicas para participarem dos leilões de energia do mercado regulado (BRASIL, 2014), as pequenas não contam com estações meteorológicas ou com facilidade de integração. Faz-se necessário adequar as necessidades de previsão à realidade, lançando mão de dados disponíveis e investimentos reduzidos.

Nesse sentido, propõe-se um método factível de previsão de geração solar fotovoltaica, para responder à necessidade dos sistemas integradores de recursos energéticos distribuídos, como as microrredes, plantas virtuais de energia e os sistemas isolados comunicáveis. A solução – não peculiar – visa atender ao crescimento gradativo do número de pequenas usinas atendidas, sem a necessidade de revisão frequente do modelo proposto, o que facilita a incorporação pelos otimizadores energéticos.

1.2 Definição do problema

É viável prever a geração de energia ativa de sistemas solares fotovoltaicos de pequeno porte a partir da associação de dados físicos e meteorológicos de empresas especializadas e dados de medição de geração de energia das usinas, com a aplicação de *perceptron* multicamadas (MLP)?

1.3 Objetivos

1.3.1 Objetivo Geral

O presente trabalho teve por objetivo desenvolver um método de previsão de geração solar fotovoltaica de curto prazo não peculiar para aplicações voltadas à operacionalização de microrredes e plantas virtuais de energia, lançando mão de *perceptron* multicamadas.

1.3.2 Objetivos Específicos

Para o alcance do objetivo geral apresentado, foram propostos os seguintes objetivos específicos:

- a) pesquisar os principais fatores físicos e meteorológicos explicativos da variação da incidência solar sobre os painéis fotovoltaicos;
- b) selecionar os dados de entrada, identificar as suas fontes e prepará-los para o desenvolvimento e os testes;
- c) propor o método de previsão de geração solar fotovoltaica e selecionar as configurações do *perceptron* multicamadas;
- d) simular previsões e analisar estatisticamente os resultados obtidos.

1.4 Estrutura do Trabalho

Este trabalho está dividido em cinco capítulos, treze apêndices e dois anexos. Neste primeiro capítulo é apresentada a introdução do trabalho, a justificativa do problema, os objetivos geral e específicos, além desta estrutura do trabalho. Procurou-se evidenciar a relevância da previsão de geração fotovoltaica diante de um cenário hipotético de inclusão dos pequenos consumidores no mercado de energia de forma agregada, o que demanda um método capaz de lidar com o cenário de monitoramento e dados limitados.

O Capítulo 2 apresenta a revisão bibliográfica. Inicialmente, busca-se a contextualização do uso da previsão de geração fotovoltaica na operação de microrredes e plantas virtuais. Então, reúnem-se alguns dos principais aspectos influenciadores da

disponibilidade final de irradiação solar, ligados ao geoposicionamento e às condições atmosféricas. A seguir, conceituam-se os modelos de previsão e algumas métricas de avaliação dos resultados, além das formas de validação do modelo. Apresentam-se os principais métodos de aprendizado de máquina utilizados para a solução do tipo de problema proposto, com destaque para as redes neurais artificiais e, notadamente, o *perceptron* multicamadas.

O Capítulo 3 parte das predefinições da construção do método de previsão e suas bases hipotéticas, e então apresenta a origem dos dados utilizados para o desenvolvimento. Então, procura apresentar a ideia geral do método proposto, seguindo ao detalhamento das decisões tomadas para sua construção e à apresentação dos experimentos realizados. Os apêndices foram utilizados em grande quantidade para tornar o texto do capítulo mais fluente, visto que o trabalho se baseia em uma série de experimentos com farta geração de informações.

O Capítulo 4 apresenta os resultados da aplicação do método proposto em três usinas fotovoltaicas de três regiões do planeta. O teste simula o início do monitoramento de geração dessas usinas e apresenta um quadro consolidado para diferentes etapas de maturidade dos dados, além de métricas mensais com o acúmulo gradativo de dados para treinamento da rede neural de acordo com o ciclo proposto de atualização, o que permite inferir sobre a qualidade da previsão em função do volume de dados históricos de geração. Curvas diárias de geração real e previsão para diferentes quartis da raiz do erro quadrático médio ajudam a visualizar graficamente o resultado para as diferentes usinas. Algumas visões cumulativas das métricas são apresentadas para mostrar a influência de um número reduzido de horas na deterioração dos valores médios; alguns casos são discutidos e é possível observar algumas vulnerabilidades do método. A capacidade intrínseca da agregação de usinas para a redução da variabilidade da previsão é abordada. Finalmente, compara-se o método proposto com o algoritmo Floresta Aleatória.

O Capítulo 5 reúne as principais conclusões do trabalho, pontos de atenção observados durante o desenvolvimento, as principais dificuldades, bem como várias sugestões para desenvolvimentos futuros.

O Apêndice A apresenta os principais aspectos relacionados à construção de um modelo de previsão baseado em *perceptron* multicamadas. O Apêndice B explica a forma de obtenção dos dados de medição de uma usina dos EUA. O Apêndice C apresenta os dados meteorológicos utilizados no trabalho. O Apêndice D apresenta as características das três usinas de referência desse trabalho. O Apêndice E comenta os histogramas dos dados utilizados durante os experimentos para a definição dos hiperparâmetros da rede neural artificial. O Apêndice F detalha o processo de formação das parcelas de validação cruzada utilizadas nos experimentos dos Apêndices G a M,

realizados para justificar as decisões da estrutura do modelo proposto.

O Anexo A – proveniente de Veen e Leijnen (2019) – reproduz os diagramas de uma parcela das redes neurais artificiais já desenvolvidas. E, finalmente, o Anexo B – de Ausgrid (2014) – reproduz a especificação dos dados de uma das usinas utilizadas no trabalho.

Com o intuito de facilitar a leitura e conexão com a bibliografia da área, algumas abreviaturas e siglas são mantidas em inglês, apesar da inserção da descrição traduzida no texto. As descrições em inglês encontram-se na Lista de Abreviaturas e Siglas.

2 REVISÃO DA LITERATURA

2.1 Contexto das previsões em Plantas Virtuais e Microrredes

Uma Planta Virtual de Energia (VPP) é uma entidade que agrega recursos remotos de geração e de carga flexível para criar um único perfil de operação técnica ou comercial a partir da união das características individuais desses recursos. Uma vez agregados, esses participantes individuais podem atingir diferentes segmentos do mercado de energia, como se fossem um consumidor ou produtor convencional – em porte e características – desses segmentos, permitindo a provisão de serviços ancilares e resposta da demanda aos mercados de energia e capacidade. Pode-se distinguir as VPPs entre técnicas e comerciais de acordo com seu foco operacional. Uma VPP técnica lida principalmente com a influência dos REDs à operação dos sistemas de transmissão e distribuição, podendo receber comandos dos operadores desses sistemas. Já uma VPP comercial é independente dos sistemas de transmissão e distribuição, sendo otimizada para agir em resposta aos comandos de preços dos mercados (LEITES *et al.*, 2018; NAVAL; YUSTA, 2021).

Atualmente, uma VPP comercial pode ser operada através da figura do Comercializador Varejista (CV) (BRASIL, 2021a) no chamado Mercado Livre de Energia – ou mais formalmente, Ambiente de Contratação Livre (ACL) (BRASIL, 2004a; BRASIL, 2004b). Esse ente é autorizado a representar as cargas e os geradores aptos a participar do ACL de forma independente, mas que não desejam ou não possuem capacitação para gerir os compromissos envolvidos (processo de adesão, acompanhamento da medição, contratação de garantias, fechamento de balanço energético neutro, aquisição ou venda de energia, cumprimento do calendário de obrigações operacionais e financeiras etc.), dentro de alguns limites (CCEE, 2022a). Mas apesar dos esforços permanentes da Câmara de Comercialização de Energia Elétrica (CCEE), da Agência Nacional de Energia Elétrica (ANEEL) (ANEEL, 2021a; ANEEL, 2021b), dos legisladores (BRASIL, 2021c; BRASIL, 2022) e dos demais atores do setor elétrico (*stakeholders*), a maioria do potencial relacionado à agregação de REDs no Brasil ainda carece de um marco legal ou alguma regulamentação para sua consolidação, principalmente nas temáticas da universalização do acesso ao ACL e aos programas de resposta da demanda, do fornecimento de serviços ancilares, do armazenamento de energia e da simplificação do Sistema de Medição para Faturamento (SMF).

Já no âmbito local, uma microrrede pode ser simplificada descrita como um grupo de REDs e cargas controláveis e não controláveis que operam sincronizadas com a rede ou não. Através do gerenciamento da demanda e a operação de

seus REDs as microrredes colaboram para a redução do consumo de combustíveis fósseis e o corte de picos de carga, contribuindo para a postergação de investimentos nas linhas de transmissão e distribuição (TENFEN; FINARDI, 2015). Assim como as VPPs, as microrredes carecem de uma evolução regulatória capaz de assegurar sua máxima eficiência e retribuição ao sistema elétrico. Dentre as principais demandas, pode-se citar:

- a) acesso universal e simplificado para a aglutinação de unidades consumidoras contíguas;
- b) possibilidade de participação da microgeração e minigeração distribuída de forma transparente, competitiva e simplificada no fornecimento ancilar às distribuidoras de energia elétrica, com iniciativa em andamento (ANEEL, 2021a);
- c) regulamentação do armazenamento de energia, com iniciativa em andamento (BRASIL, 2022); e
- d) unificação de critérios técnicos entre as distribuidoras e maior clareza nas condições para o ilhamento intencional, com a remoção de embaraços técnico-econômicos.

Tanto as VPPs quanto as microrredes operam de forma automática e segundo o despacho fornecido por seu sistema de otimização com vistas a garantir o melhor uso dos recursos energéticos disponíveis (MAKOHIN *et al.*, 2018). Esses recursos podem ser próprios ou oferecidos pelo mercado. Os recursos energéticos próprios podem estar conectados a uma mesma unidade consumidora de energia, comum no caso das microrredes atuais, ou em diferentes locais, distribuidoras ou submercados, como no caso das VPPs.

Durante cerca de quatro anos, este autor participou do quadro de pesquisadores de três projetos P&D ANEEL¹ relacionados às microrredes e plantas virtuais de energia comerciais para o mercado nacional e sob as condições locais. Nesse contexto, foi possível desenvolver uma abordagem própria da operacionalização desses sistemas, com o gerenciamento dos recursos energéticos podendo envolver, dentre outros REDs:

- a) cargas: que podem ter diferentes curvas de carga – mais, ou menos planas, ou sazonais – podendo ser comuns ou críticas. Cargas comuns são a base do consumo de energia das unidades consumidoras e podem sofrer pequenas interrupções sem grandes prejuízos. As cargas críticas

¹Os projetos PD-0054-1051/2017 (Virtual Power Plant AES com Foco em Agregação de Grandes Clientes e Mercado Energético - Fase I), PD-0064-1054/2019 (Sistema Inteligente de Controle e Otimização de Microrredes - Fase II) e PD-0064-1065/2020 (Virtual Power Plant AES com Foco em Agregação de Grandes Clientes e Mercado Energético - Fase II), de iniciativa da concessionária de geração AES Brasil, tiveram a Fundação Certi como executora (AES BRASIL, 2021).

são aquelas que envolvem processos cuja interrupção não planejada pode causar riscos à vida (equipamentos de transporte, setores hospitalares etc.), aos equipamentos (usinagem, tratamento térmico etc.) ou ao negócio (empresas de rádio e televisão, datacenters, congressos etc.) (PEREIRA, 2003);

- b) contratação de energia: compreende os valores monetários aplicáveis ao consumo de energia ativa dos mercados cativo e livre. O mercado cativo possui preços regulamentados para os horários de ponta, intermediário e fora de ponta – válidos por até um ano. Além disso, abrangem a aplicação das bandeiras tarifárias. Já no mercado livre a energia é negociada livremente entre os agentes, em condições não regulamentadas. Existe uma distinção entre o mundo físico, a operação do Sistema Interligado Nacional (SIN) – operado pelo Operador Nacional do Sistema Elétrico (ONS), e o financeiro, tratado pelos agentes sob supervisão da CCEE. A liquidação mensal do mercado, atribuída à CCEE, ocorre a preços de liquidação horários, calculados e divulgados somente na noite anterior ao dia da operação (CCEE, 2022b);
- c) contratação de demanda: é a demanda faturável da distribuidora. É comum aos mercados livre e cativo, embora com algumas particularidades técnicas (requisitos das instalações) e econômicas ou operacionais (desconto de energia incentivada, sistema de compensação de energia). Atualmente é aplicável apenas aos consumidores de média e alta tensão;
- d) resposta da demanda: quando a unidade consumidora possui algum tipo de atividade que pode ser adiantada, atrasada, ou completamente suprimida, em troca dos ganhos econômicos proporcionados pela revenda da energia previamente contratada. A viabilidade das ações de gerenciamento pelo lado da demanda (GLD) é mais comum quando o consumidor possui algum processo intensivo no uso de energia (OLIVEIRA *et al.*, 2019) ou cujo valor do produto seja flutuante, como na produção de *commodities* industriais (alumínio, cimento, produtos químicos etc.), mas o consumo residencial tende a ter grande participação no futuro (KIRKERUD; NAGEL; BOLKESJØ, 2021);
- e) geração despachável de energia: na presença de um ou mais grupos motores geradores (GMGs) a diesel ou gás natural (e futuramente hidrogênio). Esses equipamentos são controláveis, mas pouco eficientes economicamente para a produção de energia elétrica. Sua viabilidade costuma estar associada à redução do nível de contratação de demanda da distribuidora ou à garantia de disponibilidade energética (*backup*) aos

processos críticos. Tanto o custo dos combustíveis quanto o benefício proporcionado por esses geradores podem variar ao longo da operação, o que pode levar a uma operação deficitária. Novas propostas de integração com energia solar fotovoltaica podem trazer sobrevida a essa fonte de geração (HAMILTON; NEGNEVITSKY; WANG, 2019);

- f) geração não despachável: para a redução do consumo de energia da rede, seja em função dos preços da energia elétrica (e da tarifa de distribuição) ou da participação em programas de consumo responsável de energia elétrica, através da substituição das fontes tradicionais por fontes renováveis, geralmente solar fotovoltaica ou eólica. No entanto, pela variabilidade de geração dessas fontes, não descarta a necessidade de garantir a demanda de energia, em geral, pela contratação da própria rede de distribuição, mas também através de geradores despacháveis;
- g) armazenamento de energia: muito presente em sistemas isolados e raro em sistemas conectados onde a viabilidade econômica ainda é uma promessa. A tecnologia de armazenamento mais comum é por baterias. As microrredes, redes inteligentes e a eletromobilidade devem induzir o avanço tecnológico e dos sistemas produtivos das baterias. Além disso, espera-se um avanço regulatório, com o reconhecimento e aceitação desse recurso como uma das soluções ao problema dos picos de consumo, independente da fonte de energia (BRASIL, 2021c).

Assim, o controlador de microrrede ou o sistema de gerenciamento de uma planta virtual busca otimizar a oferta e a demanda de energia a todo instante, para atingir o balanço neutro, considerando as restrições existentes, os custos e as oportunidades envolvidos, como:

- a) restrições:
 - i) cargas totais do sistema;
 - ii) cargas críticas a serem mantidas a qualquer custo²;
 - iii) capacidade e demais restrições operacionais dos geradores próprios, o que no caso dos equipamentos não despacháveis coincide com a própria geração;
 - iv) fluxo de suprimento e/ou capacidade de armazenamento de combustíveis;
 - v) capacidade do sistema de armazenamento, bem como seus limites ou perfis de carga e descarga;

²Esta afirmação pressupõe que os custos do desligamento das cargas críticas, muitas vezes imensuráveis, ultrapassam o maior custo variável unitário (CVU) de geração envolvido no problema, seja próprio ou ofertado pelo mercado.

- vi) limite de capacidade do ponto de conexão com a rede de distribuição;
- vii) limites de demanda contratada da distribuidora (quando aplicável);
- viii) perfil da multa de ultrapassagem da demanda contratada da distribuidora;
- ix) perfil da multa pela contratação de energia em montante inferior ao consumo dos últimos 12 meses;

b) custos:

- i) CVU dos geradores despacháveis próprios;
- ii) custo horo-sazonal da energia proveniente da rede de distribuição (mercado cativo);
- iii) custo das bandeiras tarifárias vigentes;
- iv) custo da energia contratada no mercado livre (contratos flexíveis e/ou modulados pela carga);
- v) valor do PLD horário, em caso de déficit;
- vi) custo dos combustíveis dos geradores despacháveis;
- vii) custo marginal de depreciação do sistema de armazenamento pelo carregamento, descarregamento ou manutenção do estado de carga;

c) oportunidades:

- i) valor do PLD horário, em caso de superávit;
- ii) *Spread* da energia no mercado de curto prazo;
- iii) programas de resposta da demanda;
- iv) participação em esquemas de serviços ancilares.

Todos esses fatores alimentam uma rede de relações que pode ser resolvida por um modelo de otimização visando definir as variáveis de decisão em função dos parâmetros conhecidos (ou estimados), num processo de planejamento cíclico em que, a cada período de discretização, avança-se no tempo o planejamento. As variáveis de decisão podem envolver o desligamento de cargas, o despacho de geradores, o esquema das baterias ou a contratação de energia, por exemplo. Essas são, em última análise, dependentes da estimação de parâmetros como: consumo de energia e nível de geração não despachável. Assume-se nesses modelos que a previsão desses recursos intermitentes é de qualidade (TENFEN; FINARDI, 2015).

Os parâmetros estimados do modelo de otimização das microrredes ou plantas virtuais se tornam mais exatos, ou precisos, à medida que se aproximam do presente, pois certos aspectos já se encontram conhecidos, ou suas previsões são mais assertivas (GENDZH, 2014). Por exemplo: dado que a temperatura atual é de

20°C, o consumo do condicionador de ar torna-se mais previsível para a próxima hora do que seria para uma determinada hora do dia seguinte; da mesma forma, dado que o céu está 100% coberto de nuvens, ou está chovendo, torna-se mais provável o acerto da previsão da geração de energia solar fotovoltaica das próximas horas feita agora, do que aquela realizada no dia anterior. No entanto, o planejamento não se torna necessariamente mais preciso à medida que se aproxima da operação, porque há parâmetros do modelo de otimização que permanecem no futuro interferindo nas decisões de curto e curtíssimo prazos, como nesses exemplos:

- a) uma ultrapassagem da demanda contratada está prevista para o final do mês. Neste caso, se a ultrapassagem é inevitável, não há razão para dispendere recursos mais caros para evitá-la no dia atual, até o limite de ultrapassagem já previsto. Pode-se ainda incluir as incertezas nesta ponderação;
- b) há uma previsão de preços mais altos da energia para o próximo dia. Se os preços são previstos maiores no dia seguinte, respeitadas outras motivações técnicas ou econômicas, o sistema de armazenamento pode ter de terminar o dia atual totalmente carregado, para aproveitar este *spread* entre os preços de carga (mais baratos, atuais) e descarga (mais caros, futuros);
- c) a geração não despachável prevista para o próximo dia é baixa e o suprimento de combustível da geração despachável será insuficiente para garantir o respeito aos limites técnicos de consumo da rede. Faz-se necessário reduzir o consumo de combustível do dia de hoje.

Se alguma dessas previsões não se confirmar, o resultado será pior, visto do futuro, do que eventualmente poderia ser com previsões melhores. Assim, percebe-se a relação da eficiência da operação das microrredes e plantas virtuais com a qualidade das previsões de consumo, geração não despachável e preços.

2.2 Aspectos físicos que interferem na geração fotovoltaica

A geração solar FV é o processo de conversão de luz solar – composta de ondas eletromagnéticas de diversas frequências – diretamente em eletricidade usando células solares fotovoltaicas. Para uma visão sobre as bases do processo tecnológico envolvido na geração solar fotovoltaica pode-se sugerir o trabalho de Mellit e Kalogirou (2017). Os autores abordam o processo tecnológico envolvido na geração solar fotovoltaica, desde o processo de conversão de luz solar em energia através das células solares; os tipos e a eficiência dos materiais semicondutores; os tipos de painéis solares (incluindo filmes) e os sistemas FV; seus arranjos; a conexão à rede e

os sistemas isolados; as baterias e o sistema de armazenamento; o teste, medição e monitoramento; até os impactos ambientais e o descarte.

Os sistemas FV têm sua produção afetada por vários fatores, como: irradiação solar (e sua composição espectral), condições climáticas, pressão dos ventos, umidade relativa do ar, temperatura ambiente e temperatura dos painéis FV. As variações da irradiação solar, do ponto de vista do posicionamento do Sol em relação aos painéis FV durante o dia e ao longo do ano são previsíveis a todo instante, porém, as variações naturais do clima afetam esses parâmetros, impactando a quantidade de energia elétrica gerada (AKHTER *et al.*, 2019).

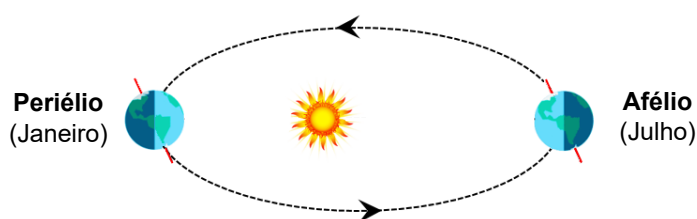
Atribui-se ainda uma parcela da variação da potência de saída dos sistemas às variações da eficiência de conversão – performance –, à degradação e aos mecanismos de falhas progressivos de longo-prazo. Um desafio para a garantia da qualidade e eficiência das instalações é salvaguardar sua performance através da identificação e quantificação dos fatores relacionados com os mecanismos de perdas. O comportamento das instalações FV em campo difere daquele previsto pelos manuais dos fabricantes, cujos dados são oriundos de testes de laboratório em condições padrão de teste (STP). Estudos empíricos demonstram que a tecnologia de silício cristalino exibe uma performance sazonal, variando ao longo do ano e ligeiramente entre os anos, principalmente em função de condições ambientais, como a irradiação global, temperatura ambiente, e a composição do espectro de luz. De forma similar, pode-se citar os filmes FV, embora alguns deles sejam menos sensíveis à temperatura (MELLIT; KALOGIROU, 2017).

2.2.1 Radiação solar

A conversão FV é dependente primariamente da quantidade (nível de intensidade) e qualidade (composição espectral) da luz solar incidente (MELLIT; KALOGIROU, 2017). Ambos são relacionados com os movimentos da Terra em torno do Sol e os fenômenos meteorológicos decorrentes.

O movimento de translação da Terra, em torno do Sol, ilustrado pela Figura 2, é elíptico e tem duração anual.

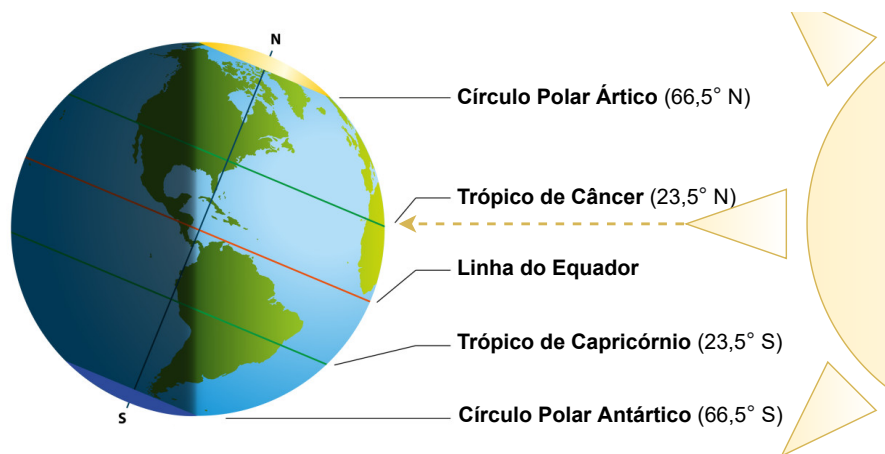
Figura 2 – Movimento de translação da Terra



Fonte: adaptado de tutorix.com (2021).

No início de janeiro de cada ano a Terra tem sua máxima proximidade do Sol (periélio) e no início de julho tem seu máximo afastamento (afélio). Isso não afeta o clima da Terra significativamente, mas os eixos de translação e rotação não são paralelos, fazendo com que a exposição ao Sol se alterne durante o ano em períodos de maior ou menor exposição em cada hemisfério. A Figura 3 ilustra o ponto de máxima exposição no hemisfério Norte, no chamado solstício de verão (do hemisfério). Nesse momento, os raios solares são perpendiculares a um plano horizontal localizado no Trópico de Câncer. Pontos ao norte desse trópico nunca apresentam o Sol a pino, ou seja, o sol do meio-dia nunca atinge o zênite – a elevação solar de 90° – como ilustra a Figura 4a.

Figura 3 – Solstício de verão no hemisfério Norte



Fonte: adaptado de kidspressmagazine.com (2022).

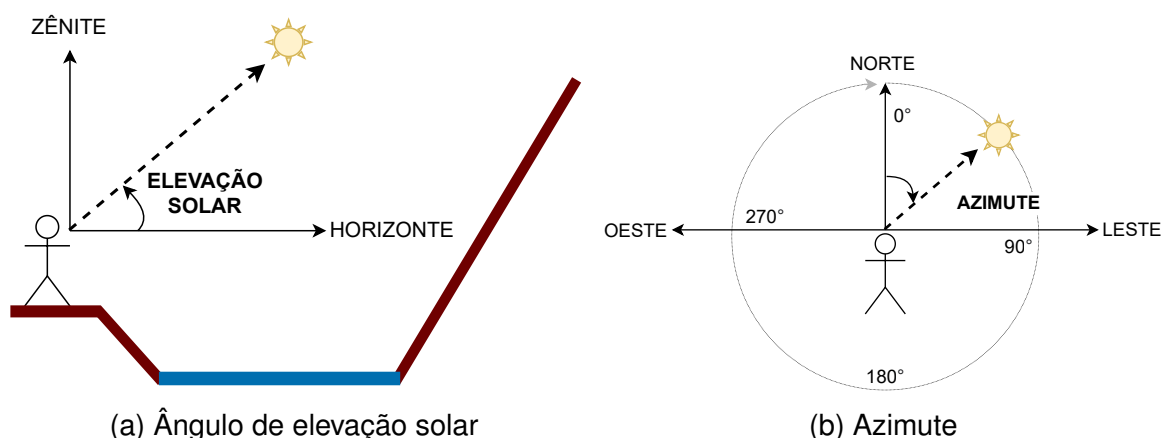
Similarmente, no hemisfério Sul o Sol nunca atinge o zênite para os pontos abaixo do Trópico de Capricórnio. Já entre os trópicos, a variação da declinação solar – inclinação de até $23,5^\circ$ do Sol em relação à linha do Equador, para o Norte ou para o Sul – faz com que os pontos entre os trópicos atinjam o zênite duas vezes por ano. Nessas latitudes da Terra, quando o Sol não está no zênite, pode estar ao Norte ou ao Sul, dependendo da época do ano.

O ângulo de elevação solar³ (Figura 4a), medido no local de interesse, do horizonte à posição do Sol, reflete a sazonalidade do posicionamento solar ao longo do ano e das horas do dia em qualquer ponto da Terra. Ele vai de 0° (horizonte) a 90° (zênite), mas para altitudes acima do nível do mar pode assumir valores negativos, abaixo do horizonte.

Para completar o georreferenciamento solar é utilizado o posicionamento do Sol em relação ao Norte geográfico, no sentido horário, através do ângulo chamado azimute, como ilustra a Figura 4b. A elevação solar e o azimute são conjuntamente

³Pela natureza do uso dos ângulos nesse trabalho, relacionada ao mundo real, e em consonância com as fontes de dados utilizadas nesse trabalho, todos os ângulos são expressos em graus.

Figura 4 – Posicionamento solar visto da Terra



Fonte: elaboração própria (2022).

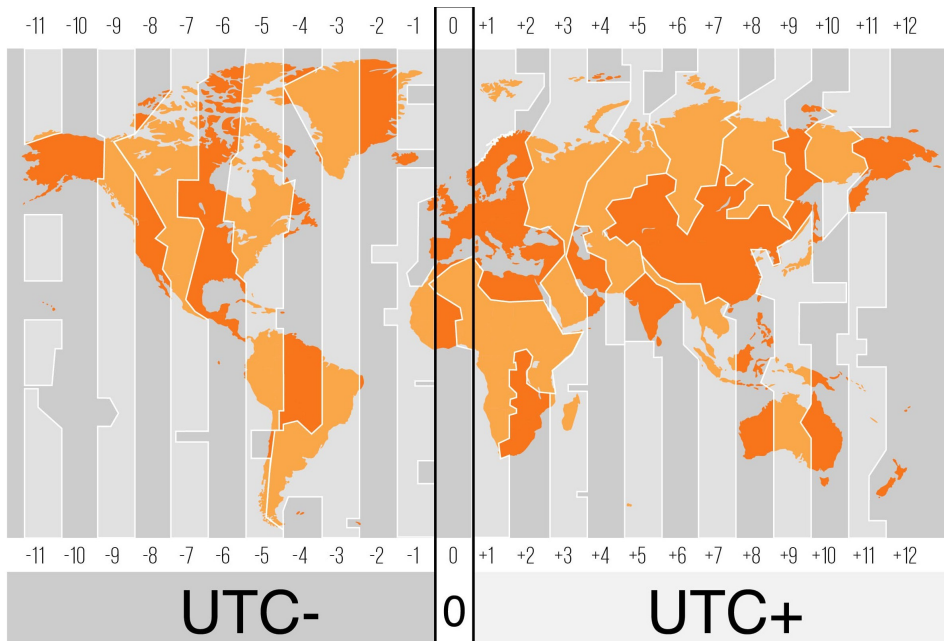
correlacionados com os ângulos horário e diário solares. O ângulo horário é uma referência em relação ao ponto de máxima elevação solar diária, quando atinge 0° . A escala cíclica vai de -180° (início da madrugada) a 180° (final da noite), ambos valores representando o mesmo momento. A passagem dos dias costuma ser descrita matematicamente ao numerar os dias continuamente ao longo do ano para produzir o chamado dia juliano (de 1 a 366). Cada dia do ano pode então ser expresso na forma angular, como ângulo diário, multiplicando-se o dia juliano por $360/365,25$. O ângulo diário é usado em várias expressões trigonométricas de cálculos de geoposicionamento solar (MELLIT; KALOGIROU, 2017).

O ângulo de elevação solar e o azimute são dados suficientes para desambiguar a posição solar a qualquer instante do ano num dado ponto da Terra e têm efeito semelhante ao uso dos ângulos horário e diário do mesmo ponto, porém com significado físico.

Vale ressaltar que a hora local é um dado convencionalizado conforme o tempo universal coordenado (UTC), tendo o meridiano de Greenwich como referência (zero) entre as longitudes, e os fusos-horários decididos politicamente, como mostra a Figura 5. Logo, não há coincidência entre as 12h (meio-dia) e o ângulo horário solar de zero graus. Pode-se perceber que a hora local convencionalizada é única dentro de uma faixa de longitudes e que para a mesma longitude podem ser encontrados fusos-horários distintos para latitudes distintas. Torna-se importante relacionar ambas métricas cuidadosamente para evitar interpretações equivocadas da radiação solar incidente no local.

O ângulo zenital (θ) é a medida entre o zênite e a direção do Sol num dado instante, e derivado matematicamente do ângulo de elevação solar (α), sendo:

Figura 5 – Fusos horários convencionados

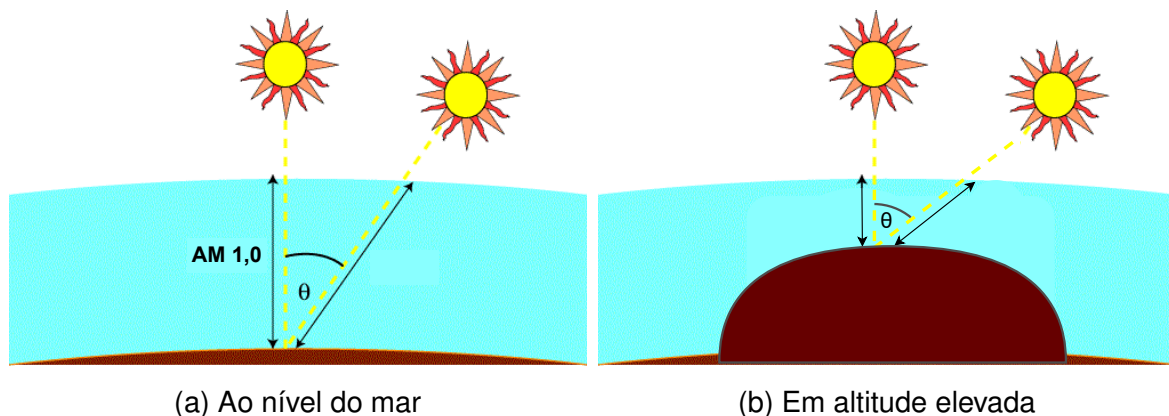


Fonte: 24timezones.com (2022).

$$\theta = 90^\circ - \alpha. \quad (1)$$

Dependendo do ângulo zenital e da altitude local, como mostra a Figura 6, as ondas eletromagnéticas do Sol tem de atravessar uma camada de massa de ar maior ou menor até atingir o solo. A massa de ar é internacionalmente definida pela sigla AM (do inglês *Air Mass*), sendo AM igual a 1,0 (para o nível do mar) quando o Sol atinge o zênite (VILLALVA; GAZOLI, 2012).

Figura 6 – Relação entre ângulo zenital e altitude com o coeficiente de massa de ar



Fonte: pveducation.org (2022).

Do feixe de radiação extraterrestre (ETR) proveniente do Sol apenas uma parcela consegue atravessar a atmosfera da Terra até atingir os painéis solares, com

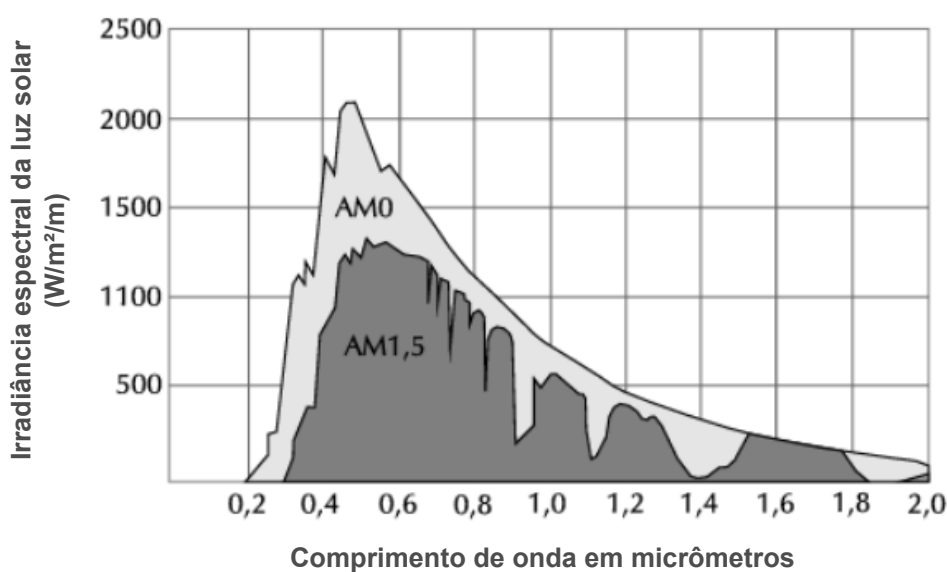
parte absorvida, refletida ou reemitida. Essa irradiação disponível na superfície terrestre tem duas componentes, chamadas irradiação difusa horizontal (DHI) e irradiação direta normal (DNI). A soma geométrica dessas componentes incidentes num plano horizontal na superfície da Terra é chamada de irradiação global horizontal (GHI) e representada por:

$$\text{GHI} = \text{DHI} + \text{DNI} \cos(\theta). \quad (2)$$

A DHI é a radiação solar sem sombra (céu aberto) coletada por uma superfície por unidade de área, a qual atinge a Terra omnidirecionalmente após ser dispersada por partículas da atmosfera ou outros obstáculos. Já a DNI pode ser descrita como um agregado da radiação solar perpendicular a uma superfície, em trajetória linear (direta) desde a posição atual do Sol (KUMAR *et al.*, 2020).

O espectro de luz da radiação solar que atinge o solo varia com a localização geográfica, a hora do dia, o dia do ano, as condições climáticas (incluindo a cobertura de nuvens), a composição da atmosfera, a altitude, o albedo local, e os efeitos do ângulo de incidência dentre outros fatores. O perfil espectral médio varia em função da massa de ar e pode ser obtido experimentalmente. A Figura 7 apresenta a distribuição AM0, que corresponde à radiação solar no espaço extraterrestre (acima da atmosfera), e a distribuição AM1,5 referente ao ângulo zenital $\theta = 48,5^\circ$, de regiões temperadas do hemisfério norte (VILLALVA; GAZOLI, 2012).

Figura 7 – Características da radiação solar para as massas de ar AM0 e AM1,5

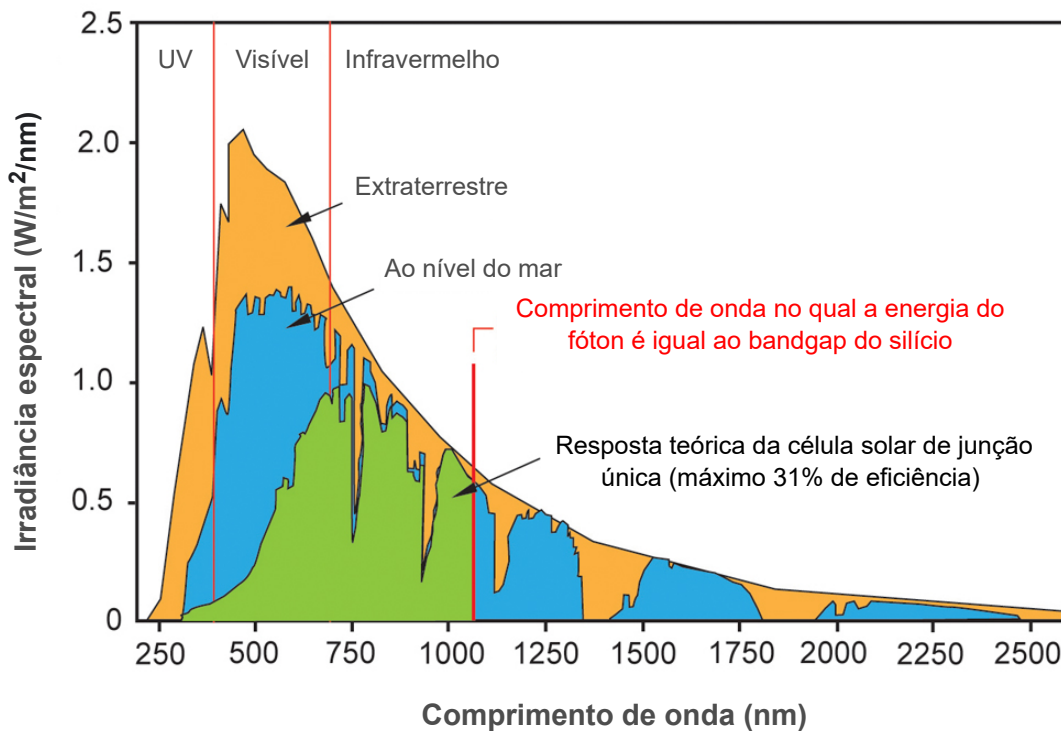


Fonte: adaptado de Villalva e Gazoli (2012).

Sob baixos níveis de irradiação a eficiência dos painéis solares decresce; a extensão varia com as tecnologias dos semicondutores utilizados. Por exemplo, como

ilustra a Figura 8, para a tecnologia de junção p-n única de silício dos painéis clássicos, os comprimentos de onda acima de 1.100 nm não contêm energia suficiente para conduzir os elétrons, ao passo que sob baixos comprimentos de onda, o excesso de energia é perdido em calor. Esses dois fatores combinados produzem um limite máximo teórico de eficiência dessa tecnologia da ordem de 31% – chamado limite Shockley-Queisser. Uma nova tecnologia baseada em junções agregadas para diferentes comprimentos de onda é uma forma de ultrapassar esse limite (VIRIDIAN SOLAR, 2022).

Figura 8 – Espectro de luz solar



Fonte: adaptado de Viridian Solar (2022).

O impacto da altitude na irradiação solar costuma ser tratado pelo chamado efeito altitude (AE), um aumento percentual por 1.000 metros em relação a um local num vale próximo, definido por:

$$AE = \left(\frac{I_m}{I_v} - 1 \right) \times \frac{\Delta(\text{altitude})}{1000} \times 100\%, \quad (3)$$

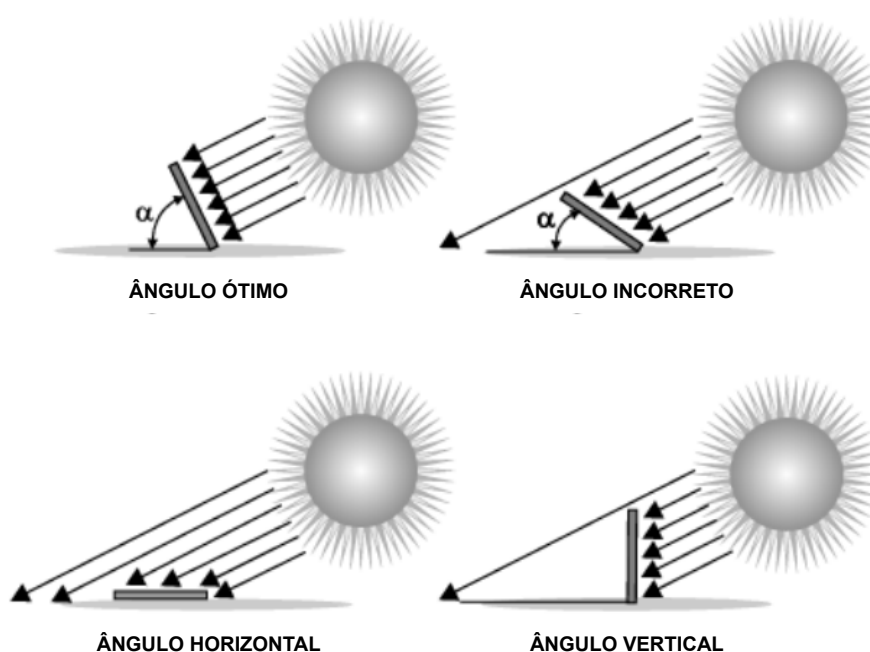
sendo que I_m é a irradiância na região elevada (montanha); I_v é a irradiância na região baixa (vale); e $\Delta(\text{altitude})$ a diferença de altitude entre as regiões, em metros. Para se ter uma ideia da dimensão dessa variação, Blumthaler, Ambach e Ellinger (1997) encontraram efeitos altitude totais (UVA + UVB) da ordem de 8% a 15% por 1.000 metros em estudo realizado na Áustria e Suíça envolvendo uma diferença de 3.000 metros entre os locais. Os principais fatores que influenciam no AE são os aerossóis –

partículas finas sólidas ou vapores – na camada atmosférica entre as altitudes, e os albedos locais. O albedo pode ser entendido como a capacidade de refletir a luz, ou ainda, a proporção de luz incidente na superfície ao redor que é refletida.

2.2.2 Posicionamento dos painéis solares FV

Uma vez que haja irradiação disponível no local da instalação do sistema solar FV, é preciso que os painéis solares estejam posicionados de forma adequada para o melhor aproveitamento da energia solar, e para que não haja sombra. Alguns sistemas FV utilizam mecanismos capazes de manter a face dos painéis solares perpendiculares à elevação solar e/ou ao azimute, os chamados seguidores solares ou *trackers*. Mas a maioria dos sistemas é estática e, nesse caso, possui eficiência reduzida em função do posicionamento, qualquer que seja. A Figura 9 apresenta quatro posições para uma mesma elevação solar máxima do dia. Com ângulo ótimo de inclinação, há 100% de aproveitamento dos raios solares para a área do painel.

Figura 9 – Inclinação dos painéis solares



Fonte: adaptado de Villalva e Gazoli (2012).

No entanto, o Sol está em constante movimento, logo, somente por um instante no ano ele poderia estar na condição de ângulo ótimo dos sistemas estáticos. E durante o dia, o azimute varia, outro fator contra a eficiência desses sistemas. Em termos matemáticos, a eficiência instantânea (η) de aproveitamento da irradiação direta dos painéis solares, quanto ao posicionamento, poder ser vista por:

$$\eta(\alpha, \beta, \psi, \Theta) = [\cos(\alpha) \operatorname{sen}(\beta) \cos(\psi - \Theta) + \sin(\alpha) \cos(\beta)] \times 100\%, \quad (4)$$

sendo que α é o ângulo de elevação solar; β é o ângulo de inclinação dos painéis solares; ψ é o ângulo azimute perpendicular à face dos painéis, geralmente posicionados em 0° (Norte) para instalações no hemisfério Sul e 180° (Sul) para as do hemisfério Norte; e Θ é o azimute (HONSBURG; BOWDEN, 2022).

Quando a elevação solar e o ângulo de inclinação estão perpendiculares, assim como o azimute e a face dos painéis, o aproveitamento da irradiação solar é máximo. É comum a instalação dos painéis solares sobre telhados, onde nem sempre é possível manter o azimute ou a inclinação ideais. Além disso, muitos sistemas utilizam *strings* em diferentes posições, como ilustra a Figura 10, com os painéis distribuídos para leste e oeste.

Figura 10 – Telhado solar comercial



Fonte: elysia.com.br (2021).

2.2.3 Limpeza dos painéis solares FV

Quando um painel solar pertencente a um conjunto ligado em série falha ou é afetado por sombra ou poeira, a corrente máxima é reduzida para todo o conjunto, prejudicando a potência gerada pela usina. Sombras causadas por árvores ou edifícios podem ser prevenidas através da seleção cuidadosa do local de instalação dos painéis, porém o sombreamento causado pela deposição de poeira tem um impacto contínuo na geração. Os painéis sombreados ou com falha podem atingir temperaturas elevadas devido ao chamado *hotspot phenomenon*, podendo ser irreversivelmente danificados. Como medida de controle e segurança, diodos *by-pass* são utilizados para isolar esses painéis do restante do conjunto, criando circuitos alternativos. No entanto, eles dão origem a novas perdas e criam máximos locais nas curvas PV, de forma que o algoritmo

MPPT que controla a operação do conversor pode se confundir e adotar o novo ponto de operação como máximo global (LI VIGNI *et al.*, 2015; WANG *et al.*, 2020).

O intervalo de limpeza dos painéis solares depende da severidade das deposições de poeira e do tamanho do sistema FV (MICHELI *et al.*, 2020). Em experimento realizado com um sistema FV localizado em Shanghai na China, composto de 2 *strings* de 4,05 kWp, buscou-se testar 4 metodologias de programação ótima da limpeza dos painéis solares. Cada string continha 18 módulos policristalinos, ou cerca de 30 m². Os resultados apontaram períodos ótimos de limpeza entre 28 e 49 dias dependendo da sensibilidade ao custo de limpeza (WANG *et al.*, 2020).

Para melhorar a performance das usinas FV sujeitas a irradiação irregular de seus painéis, várias arquiteturas de conversores são propostas na literatura. Quando a irradiação é uniforme, um único inversor (centralizado) pode ser usado em pequenas e médias usinas. Quando não é o caso, são usados inversores *string* ou micro inversores. Além disso, o uso de diferentes topologias de conexão da usina colabora para a melhoria de performance, como: série-paralelo, *total-cross-tied*, *bridge-linked* e *honey comb*. Mais recentemente vêm surgindo técnicas baseadas em sistemas de reconfiguração dinâmica: medindo ou estimando a radiação dos painéis e alterando as conexões elétricas entre eles (LI VIGNI *et al.*, 2015).

2.2.4 Degradação dos painéis solares FV

Como os sistemas FV consistem de vários componentes (ex.: painéis solares, conexões elétricas entre os painéis, inversor solar etc.), uma falha em qualquer um deles pode afetar drasticamente a eficiência e a confiabilidade do sistema inteiro. Apesar de haver vários sensores disponíveis para detecção de falhas, eles nem sempre funcionam corretamente em função das características não lineares da geração FV, ou por exemplo, na presença de controle do tipo MPPT – capaz de manter a operação em máxima potência, qualquer que seja a condição de radiação solar ou temperatura de trabalho dos painéis solares (LEE; KO; LEE, 2020; VILLALVA; GAZOLI, 2012).

2.3 Métodos de previsão de geração fotovoltaica

Os pesquisadores investigam meios de prever a intensidade de irradiação solar há muito tempo, com bons resultados aplicados à agricultura, construção, geração de energia etc. (SHI *et al.*, 2012). O escopo de previsão pode ser o sistema elétrico como um todo, uma região, um alimentador, uma subestação, uma usina ou mesmo um equipamento. Em geral, a variabilidade da geração renovável costuma ser muito maior que a das cargas e, portanto, mais difícil de prever (WANG *et al.*, 2018).

Em condições de céu aberto, a energia produzida por um sistema solar PV

pode ser estimada, com razoável precisão, em função da posição dos painéis solares em relação ao Sol e algumas métricas de eficiência do sistema. O desafio é prever os efeitos de nuvens, aerossóis e outras condições atmosféricas, além da variação da eficiência dos painéis em função da temperatura ambiente (TUOHY *et al.*, 2015). Para lidar com isso, pesquisadores e climatologistas costumam combinar dados, imagens, modelos numéricos ou híbridos. As diferentes técnicas de previsão de irradiação solar são necessárias para a produção de previsões em diferentes horizontes de tempo com vistas a uma integração confiável, contínua e eficiente da energia solar às redes (KUMAR *et al.*, 2020).

Os métodos de previsão dependem sobretudo de três aspectos: o período de previsão (antecedência e discretização), as informações disponíveis sobre o sistema (tecnologia, configuração, eficiência, histórico de geração etc.) e a aplicação (TUOHY *et al.*, 2015). Cada necessidade de decisão requer um horizonte de tempo de previsão diferente (JEBLI; BELOUADHA; KABBAJ, 2020; WAN *et al.*, 2016), como mostra a Tabela 1.

Tabela 1 – Horizontes de previsão de geração de energia

Tipo	Antecedência	Utilização
Curtíssimo prazo	Segundos a minutos	Controle de estoque de energia e fechamento de balanço no mercado de eletricidade. Tem especial importância para as microrredes e redes inteligentes
Curto prazo	Até 48h ou 72h	Crucial para diferentes tomadas de decisão no mercado de eletricidade e operação segura da geração e dos serviços de distribuição de energia, incluindo o despacho econômico, comissionamento, etc. Neste horizonte se encaixam os planejamentos intradiário e do dia seguinte com base no PLD horário
Médio prazo	Até 1 semana	Planejamento da manutenção dos sistemas FVs, usinas convencionais, transformadores e linhas de transmissão
Longo prazo	Meses a anos	Planos de expansão do sistema elétrico e planejamento da operação fotovoltaica

Fonte: adaptado de Jebli, Belouadha e Kabbaj (2020) e Wan *et al.* (2016).

De forma geral, as entradas dos métodos de previsão podem ser a intensidade solar e outras métricas (métodos indiretos) ou a própria saída do sistema

(métodos diretos) (SHI *et al.*, 2012). O conhecimento do perfil de geração solar fotovoltaica de uma usina é normalmente obtido através da medição direta da potência injetada na rede pelo inversor fotovoltaico. No entanto, as medições locais podem ser afetadas por ações exógenas de controle, as quais podem tornar as observações não correlacionadas com a verdadeira potência máxima da usina de geração fotovoltaica (UFV), e conseqüentemente, com a irradiância. Assim, o uso dessas observações como medidas de referência ou dados de treinamento para previsão de geração solar fotovoltaica podem ter um impacto prejudicial na estimativa do estado atual da usina e no desempenho das previsões (SCOLARI; SOSSAN; PAOLONE, 2018).

A abordagem desses métodos pode envolver técnicas determinísticas ou probabilísticas. As determinísticas têm sido largamente utilizadas para geração solar fotovoltaica, entretanto, devido às características intrínsecas, estes métodos possuem algumas deficiências. Como a confiabilidade das previsões é predeterminada pela performance histórica do modelo, a incerteza em tempo real é difícil de ser estimada, especialmente para uso das microrredes, cuja confiabilidade da operação depende principalmente das condições climáticas e ambientais. Assim, mais recentemente, tem crescido o interesse pelos métodos probabilísticos para cobrir as necessidades econômicas e de confiabilidade dos sistemas de potência (SHENG *et al.*, 2018). Alguns estudos procuram construir uma função de distribuição de probabilidade (PDF) da geração dos sistemas fotovoltaicos, ao invés de apenas definir um valor determinístico (SANJARI; GOOI, 2017).

Os métodos de previsão indiretos e determinísticos – foco deste trabalho – costumam envolver as mais variadas composições de variáveis de entrada. Um estudo conduzido na Coreia do Sul concluiu que o uso excessivo de variáveis correlacionadas como entrada dos modelos de previsão de geração solar pode aumentar o erro da previsão. A combinação de temperatura ambiente, temperatura dos painéis solares e umidade relativa do ar deram os melhores resultados entre as combinações testadas por Kim *et al.* (2019). Liu *et al.* (2015) acreditam que a relação entre a potência de saída e os vários fatores meteorológicos envolvidos é estocástica e não linear, onde diferentes tipos de clima levam a variações significativas da energia produzida. Logo, a modelagem deve ser realizada em função dos diferentes parâmetros meteorológicos, e a previsão de geração deve ser segmentada para diferentes condições de tempo: ensolarado, parcialmente nublado, nublado, chuvoso/nevando.

No início deste trabalho, duas principais referências nortearam a introdução ao estado da arte da previsão FV. Akhter *et al.* (2019) realizaram uma revisão bibliográfica extensa dos métodos de aprendizado de máquina (AM) e meta-heurísticos utilizados para a previsão solar fotovoltaica, e concluíram que os métodos combinados performam melhor que os métodos de AM ou matemáticos (modelos explícitos). Já

Kumar *et al.* (2020) revisaram os métodos sob a ótica da previsão de irradiação solar disponível aos painéis FV, o que é muito correlacionado com a geração de energia. Recomenda-se ainda a leitura de Alkhayat e Mehmood (2021), pois eles revisaram especificamente a aplicação de AP para a previsão de geração fotovoltaica e eólica. Suas conclusões parecem acompanhar os achados de Akhter *et al.* (2019) ao identificar a prevalência dos métodos combinados (ou híbridos), seguidos das redes neurais recorrentes (RNNs) e redes neurais convolucionais (CNNs). Os métodos probabilísticos e os *multi-steps* – vários intervalos à frente e não apenas o próximo intervalo de discretização –, como no presente trabalho, vêm ganhando destaque.

2.3.1 Características do modelo de previsão

O termo modelo, usado em múltiplos contextos, pode referir-se a uma declaração conceitual ou descrição textual de causa e efeito, a um modelo em pequena escala de um fenômeno físico testado em laboratório, ou a uma representação matemática de um conhecimento conceitual humano sobre o funcionamento de algum aspecto da natureza, dentre outros usos (RHINERHART, 2016). Os modelos matemáticos geralmente relacionam saídas às entradas, efeitos às causas, resultados às influências, respostas aos estímulos. Uma vez conhecido, pode ser usado para previsão ou antecipação de fenômenos naturais e, de outra maneira, pode-se reverter a sentença matemática do modelo de forma a saber o que fazer para obter determinado resultado. Rhinerhart (2016) classifica os modelos matemáticos segundo uma série de atributos opostos, os quais criam categorias de problemas que direcionam os métodos apropriados de regressão.

No caso da previsão do valor de geração solar fotovoltaica por *perceptron* multicamadas, a partir de dados meteorológicos e de geoposicionamento solar, compreende-se a seguinte adequação aos atributos propostos por aquele autor:

- a) em equilíbrio (em oposição ao dinâmico): embora as condições atmosféricas sejam regidas por um sistema dinâmico, e a irradiação solar apresente características cíclicas, a geração de energia dos sistemas FV – em si – pode ser tratada como um sistema em equilíbrio, que só depende das condições instantâneas das entradas, e cujas características não variam no tempo (no curto e médio prazos)⁴;
- b) empírico (em oposição ao fenomenológico): a forma do modelo não é baseada em fenômenos físicos envolvidos no problema;
- c) determinístico (em oposição ao estocástico): espera-se um valor único de saída e não uma distribuição de probabilidades;

⁴Desprezando-se a perda natural da eficiência do sistema com a vida útil e outras condições extremas de utilização (temperatura, poeira, umidade etc.) que causem sua degradação.

- d) não linear: os impactos dos coeficientes são não lineares, o que combinado com as restrições nas variáveis, impedem o uso de métodos de solução genéricos;
- e) contínuo (em oposição ao discreto): as variáveis descritivas do modelo são contínuas em nível macro, em sua maioria, embora algumas delas sejam obtidas a partir de medições de baixa precisão. Por exemplo: a cobertura de nuvens, uma possível variável do modelo de previsão, é obtida a partir de avaliação humana ou automatizada do número de quadrantes cobertos de nuvens em relação a um número total de quadrantes imaginados na abóboda celeste;
- f) não analítico: pela presença de descontinuidades, ou seja, pontos não deriváveis no modelo, característica das RNAs;
- g) implícito (em oposição ao explícito): requer algum procedimento iterativo e um critério de convergência, e não pode ser resolvido pela simples execução de uma sequência de operações matemáticas;
- h) com valor de saída único (em oposição ao multivalorado);
- i) multivariáveis: o número de atributos de entrada não é maior que a unidade;
- j) restrito: as entradas possuem limites físicos que mantêm o sistema estável;
- k) valoração (em oposição à classificação): a saída é quantitativa, um número real, e não qualitativa;
- l) intratável: não pode ser representado e analisado a partir de uma equação matemática.

A configuração de um modelo de previsão é o principal desafio envolvido. Uma rede neural artificial, por exemplo, envolve a seleção de aspectos quantitativos como o número de atributos de entrada, o número de camadas ocultas e o número de neurônios em cada camada; ou funcionais, como o tipo de função de ativação e de otimização. Uma vez definido, o algoritmo de regressão pode ser facilmente aplicado rotineiramente (RHINERHART, 2016).

2.3.2 Métricas de avaliação de desempenho

As métricas adequadas para a avaliação do desempenho dos modelos de previsão variam de acordo com o tipo de problema envolvido e a ferramenta utilizada. Os modelos de regressão determinísticos para a previsão de geração solar fotovoltaica geralmente são verificados através, dentre outras, de métricas como: erro médio (MBE),

erro absoluto médio (MAE), raiz do erro quadrático médio (RMSE) e coeficiente de determinação (R^2) (KOSTYLEV; PAVLOVSKI, 2011; ALKHAYAT; MEHMOOD, 2021).

Essas métricas estatísticas são de uso geral e largamente difundidas. O erro médio (MBE) – em inglês *mean bias error* – é uma métrica do viés médio das previsões, o qual pode indicar se o modelo subestima ou superestima a geração de energia ao longo de um período de avaliação, mas diz pouco sobre o poder de previsão em si (MELLIT; KALOGIROU, 2017; KOSTYLEV; PAVLOVSKI, 2011). Sua formulação é dada por:

$$\text{MBE}(X, h) = \frac{1}{m} \sum_{i=1}^m h(x_i) - y_i, \quad (5)$$

sendo que:

X – é o conjunto de atributos da amostra;

h – é a transformação avaliada (modelo);

m – é o número de exemplos da amostra;

x_i – é o conjunto de atributos do exemplo i da amostra;

y_i – é o valor de saída verdadeiro do exemplo i da amostra.

Erros médios reduzidos nas previsões de curto e médio prazos são importantes para as plantas virtuais porque alguns dos aspectos do gerenciamento energético local dependem do balanço de energia das fontes de produção e consumo ao longo de um período de tempo. Esse viés pode interferir em decisões que envolvem várias horas, como uma estratégia de armazenamento de energia dos sistemas PV, por exemplo, ou a cobertura do lastro de energia no Mercado Livre. No entanto, um viés médio global reduzido não é suficiente para garantir que não haja erros médios elevados na comparação entre os erros médios em diferentes extratos da amostra avaliada: períodos do dia, estações do ano ou condições atmosféricas, por exemplo.

O erro absoluto médio (MAE) – em inglês *mean absolute error* – é uma métrica da distância entre os vetores de previsão e de valores reais, também conhecido por norma ℓ_1 ou norma *Manhattan* (MELLIT; KALOGIROU, 2017). É dado por:

$$\text{MAE}(X, h) = \frac{1}{m} \sum_{i=1}^m |h(x_i) - y_i|. \quad (6)$$

Diferentemente do MBE, o MAE valora a média absoluta dos erros, o que se relaciona com a precisão esperada para um exemplo de uma amostra. Os erros extremos são tratados com mesmo peso que os erros menores, o que pode esconder variações importantes entre extratos da amostra.

A raiz do erro quadrático médio (RMSE) – em inglês *root mean squared*

error – equipara-se à norma ℓ_2 ou norma Euclidiana, e ajuda a valorar essas eventuais discrepâncias entre os extratos da amostra. É dada por:

$$\text{RMSE}(X, h) = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x_i) - y_i)^2}. \quad (7)$$

A RMSE é uma das métricas mais utilizadas para os problemas de regressão, e assim como o MAE, fornece uma referência do erro esperado do modelo de previsão, porém com um peso maior para os erros maiores, ao utilizar o quadrado dos erros. Ela é capaz de ponderar a distribuição dos erros, resultando num valor absoluto tão mais próximo do MAE quanto forem uniformes os erros dos exemplos da amostra avaliada. No limite pode ser numericamente igual ao MAE mas nunca inferior.

O coeficiente de determinação (R^2) é um número usualmente entre zero e um que mede o grau de associação entre duas variáveis, ou quanto da variabilidade observada é explicada pelo modelo de previsão (MOUSTRIS *et al.*, 2016). Ele pode assumir valores negativos quando os elementos do vetor de previsão estiverem pior correlacionados com o vetor de valores reais que uma reta horizontal média desses valores reais. Sua formulação é dada por:

$$R^2(X, h) = 1 - \frac{\sum_{i=1}^m (h(x_i) - y_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2}, \quad (8)$$

sendo que: \bar{y} é a média dos valores de saída verdadeiros (conhecidos) de todos exemplos da amostra.

É importante notar que para comparação entre resultados de amostras de tamanhos distintos, ou de usinas distintas, MBE, MAE e RMSE não são apropriados em suas versões originais pois estão relacionados com a escala das variáveis medidas. É preciso normalizá-las para este fim. A normalização dessas métricas não é única na literatura, sendo comuns as formas baseadas na média dos valores da amostra verificada ou na capacidade de produção da usina fotovoltaica (BLAGA *et al.*, 2019).

O R^2 , por outro lado, é uma métrica relativa e pode ser comparada entre amostras distintas, exceto quando o número de atributos de entrada é diferente entre elas, pois o R^2 sempre aumenta com o aumento do número de atributos (JAMES *et al.*, 2013). Nesses casos, recomenda-se o uso do R^2 ajustado, uma versão que penaliza a quantidade de variáveis de entrada. Sua formulação é dada por:

$$R^2 \text{ ajustado}(X, h) = 1 - \frac{\frac{\sum_{i=1}^m (h(x_i) - y_i)^2}{(m - d - 1)}}{\frac{\sum_{i=1}^m (y_i - \bar{y})^2}{(m - 1)}}, \quad (9)$$

sendo que d é o número de atributos de entrada do modelo.

2.3.3 Critério de avaliação da capacidade de generalização do modelo

Um dos maiores problemas dos modelos de regressão é saber se o resultado obtido a partir de uma amostra é representativo e pode ser estendido à população de origem desta amostra, ou seja, a capacidade de generalização do modelo (GIANCRISTOFARO; SALMASO, 2003).

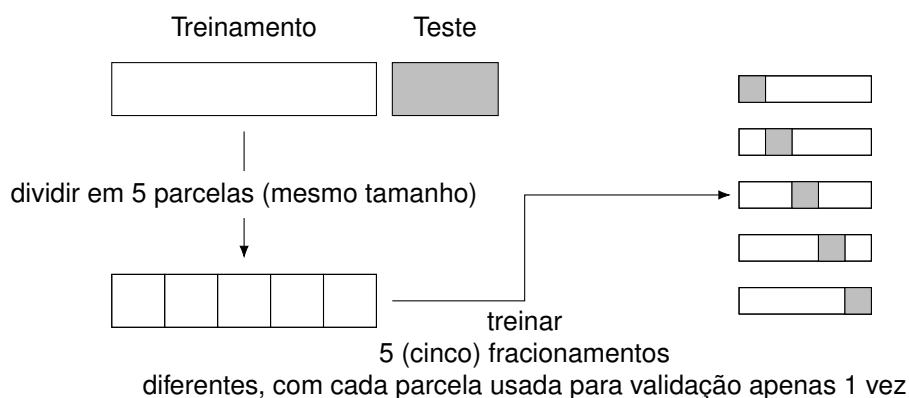
Idealmente, deveria haver uma amostra de treinamento e outra de teste em quantidade e diversidade de exemplos suficientes para garantir a confiabilidade dos resultados. Na prática isso não acontece, o que torna necessário encontrar formas de estimar o erro esperado dos modelos. Em geral, modelos complexos apresentam reduzido viés e elevada variância enquanto os mais simples apresentam elevado viés e reduzida variância. Busca-se o equilíbrio (*trade-off*) entre essas duas métricas que proporcione a melhor solução para o problema (KRSTAJIC *et al.*, 2014).

Quando se treina um modelo de previsão não se tem (supostamente) acesso à amostra de teste (novos exemplos de produção, ou uma parcela dos dados reservada para teste). No entanto, pode-se particionar os dados da amostra de treinamento em duas, para a chamada validação cruzada *holdout*. Uma proporção usual é 80% (treinamento) e 20% (validação). Este método não é satisfatório quando a quantidade de exemplos é insuficiente para garantir o aprendizado do problema específico (MURPHY, 2012).

Uma das técnicas simples e mais universalmente empregadas para a estimativa do viés e variância dos modelos de previsão, principalmente durante as fases de seleção entre alternativas de parâmetros e hiperparâmetros, é a chamada validação cruzada *k-fold* (KRSTAJIC *et al.*, 2014). A técnica consiste em dividir a amostra de treinamento em K parcelas (*folds*); então, para cada parcela $k \in \{1, \dots, K\}$, treina-se o conjunto das parcelas exceto a k -ésima, e testa-se com a k -ésima parcela, de forma rotativa, até que todas as parcelas tenham estado apenas uma vez fora do conjunto de treinamento, e apenas uma vez no conjunto de teste, como mostra a Figura 11 (MURPHY, 2012). Com os resultados de teste das K parcelas calculam-se a média e o desvio-padrão (DP) – ou variância – das métricas de interesse. O valor de K é geralmente reduzido em relação ao número de exemplos de treinamento, com $K=5$ e $K=10$ sendo dos mais populares na literatura pois já mostraram empiricamente que são

capazes de produzir vieses e variâncias moderados, embora não sejam regra (JAMES *et al.*, 2013; KUHN; JOHNSON, 2013). Nesse contexto, o viés e a variância referem-se à diferença entre a estimativa e o valor real da métrica de interesse e não aos próprios resultados do modelo em análise.

Figura 11 – Exemplo do particionamento da amostra para a validação cruzada 5-fold



Fonte: elaboração própria (2022).

Uma das variações do método, quando $K=N$, sendo N igual ao número de exemplos da amostra de treinamento, recebe o nome de validação cruzada *leave-one-out* (LOOCV). Este é muito intensivo computacionalmente visto que o treinamento e o teste devem ser repetidos N vezes. Há ainda variações como a validação cruzada aninhada e métodos similares como o *bootstrap* (KOHAVI, 1995).

A melhor opção obtida para uma série de bases de dados reais testadas por Kohavi (1995), mesmo quando o poder computacional permitia um fracionamento maior, foi a validação cruzada estratificada 10-fold. Na validação cruzada estratificada, os exemplos de cada parcela são selecionados de forma a garantir aproximadamente a mesma distribuição em todas as parcelas de validação, e similar àquela vista na população. Molinaro, Simon e Pfeiffer (2005) também obtiveram menores vieses e variâncias de validação cruzada *k-fold* ao trabalhar com a média de 10 realizações do processo completo.

Apesar de sua utilidade na modelagem do aprendizado de máquina, é preciso atentar-se ao fato de que a validação cruzada pode sofrer alta variabilidade, e levar a más escolhas na seleção do modelo e a um comportamento errático da estimativa do erro esperado. Bengio e Grandvalet (2004) alertam que a estimativa de variância de um modelo só pode ser feita a partir de realizações independentes, ou dependentes desde que a correlação entre as amostras seja conhecida, e mostram que a validação cruzada *k-fold* produz erros de teste interdependentes com vieses que podem chegar, em casos simples, à magnitude da variância.

2.3.4 Aprendizado de Máquina

O aprendizado de máquina (AM) é uma forma de inteligência artificial (IA) que permite a um sistema aprender a resolver problemas a partir do emprego de uma variedade de algoritmos que consomem dados de entrada para gerar melhorias, descrever dados e prever resultados, sem o uso de programação explícita. Usualmente, o AM é segmentado segundo duas abordagens (HURWITZ; KIRSCH, 2018):

- a) aprendizado supervisionado (ou preditivo): onde o objetivo é encontrar a correlação existente entre os dados disponíveis com relação a um dado de medição, significado ou referência – às vezes chamado de etiqueta, ou *label* –, a fim de realizar previsões ou classificações. É o mais usado em sistemas de previsão;
- b) aprendizado não supervisionado: geralmente utilizado em problemas de classificação e reconhecimento que envolvem volumes excessivos de dados, difíceis de serem etiquetados. Esses algoritmos segmentam os dados em grupos de exemplos ou grupos de atributos, e criam automaticamente as etiquetas, o que os torna, então, algoritmos de aprendizado supervisionado.

Pode-se somar uma terceira abordagem chamada Aprendizado por Reforço. Nesse caso, não há dados iniciais de treinamento e o sistema adquire conhecimento a partir de tentativa e erro. Os acertos são premiados e representam o único *feedback* recebido. Em geral, são usados quando representam sistemas muito complexos para serem modelados, mas fáceis de serem avaliados. É utilizado na área de jogos eletrônicos, robótica, carros autônomos e controle de processos em geral (AGGARWAL, 2018).

Devido às suas especificidades e ao crescente interesse, pode-se destacar o aprendizado profundo (AP), uma variação aplicável às três abordagens mencionadas, que incorpora redes neurais em camadas sucessivas para aprender com os dados de forma iterativa a partir de uma combinação de aprendizado supervisionado ou não supervisionado, com inspiração no funcionamento do córtex cerebral. O AP é usado em sistemas de reconhecimento de voz, aplicações de internet das coisas (IoT), dentre outros.

A Tabela 2 apresenta alguns dos principais métodos de AM supervisionado utilizados para a previsão de geração solar fotovoltaica determinística – a mais comum.

O modelo linear generalizado (GLM) – ou *generalised linear model* – é um AM baseado na combinação de regressões lineares simples que permite variáveis de saída cuja distribuição dos erros não acompanha uma distribuição normal. Máquina de aumento de gradiente (GBM) – ou *gradient boosting machine* – é um AM

Tabela 2 – Síntese comparativa dos métodos de aprendizado de máquina supervisionado utilizados na previsão de geração de energia solar fotovoltaica

Métodos ^a	Critério					
	Precisão	Velocidade	Linear	Nº de parâmetros	Complexidade	Facilidade de uso
GLM	baixa	baixa	sim	baixo	baixa	sim
GBM	média	baixa	sim/não	baixo	baixa	sim
AD	média	baixa	não	baixo	média	sim
RF	alta	média	não	baixo	média	sim
SVM	alta	alta	não	baixo	média	não
AP/RNA	alta	alta	não	baixo	alta	não

^a Modelo linear generalizado (GLM), máquina de aumento de gradiente (GBM), árvore de decisão (AD), floresta aleatória (RF), máquina de vetores de suporte (SVM), aprendizado profundo (AP), rede neural artificial (RNA).

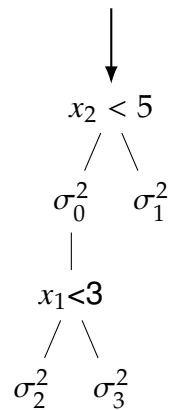
Fonte: adaptado de Jebli, Belouadha e Kabbaj (2020).

altamente eficiente, baseado no princípio do aumento do gradiente, que consiste num meta-algoritmo que combina vários algoritmos fracos, iterados sucessivamente para redução do erro, resultando numa combinação linear desses algoritmos fracos (JEBLI; BELOUADHA; KABB AJ, 2020).

Árvore de decisão (AD) – ou *decision tree* – é um modelo de AM popular em problemas de classificação e regressão que usa os dados de entrada para construir recursivamente uma árvore enraizada na qual o vértice superior é designado como raiz e há um caminho único para cada vértice direcionado para fora da raiz. Começando na raiz com todos os dados em uma classe, é tomada a decisão de dividir os dados em duas subclasses usando o atributo e ponto de corte que resulta no maior ganho de informação, segundo algum critério – como o somatório das variâncias (σ^2) dos erros das amostras resultantes, no caso da regressão (MONCADA; RICHARDSON; VEGA-AVILA, 2018). No exemplo da Figura 12, o atributo x_2 foi o que apresentou um ponto de separação ($x_2=5$) dos exemplares da amostra completa que resulta em duas subamostras de menor somatório das variâncias σ_0^2 e σ_1^2 . A subamostra relativa à variância σ_0^2 apresentou no atributo x_1 o ponto de separação ($x_1=3$) que resulta no menor somatório das variâncias σ_2^2 e σ_3^2 das subamostras resultantes, quando a saída é o valor médio de saída dos elementos de cada subamostra. E assim sucessivamente, para todos ramos da árvore de decisão, até que a redução da variância total para o próximo nível seja menor que um valor predefinido, ou que um número máximo de níveis seja atingido.

A floresta aleatória (RF) – ou *random forest* – é uma ferramenta muito utili-

Figura 12 – Exemplo de uma árvore de decisão



Fonte: elaboração própria (2022).

zada em ciência de dados em aplicações de diversos setores, e consiste num conjunto de ADs independentes construídas a partir de amostras aleatórias da população e visando uma visão fragmentada do problema⁵. Para fugir do sobreajuste, o resultado final – no caso da regressão – é a média dos resultados individuais das árvores criadas, e não o melhor resultado obtido (JEBLI; BELOUADHA; KABBAJ, 2020).

Uma máquina de vetores de suporte (SVM) – ou *support vector machine* – constrói hiperplanos em um espaço n-dimensional para classificação ou regressão de dados. O método utiliza, para os problemas linearmente não separáveis, uma transformação das entradas do espaço original para um espaço mais complexo em que o problema seja linearmente separável. A SVM é um modelo não linear efetivo de IA que comparado com as RNAs possui maior velocidade de treinamento e oferece bom resultado de convergência, livre dos mínimos locais (SHI *et al.*, 2012).

Para uma compreensão mais ampla dos aspectos relacionados ao aprendizado de máquina e às redes neurais artificiais, dois autores são sugeridos. Géron (2019) oferece um ponto de partida para os iniciantes em aprendizado de máquina, pois consegue reunir os conceitos de algoritmos clássicos e modernos com as ferramentas computacionais e exemplos de aplicações. O problema é que não possui um sumário organizado, o que dificulta visitar alguns conceitos quando demandados. Já Aggarwal (2018) foca nas redes neurais artificiais de aprendizado profundo, e possui um conteúdo bem organizado e desenvolvido.

2.3.5 Redes neurais artificiais

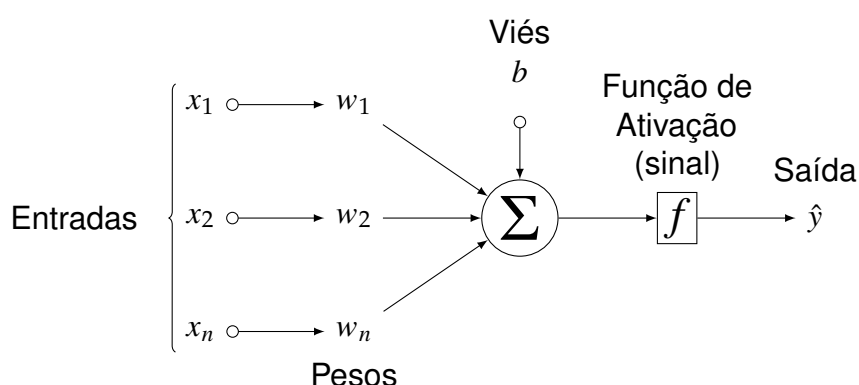
As redes neurais artificiais (RNAs) foram desenvolvidas após o surgimento dos primeiros computadores, com inspiração no sistema nervoso humano, visando

⁵A floresta aleatória é abordada em um comparativo do Estudo de Casos (seção 4.3).

solucionar problemas através da simulação do que se pensava serem os mecanismos de nossos neurônios. Inicialmente a capacidade computacional não permitia grandes avanços, o que levou a um período de pouco interesse até próximo da virada do século, quando uma grande disponibilidade de dados (armazenados) e uma crescente capacidade computacional deu nova vida às RNAs, sob a grife aprendizado profundo (AP) (AGGARWAL, 2018). Russell e Norvig (2016) apresentam o ponto de vista histórico das RNAs, sua origem a partir de 1943 e seu relacionamento com o termo inteligência artificial, as novas técnicas e usos, a adoção do método científico em seu desenvolvimento, novas bases de dados compartilhadas, etc.

A unidade básica de uma RNA tradicional é o *perceptron*. Trata-se de uma das arquiteturas mais simples e baseada na chamada unidade lógica de limiar (TLU), com uma camada de neurônios de entrada $\bar{X}^T = [x_1 \dots x_n]$, uma camada computacional (a TLU) e uma única saída (\hat{y}), podendo ou não contar com um viés adicionado (b), como mostra a Figura 13 (GÉRON, 2019).

Figura 13 – Estrutura básica de rede neural artificial com TLU – o *perceptron*



Fonte: elaboração própria (2022).

O funcionamento da TLU é descrito por:

$$\hat{y} = \text{sinal} \left(\bar{X}^T \bar{W} + b \right) = \text{sinal} \left(\sum_{j=1}^n x_j w_j + b \right), \quad (10)$$

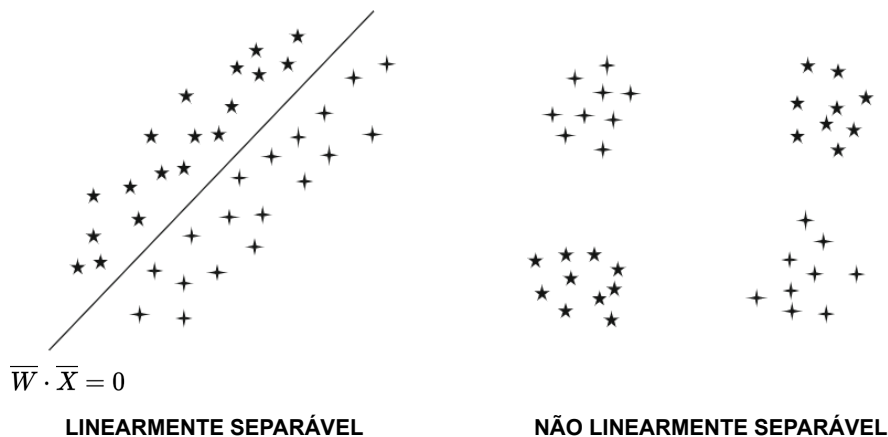
sendo que: $\hat{y} = \{-1, 0, +1\}$ é a previsão computada pela função sinal⁶ (função de ativação); $\bar{X}^T = [x_1 \dots x_n]$ é o conjunto dos valores dos n atributos de entrada; $\bar{W}^T = [w_1 \dots w_n]$ é o conjunto de pesos aplicados sobre os valores dos atributos de entrada; e b é o viés, um valor real adicionado (ou zero) (AGGARWAL, 2018; GÉRON, 2019).

Os dados de entrada para treinamento (\bar{X}) são conhecidos, e os pesos (\bar{W}) e o viés (b) devem ser encontrados através de otimização numérica, embora nos

⁶A função sinal – ou *sign* – retorna 1 (um) para entrada positiva, 0 (zero) se a entrada for zero ou -1 (um negativo).

primórdios tenha sido heurística, por circuitos eletrônicos (*hardware*) (AGGARWAL, 2018). O modelo proposto no *perceptron* é linear, com a Equação 10 definindo um hiperplano. Nesta configuração a rede neural é capaz de solucionar apenas problemas de classificação binários simples, linearmente separáveis, como ilustra a Figura 14, pois seus limiares de decisão possuem limites lineares (GÉRON, 2019).

Figura 14 – Exemplos de classes linearmente separáveis e inseparáveis



Fonte: adaptado de Aggarwal (2018).

Vários incrementos de complexidade foram realizados para solucionar novos problemas, com menores recursos computacionais ou com número reduzido de exemplos de treinamento. Dentre as inovações ao modelo apresentado, pode-se citar:

- a) novas funções de ativação⁷, para possibilitar a solução de variáveis reais, aumentar a não linearidade do modelo (sigmoide, tangente hiperbólica, etc.), aumentar a tolerância aos ruídos (CLEVERT; UNTERTHINER; HOCHREITER, 2016) ou apenas delimitar o valor de saída (softmax, ReLU etc.);
- b) paralelização de neurônios na camada oculta, para aumentar a capacidade de aprendizagem do modelo;
- c) serialização de camadas ocultas, ou *perceptron* multicamadas (MLP), para aumento da representatividade do modelo e da capacidade e velocidade de aprendizagem (AGGARWAL, 2018);
- d) utilização de funções de otimização mais especializadas para aumentar a chance de aprendizagem;
- e) desenvolvimento de métodos de otimização mais especializados (para solução de novos problemas) ou mais genéricos (para garantir o sucesso

⁷As funções de ativação são tratadas na seção A.3.

em um número maior de tipos de problemas), com vistas ao aumento da velocidade de treinamento;

- f) novas formas de conexão entre os neurônios das várias camadas para solução de problemas específicos, em oposição às chamadas redes densas, onde todos os elementos de uma camada estão conectados aos elementos da camada seguinte;
- g) conexões recorrentes, ou seja, saídas retroalimentando alguns neurônios de entrada ou intermediários;
- h) funções de otimização especializadas;
- i) soluções empacotadas, ou seja, várias alternativas integradas na solução de um problema (duas ou mais estruturas RNA conectadas, saídas derivadas da média entre mais de um modelo, etc.); etc.

O Anexo A apresenta os diagramas das principais estruturas RNA já criadas, extraídos de Veen e Leijnen (2019), os quais também apresentam uma referência rápida de todas elas, inclusive com citação dos artigos de origem. Algumas das estruturas mais comuns na área de geração solar fotovoltaica são: *perceptron* multicamadas, rede neural convolucional e rede neural recorrente (MASSAOUDI *et al.*, 2021).

O *perceptron* multicamadas (MLP) é uma rede neural básica de AP e supervisionado, do tipo *feedforward* – ou seja, cada camada se conecta à próxima camada, da entrada para a saída, numa única direção – e densa – ou seja, com todos os neurônios de uma camada conectados a todos os neurônios da camada anterior –, que aplica o método chamado *back propagation* para minimização da função de custo. Possui boa capacidade de generalização de problemas complexos, mas com limitação na representação de relacionamentos intertemporais entre os dados, pois não salva informações prévias. Essa estrutura, de interesse nesse trabalho, é detalhada no Apêndice A, assim como alguns dos principais aspectos relacionados à sua operacionalização, como: métodos de otimização, funções de custo, funções de ativação, critérios de parada do treinamento, métodos de regularização e normalização de variáveis de entrada.

A rede neural convolucional (CNN) (LECUN *et al.*, 1998) também é do tipo *feedforward*, composta de três elementos principais – convolução, computação de agrupamento e camadas densas. A convolução possibilita a extração de atributos provenientes de dados contíguos, enquanto a computação de agrupamento visa uma transformação que agrupa e unifica atributos similares, e as camadas densas finais organizam e resumem a saída. Essa rede neural trouxe uma grande evolução ao processamento de imagens (YALÇIN, 2021). Na área de geração solar FV são usadas para lidar com dados sequenciais (MASSAOUDI *et al.*, 2021).

Já a rede neural recorrente (RNN) (RUMELHART; HINTON; WILLIAMS,

1986) é um caso particular de rede neural de aprendizado profundo baseada na teoria de controle, composta de uma rede de neurônios cujas saídas são conectadas não apenas à próxima camada mas também a um ponto de retroalimentação. Em sua estrutura básica é considerada limitada, pois não é capaz de manter memórias de longo prazo. Assim, foram desenvolvidas arquiteturas especializadas para este fim, como a chamada rede de memória de curto prazo longa (LSTM) (HOCHREITER; SCHMIDHUBER, 1997). As redes recorrentes são especialmente úteis para previsões relacionadas às séries temporais.

2.4 Conclusão parcial

Nesse capítulo foi apresentado o contexto em que a previsão de geração FV proposta se insere – o das pequenas usinas reunidas em uma microrrede ou VPP para ganharem escala e previsibilidade para participarem do mercado de energia de forma competitiva. Evidenciou-se o papel que as previsões de qualidade têm para o sucesso desse modelo de operação.

Os principais aspectos físicos do geoposicionamento solar foram apresentados e relacionados com as condições atmosféricas, assim como foram conceituadas algumas das medidas que são aproveitadas como entradas do método proposto: ângulo de elevação solar, ângulo azimute, ângulo horário solar e ângulo diário solar.

Por fim, os aspectos mais importantes para a seleção dos métodos de previsão de geração FV são abordados e contextualizados na área do AM. São apresentadas as principais estruturas RNAs utilizadas para o tipo de problema, destacando-se o detalhamento no Apêndice A da estrutura de interesse desse trabalho, o MLP, porta de entrada para o universo das RNAs e um legítimo representante do AP.

3 PREVISÃO DE GERAÇÃO SOLAR FOTOVOLTAICA

O método proposto para a previsão de geração solar fotovoltaica possui três predefinições: o objetivo da previsão, a técnica adotada e a disponibilidade de dados. Cada um desses aspectos interfere nas decisões tomadas durante o desenvolvimento do método e vale a pena detalhá-los.

Iniciando-se pelo objetivo, a previsão proposta visa dotar os sistemas de gerenciamento econômico das plantas virtuais de energia de prognósticos de geração das usinas fotovoltaicas participantes do *pool*, destinados ao planejamento da próxima hora ao dia seguinte, em base horária¹, de forma contínua. Esses dados complementam outros esquemas de previsão de horizontes maiores (dias, meses ou anos) ou menores (semi-horários, quartos de hora, minutos etc.).

Quanto à técnica adotada, as redes neurais artificiais são cada vez mais utilizadas em diversas aplicações, no entanto, há uma longa curva de aprendizado e alguma incerteza entre os potenciais usuários quanto ao custo-benefício de sua utilização. Mas a cada problema investigado, vencidos ou não os desafios estabelecidos, avança-se – acredita-se – no sentido da generalização de seu uso, em função da experiência adquirida. Assim, a aplicação do MLP, um dos objetivos centrais deste trabalho, independe da oportunidade de outras técnicas.

Os dados históricos disponíveis para a aplicação de técnicas de inteligência artificial são a geração de energia com etiqueta temporal e dados meteorológicos padronizados. Os dados de geração de energia obtidos para o desenvolvimento deste trabalho são horários. Para a obtenção de dados de medição em fluxo ininterrupto das usinas durante a aplicação do método em produção, a forma mais simplificada – sem a necessidade de investimentos extras – parece ser através de consulta direta via Plataforma de Integração da CCEE (CCEE, 2022c), também com discretização horária. Alguns serviços especializados de previsão das condições do tempo oferecem – em geral – dados instantâneos semi-horários ou horários da temperatura ambiente, umidade, pressão atmosférica, velocidade e direção dos ventos, índice pluviométrico, cobertura de nuvens etc. Desta forma, assume-se a discretização horária, sabendo-se que intervalos menores, se houvesse a disponibilidade de dados, poderiam trazer melhores resultados (DUCHAUD *et al.*, 2020).

¹O ACL do mercado de energia do Brasil opera com discretização horária.

3.1 Bases do método proposto de previsão fotovoltaica

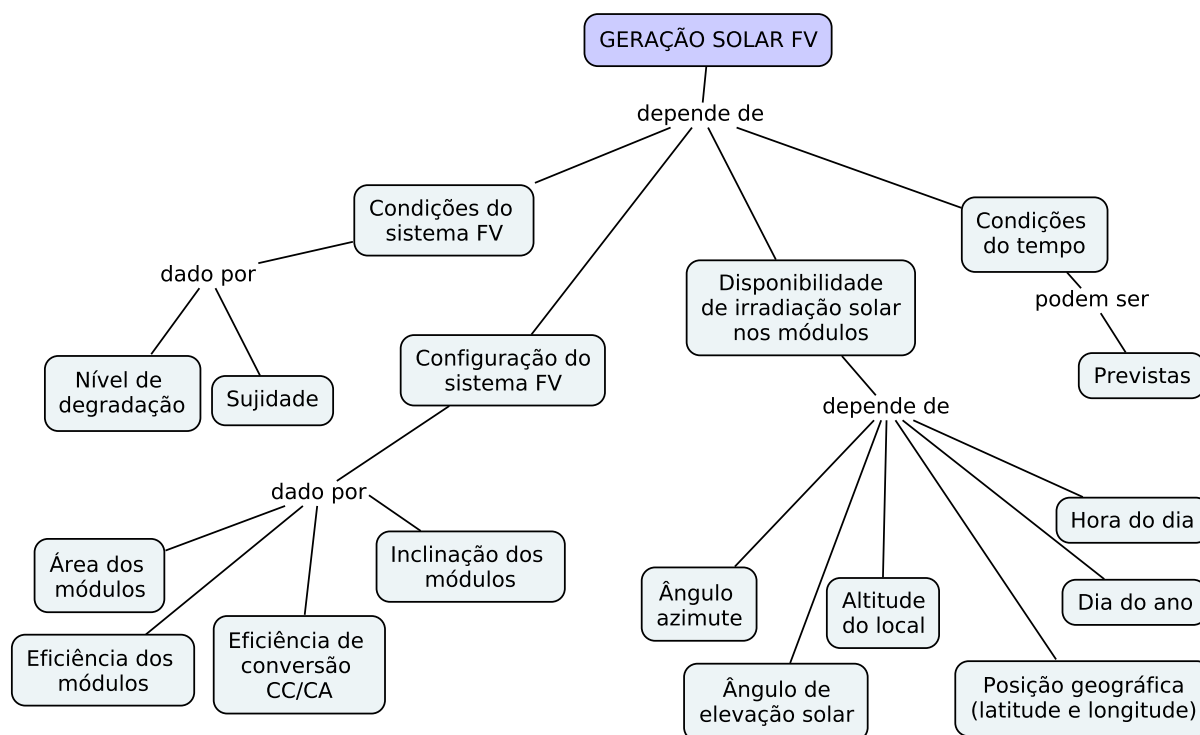
O método de previsão de geração solar fotovoltaica desenvolvido para as VPPs parte do pressuposto de que o nível de produção de energia de um sistema FV responde aos aspectos relacionados a quatro principais direcionadores, ilustrados pelo mapa mental da Figura 15. São eles:

- a) disponibilidade de irradiação solar: que, devido à característica cíclica do posicionamento entre a Terra e o Sol, varia principalmente com a hora do dia, dia do ano e a posição geográfica na Terra (latitude, longitude e altitude);
- b) condições do tempo: que reduzem ou ampliam os efeitos da irradiação solar disponível, pois interferem tanto na eficiência dos painéis FV e da eletrônica de potência envolvida (temperatura ambiente, por exemplo), quanto como um filtro da irradiação solar em função da cobertura de nuvens, além de participarem em outros efeitos, como a reflexão;
- c) configuração do sistema FV: depende de uma série de características próprias de cada instalação, uma combinação de:
 - i) tecnologia dos painéis solares (curva de eficiência de conversão);
 - ii) quantidade (área instalada) e esquema de ligação dos painéis solares (com maiores ganhos ou perdas do arranjo);
 - iii) capacidade e tecnologias presentes nos conversores (MPPT, eletrônica de potência no nível do módulo (MLPE) etc.);
 - iv) inclinação e azimute dos painéis solares da instalação (ou função de seguimento solar);
- d) condições do sistema FV: no sentido do nível de degradação dos equipamentos envolvidos e outros defeitos técnicos ou de instalação (conexões elétricas, sombras, sujidade etc.).

É fato que algumas dessas condições que afetam o resultado final da geração solar fotovoltaica de uma usina, não variam com o tempo. Para uma suposição de causa para o efeito da variação da geração de energia elétrica de uma usina FV, ao longo de um período, pode-se descartar a configuração do sistema FV. Ele afeta a saída do sistema, mas é um parâmetro constante, incapaz de contribuir para a explicação das variações da geração da usina ao longo do tempo. Pode-se dizer o mesmo da altitude local. Embora este dado interfira no posicionamento solar em relação à usina, e afete o resultado cotidiano do sistema FV, não varia com o tempo, portanto não explica as variações horárias da geração solar fotovoltaica de uma usina.

As condições do sistema FV costumam ser difíceis de monitorar. Talvez o

Figura 15 – Principais variáveis do problema da geração solar fotovoltaica



Fonte: elaboração própria (2022).

dado mais objetivo seja o tempo de vida dos equipamentos, mas, ainda assim, seus componentes possuem diferentes vidas úteis e uma usina pode passar por manutenção preventiva ou corretiva, tornando o tempo de vida dos equipamentos híbrido, o que, do ponto de vista de um sistema gerenciador de plantas virtuais, é difícil de generalizar entre as aplicações. Quanto à sujidade dos painéis solares, poder-se-ia acompanhar o tempo decorrido desde sua última limpeza e/ou desde a última chuva (média a forte) (APPELS *et al.*, 2012), dependendo da inclinação desses painéis. Mas esta informação não está disponível para ser considerada neste trabalho.

Sobram os efeitos da disponibilidade de irradiação solar e das condições do tempo. Ambos são interrelacionados, pois muitos dos fatores climáticos são sazonais, em função dos ciclos solares, em última análise. A disponibilidade de irradiação solar pode ser simplificada representada pela posição solar (elevação e azimute) em relação à usina, calculada a partir de modelos estabelecidos (NREL, 2021b). É verdade que não difere em última análise da simples anotação dos ângulos diário e horário solares. Mas se pressupõe que as variáveis de posicionamento solar podem contribuir para a aceleração da convergência ou redução do tamanho da amostra mínima necessária para uma mesma exatidão da previsão.

Finalmente, os parâmetros das condições do tempo podem ser obtidos diretamente de empresas e serviços especializados em previsões meteorológicas,

o que conta com longa tradição e esforços permanentes de melhoria, visto que as previsões são necessárias para o sucesso de algumas atividades econômicas e sociais, da agricultura e produção de energia à prevenção de desastres.

Este método de previsão de geração solar fotovoltaico é voltado para o curto prazo, objetivando o planejamento intradiário (ou em inglês, *intraday*) e do dia seguinte (ou em inglês, *day-ahead*) – também referido por D+1 – com discretização horária. Esses prazos estão relacionados com o fechamento do balanço energético diário no ACL, o qual combinado com a previsão do PLD horário, geram oportunidades de resultados econômicos aos proprietários dos REDs, seja através da compra e venda de energia no mercado de curto prazo ou da participação em programas de resposta da demanda (EMARATI; KEYNIA; RASHIDINEJAD, 2020; TAN *et al.*, 2021), por exemplo. O processo visa atender novas usinas sem que seja necessário alterar sua estrutura fundamental.

3.2 Prospecção de dados para o desenvolvimento da previsão de geração

Três usinas de diferentes capacidades e instaladas em regiões distintas foram selecionadas para referenciar a construção do modelo e a verificação dos resultados. Elas são referidas por AUS (AUSGRID, 2014), BRA e EUA (NREL, 2021a) de acordo com a sigla do país da instalação e são classificáveis como microgeração (AUS) ou minigeração distribuída (BRA e EUA) nos moldes da legislação brasileira (BRASIL, 2022). Privilegiou-se a utilização de dados abertos (exceção da usina BRA) e de ampla abrangência geográfica para maior atratividade deste trabalho. A Tabela 3 resume as características dessas usinas, detalhadas no Apêndice D.

Tabela 3 – Usinas de referência do trabalho

Característica	AUS ^a	BRA	EUA
Tipo	Residencial	Comercial ^b	Comercial ^b
Cidade/Estado	The Entrance North/NSW	Bauru/SP	Cherry Hill/NJ
Latitude	-33,36°	-22,2886°	39,9283°
Altitude	25 m	600 m	18 m
Potência Nominal	2,04 kWp	95 kWp	140,14 kWp
Fator de Capacidade	27,9%	32,0%	25,1%

^a A localização da usina AUS é estimada por um ponto central arbitrário da região de The Entrance North em New South Wales (Austrália), pois a localização exata da residência não é informada.

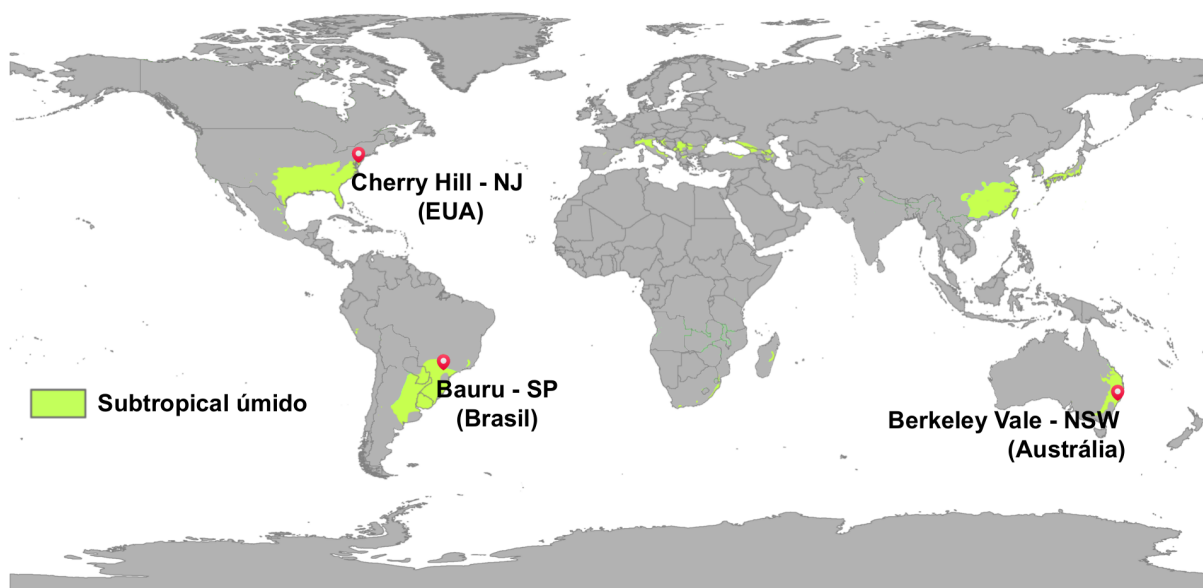
^b As usinas são utilizadas para consumo próprio dos prédios comerciais.

Fonte: elaboração própria (2022).

Embora as usinas tenham sido selecionadas em função da disponibilidade de dados, coincidentemente, elas pertencem a regiões de mesma classificação do

clima: subtropical úmido, como mostra a Figura 16. O clima subtropical úmido, ou Cfa na escala de Köppen-Geiger, é temperado, sem estação seca e com o verão quente (BECK *et al.*, 2018). Essa informação é apenas acessória e não influenciou as etapas de desenvolvimento apresentadas neste trabalho, mas cabe observar que algumas semelhanças entre os melhores parâmetros encontrados para a solução do problema nas três usinas podem ter relação com essa coincidência climática.

Figura 16 – Regiões de clima subtropical úmido



Fonte: Beck *et al.* (2018).

As informações necessárias para o desenvolvimento e a aplicação do modelo de aprendizado de máquina envolvem dados históricos e prognósticos futuros, o que restringe o arcabouço de soluções. Os atributos com histórico de acompanhamento, mas sem uma rotina de previsão própria, ou são descartados, ou são inferidos pelo próprio método, através do tratamento de séries temporais, por exemplo. Os dados meteorológicos provenientes das estações METAR (World Meteorological Organization, 2018) foram cogitados por serem muito apropriados para o acompanhamento mais refinado da cobertura de nuvens – camadas em diferentes altitudes – e da visibilidade (em distância). No entanto, não há prognóstico futuro desses atributos, o que impediria a aplicação do tipo de rede neural desejada (MLP). Com isso em vista, de acordo com a disponibilidade, foram definidos os seguintes dados para o desenvolvimento:

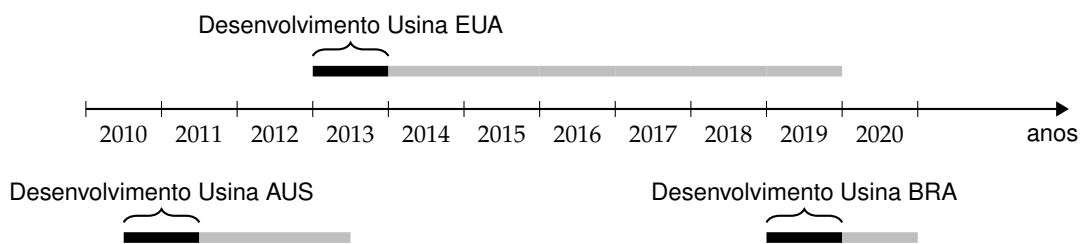
- a) potência nominal da usina (kWp);
- b) informações de georreferenciamento (latitude, longitude e altitude);
- c) histórico de medição de geração (kWh), com discretização horária;
- d) histórico de condições meteorológicas instantâneas locais em intervalos horários, envolvendo: temperatura ambiente, umidade relativa do ar, co-

bertura de nuvens, pressão atmosférica, velocidade dos ventos, direção dos ventos, índice de chuvas e de neve;

- e) previsão meteorológica, em termos dos mesmos atributos do Item d);
- f) ferramenta de cálculo das métricas de geoposicionamento solar, envolvendo: ângulo diário solar, ângulo horário solar, ângulo de elevação solar e azimute.

Os dados de localização das usinas, potência nominal e o histórico de medição – com etiqueta temporal: ano, mês, dia e hora – são triviais a qualquer sistema de gerenciamento energético, como os sistemas de gerenciamento das plantas virtuais. No caso do presente trabalho, a disponibilidade de dados de medição é diferente para cada usina, como mostra a Figura 17. A usina EUA possui alguns dados faltantes ao longo dos sete anos, mas distribuídos e pouco representativos (0,2%). No caso da usina BRA há dois períodos prolongados de interrupção – 8 dias em maio de 2019 e 15 dias em janeiro de 2020 – nos dois anos de dados. Os dados da usina AUS estão íntegros no período de três anos.

Figura 17 – Cronograma da amostra de desenvolvimento frente à amostra completa



Fonte: elaboração própria (2022).

Para uma melhor dedução e a comparabilidade dos resultados de usinas diferentes ao longo do desenvolvimento, optou-se pela conversão dos dados absolutos de geração de energia (S_h) para o sistema por unidade (p.u.),

$$S_h \text{ (p.u.)} = \frac{S_h \text{ (kWh/h)}}{S_{base} \text{ (kWp)}}, \quad (11)$$

sendo S_{base} definido pela capacidade instalada de geração de cada usina, como mostra a Tabela 4.

Em nenhuma das usinas foram observados *outliers* evidentes, ou seja, pontos fora da curva resultantes de medições incorretas ou suspeitas. Mas se sabe que os dados meteorológicos carregam alto grau de heterocedasticidade, ou seja, a variância dos erros não é a mesma em todas as observações feitas. Esses *outliers* inevitáveis podem impactar o resultado da previsão. Embora alguns estudos tenham

Tabela 4 – Base do sistema por p.u. de cada usina

Usina	Potência (kWp)
AUS	2,040
BRA	95,000
EUA	140,140

aplicado métodos para remoção desses dados, a seleção de um ponto de corte ainda é uma discussão em aberto (SHENG *et al.*, 2018).

A participação das observações em que a geração foi menor ou igual a zero foi em torno de 5% a 10% das horas claras, como mostra a Tabela 5. Isso inclui a primeira ou última hora dos dias em que a irradiação disponível para os painéis solares é insuficiente para a geração acima das perdas do inversor, além de eventuais momentos de extrema nebulosidade ou sombra ao longo do dia. Os dados disponíveis da usina AUS apresentam apenas valores nulos ou positivos, pois foram tratados pela Ausgrid (AUSGRID, 2014).

Tabela 5 – Percentual de horas sem geração de energia na amostra de desenvolvimento

Usina	Menor que zero	Igual a zero
AUS	0,0%	4,8%
BRA	7,9%	0,0%
EUA	9,4%	0,2%

Fonte: elaboração própria (2022).

O histórico das condições de tempo dos locais das usinas, bem como as previsões meteorológicas, podem ser obtidos via interface de programação de aplicativos (API) de um fornecedor deste tipo de serviço. O Apêndice C apresenta a alternativa utilizada neste trabalho (OPENWEATHER, 2020). No entanto, como os dados são provenientes de registros instantâneos realizados em cada hora cheia (com base em UTC), optou-se por tratar cada atributo horário como sendo o par de valores de medição inicial e final de cada hora. Apesar de a integração desses atributos meteorológicos não estar disponível – não só em histórico, mas como prognóstico –, acredita-se que o par de dados possui uma informação adicional: o sentido de variação no período.

Já as métricas de geoposicionamento solar podem ser calculadas. Optou-se pela utilização da biblioteca especializada *Solar Position and Intensity Calculator* (SOLPOS), disponibilizada pelo *National Renewable Energy Laboratory* (NREL)². Ela

²O NREL é um laboratório do Departamento de Energia dos Estados Unidos criado para fomentar o

contém funções em linguagem C que visam calcular o posicionamento solar aparente e sua intensidade com base na data, hora e local da Terra. A documentação contém as referências bibliográficas que justificam cada opção técnica escolhida ao longo dos cálculos (NREL, 2021b). Como as saídas são instantâneas, foi feita a integração numérica de 60 minutos por hora para construir o valor médio horário de cada métrica.

O dia e hora locais são utilizados como etiquetas temporais utilizadas para sincronização entre as várias fontes, mas não são considerados como entradas para o modelo pois não têm sentido físico, podendo diferenciar-se pela opção política de fuso horário e pela longitude em relação à do início do fuso. Dessa forma poderiam causar alguma dificuldade para a uniformização dos ajustes do modelo para todas usinas.

Os atributos de direção dos ventos e do azimute tiveram suas escalas de referência alteradas simultaneamente para uniformização do sentido, independentemente da latitude, conforme a transformação:

$$T(x) = \begin{cases} (450 - x) \bmod 360, & \text{latitude} \leq 0^\circ \\ (270 + x) \bmod 360, & \text{latitude} > 0^\circ \end{cases}, \quad (12)$$

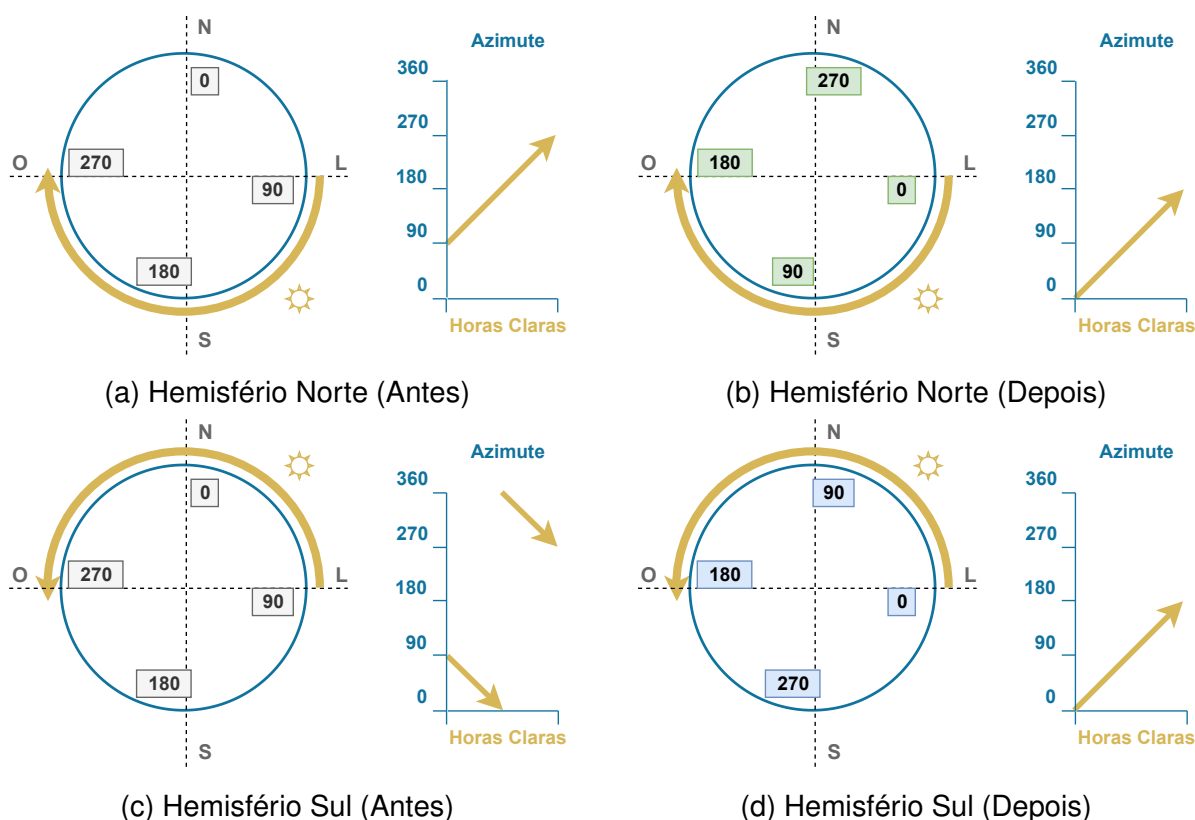
para que do nascer ao pôr do Sol se tenha um mesmo sentido de avanço das escalas em ambos hemisférios. Com isso, espera-se, mais uma vez, aproximar as estruturas neurais envolvidas na solução das diferentes usinas, no que diz respeito à convergência entre os ajustes dos modelos. A Figura 18 ilustra o problema e a solução.

Antes da transformação – Figuras 18a e 18c – o avanço do azimute ao longo das horas claras do dia segue curvas diferentes para um mesmo comportamento no hemisfério oposto. Após a transformação – Figuras 18b e 18d – o avanço é similar.

Finalmente, sabe-se que a ocorrência de neve tem um efeito positivo sobre o albedo local, embora possa encobrir temporariamente os painéis solares, ao passo que a chuva tem o efeito oposto, reduz o albedo e limpa os painéis solares. No entanto, chuva, granizo e neve não são facilmente identificados pelos sistemas de monitoramento das estações meteorológicas, que muitas vezes, utilizam modelos para desambiguação (DING *et al.*, 2014). Embora os índices de chuva e neve apareçam simultaneamente para uma mesma hora nos registros meteorológicos utilizados para a usina EUA, são pouco críveis. Dessa forma, o atributo do índice de neves foi incorporado ao atributo do índice de chuvas; ambos possuem a mesma unidade de medida – mm (estado líquido). Portanto, o modelo de previsão adotado deve inferir sobre o significado diferenciado (não linear) deste dado em diferentes condições meteorológicas e épocas do ano. Evita-se a manutenção de diferentes quantidades de atributos de entrada, visto que

avanço da ciência e engenharia nas áreas da eficiência energética, sustentabilidade da distribuição e tecnologias de energia renovável, além da provisão de conhecimentos para integração e otimização dos sistemas de energia.

Figura 18 – Transformação da escala do azimute por hemisfério



Fonte: elaboração própria (2022).

para algumas usinas o valor relativo ao índice de neves é constante – sempre zerado – e a ocorrência dos dois fenômenos é relativamente pequena.

A lista completa dos atributos para a previsão de geração FV proposta é apresentada pela Tabela 6. As abreviações são padronizadas em inglês para manter a compatibilidade com as fontes de origem dos dados.

A única saída do modelo será a previsão da geração de energia no sistema por p.u. para cada período de discretização (Tabela 7).

3.2.1 Sincronização dos dados de diferentes fontes

Uma etapa exaustiva da preparação dos dados de entrada foi a compreensão dos sistemas de referência de datas utilizados pelas várias fontes e a sincronização entre estes.

As informações da OpenWeather (OPENWEATHER, 2020) – fornecedora dos dados meteorológicos – possuem referência ao horário UTC em qualquer coordenada geográfica consultada, e os valores resultantes do cálculo de posicionamento e intensidade solares têm as coordenadas geográficas e o fuso-horário como variáveis de entrada, e resultam em dados em hora local ou em UTC. Desta forma, são facilmente adaptáveis a qualquer algoritmo computacional de uso contínuo.

Tabela 6 – Lista dos atributos de entrada do modelo

Atributo	Unidade	Abreviação
Ângulo diário solar	graus (0 a 359)	<i>dayang</i>
Ângulo horário solar	graus (-180 a 180)	<i>hrang</i>
Ângulo de elevação solar	graus (horizonte=0)	<i>elevref</i>
Azimute ^a	graus (Leste=0)	<i>azim</i>
Temperatura ambiente	°C	<i>temp</i>
Umidade relativa do ar	%	<i>humidity</i>
Cobertura de nuvens	%	<i>clouds</i>
Pressão atmosférica	hPa	<i>pressure</i>
Velocidade dos ventos	m/s	<i>wind_speed</i>
Direção dos ventos ^a	graus (Leste=0)	<i>wind_deg</i>
Índice de chuvas (1 hora)	mm	<i>rain_1h</i>

^a Versão transformada. O crescimento é horário para o hemisfério Norte e anti-horário para o hemisfério Sul.

Fonte: elaboração própria (2022).

Tabela 7 – Saída do modelo

Descrição	Unidade	Abreviação
Geração de energia ^a	p.u.	<i>gen</i>

^a A base é a potência nominal das usinas.

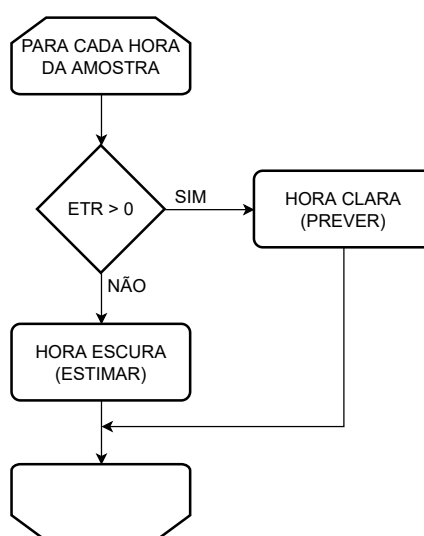
Fonte: elaboração própria (2022).

Já os dados de medição de geração de energia, os quais seriam facilmente obtidos pelo sistema de gerenciamento da VPP, no caso do estudo, foram inconvenientes. A Ausgrid – fornecedora dos dados da usina AUS – e a AES Brasil – fornecedora dos dados da usina BRA – utilizam a etiqueta de horário local com a alteração do horário de verão. Foi preciso verificar o histórico de transições entre o horário padrão e de verão nas localidades envolvidas e acrescentar a etiqueta do horário UTC a cada linha de dados para garantia da sincronização com as variáveis meteorológicas. No caso da NREL, o fuso horário é supostamente padronizado no UTC, conforme especificação da API, o que não se verifica na prática. Foi necessário um estudo e ajuste manual dos dados utilizados, já que não estavam no fuso horário local e nem em UTC. Pode ser que o relógio de algum sistema utilizado no processo de medição ou armazenamento dos dados estivesse desajustado.

3.2.2 Pré-processamento de entradas

O conjunto original de dados de medição de geração de energia das usinas contém registros para todas as horas do dia. Mas as horas escuras não produzem informações relevantes para o método, diluem as métricas utilizadas, ocupam tempo de processamento, e podem interferir negativamente no índice de acerto das horas claras. Assim, os registros da amostra passam por uma triagem, como mostra a Figura 19, em que apenas as horas claras do dia são previstas com base no modelo RNA. As horas escuras são estimadas³ como sendo a média histórica das horas escuras.

Figura 19 – Fluxo de triagem das horas claras



Fonte: elaboração própria (2022).

A definição de hora clara e hora escura é feita com base no resultado da integração numérica (60 minutos em cada hora) da ETR relativa ao local de cada usina a cada hora⁴, mas não participa das entradas do modelo, apenas da triagem das horas claras.

3.2.3 Estatística dos dados de desenvolvimento

Nesta subseção são apresentadas algumas métricas estatísticas dos atributos da amostra de desenvolvimento de cada usina, bem como algumas diferenças e similaridades entre elas. Essas informações são relevantes porque a modelagem é baseada nesse conjunto de dados e, como visto, são os dados que definem o modelo de aprendizado de máquina.

As amostras, relativas a 12 meses contíguos de cada usina para que ao menos um ciclo solar anual esteja completamente considerado, somam um total de

³Previsão é um tipo de estimativa que envolve incertezas, ao passo, que a estimativa em si, decorre de efeitos fixos.

⁴A ETR é um dos parâmetros calculáveis via SOLPOS (NREL, 2021b).

14.252 registros relativos ao mesmo número de horas claras para 12 meses de dados, sendo: 4.779 registros da usina AUS; 4.680 registros da usina BRA; 4.793 registros da usina EUA. A ETR é calculada pelo mesmo algoritmo utilizado para o cálculo dos demais atributos solares (NREL, 2021b), mas não é utilizada como atributo de entrada do modelo por ser muito correlacionada – matematicamente – com os demais atributos. Locais de latitudes mais distantes da linha do equador possuem um número maior de horas claras no verão e menor no inverno, em função dos ciclos solares.

A Tabela 8 apresenta as principais características dos dados do período de desenvolvimento (Figura 17) para as três usinas, com notação flexibilizada para ajustar-se aos dados de origem ou à facilidade de leitura.

Tabela 8 – Estatística dos atributos das usinas para o desenvolvimento

Atributo	Usina	Média	DP	Mín	25%	50%	75%	Máx
<i>dayang</i>	AUS	180,9	109,3	0	80	185	281	359
	BRA	181,1	108,0	0	82	186	277	359
	EUA	176,5	96,7	0	98	175	254	359
<i>hrang</i>	AUS	-0,5	58,0	-113,2	-50,7	-2,9	51,4	114,6
	BRA	-0,8	57,3	-105,1	-52,8	1,7	48,1	108,1
	EUA	-0,1	58,9	-114,3	-51,8	3,8	50,9	113,2
<i>elevref</i>	AUS	30,4	21,8	-5,9	13,3	30,0	45,6	78,5
	BRA	34,4	24,2	-6,2	13,9	34,6	52,2	86,4
	EUA	28,2	20,4	-5,7	11,0	26,7	42,8	72,2
<i>azim</i>	AUS	133,4	101,3	0,0	42,1	131,3	177,1	360,0
	BRA	153,2	107,7	0,0	47,8	155,8	194,5	360,0
	EUA	129,1	95,5	0,1	51,0	121,8	172,8	360,0
<i>temp</i>	AUS	19,21	4,43	6,87	16,04	19,22	22,45	36,77
	BRA	25,33	5,35	4,64	21,58	25,61	29,28	37,77
	EUA	15,06	10,05	-11,19	6,83	16,28	23,49	35,14
<i>humidity</i>	AUS	72,3	14,5	21	64	74	83	98
	BRA	68,2	21,5	15	51	69	88	100
	EUA	58,2	19,8	13	42	56	75	100
<i>clouds</i>	AUS	53,0	37,1	0	17	54	92	100
	BRA	25,6	32,7	0	0	1	40	100
	EUA	59,6	28,5	0	40	75	90	100
<i>pressure</i>	AUS	1016,9	7,2	989	1012	1016	1022	1038
	BRA	1015,9	4,1	997	1013	1015	1018	1030
	EUA	1017,9	7,2	992	1013	1018	1023	1041
<i>wind_speed</i>	AUS	5,0	3,0	0,3	2,6	4,1	6,7	19,5
	BRA	3,2	1,6	0,1	2,1	3,1	4,1	11,8
	EUA	4,3	2,2	0,14	2,6	4,1	5,7	14,4
<i>wind_deg</i>	AUS	171,2	103,8	0	60	175	260	358
	BRA	140,9	113,3	0	40	120	230	359
	EUA	183,1	88,2	0	120	180	240	357
<i>rain_1h</i>	AUS	0,1	0,4	0	0	0	0	6,17
	BRA	0,1	0,4	0	0	0	0	8,00
	EUA	0,3	1,5	0	0	0	0	46,99

Fonte: elaboração própria (2022).

Alguns pontos chamam a atenção:

- a elevação solar média (*elevref*) está de acordo com a latitude e o fator de capacidade das usinas: AUS (27,9%), BRA (32,0%) e EUA (25,1%); visto que a produção é maior próximo à linha do equador;
- a temperatura média em Bauru (BRA) é mais elevada, mas as máximas são parecidas. A mínima bem abaixo de 0° e o desvio-padrão elevado em Cherry Hill (EUA) são notáveis em relação às outras duas usinas;
- a cobertura de nuvens (*clouds*) em Bauru (BRA) é muito menor que nas demais usinas, com 75% do tempo abaixo de 40%, ao passo que em Cherry Hill (EUA) apenas 25% do tempo está abaixo de 40%;
- o índice médio de chuva⁵ na área da usina EUA é bem superior em média, inclusive com pico bem maior que as demais. Em pelo menos 75% do tempo não chove ou neva nas usinas⁶.

Alguns dos atributos das usinas possuem alta correlação linear (de Pearson) com outros deles, como mostra a Figura 20, relativa à usina BRA, por exemplo.

Figura 20 – Matriz de colinearidade entre os atributos da usina BRA (2019)

dayang		-0.01	0.03	0.08	-0.06	-0.09	-0.10	-0.09	0.15	-0.09	-0.02
hrang	-0.01		0.02	0.05	0.64	-0.60	0.12	-0.34	-0.09	0.22	0.07
elevref	0.03	0.02		-0.05	0.37	-0.19	0.17	-0.02	0.04	-0.05	-0.03
azim	0.08	0.05	-0.05		0.20	0.05	0.14	-0.35	-0.08	0.06	0.04
temp	-0.06	0.64	0.37	0.20		-0.71	0.03	-0.58	-0.18	0.17	-0.06
humidity	-0.09	-0.60	-0.19	0.05	-0.71		0.29	0.15	0.06	-0.07	0.18
clouds	-0.10	0.12	0.17	0.14	0.03	0.29		-0.26	-0.20	0.13	0.22
pressure	-0.09	-0.34	-0.02	-0.35	-0.58	0.15	-0.26		0.22	-0.23	-0.09
wind_speed	0.15	-0.09	0.04	-0.08	-0.18	0.06	-0.20	0.22		-0.14	-0.03
wind_deg	-0.09	0.22	-0.05	0.06	0.17	-0.07	0.13	-0.23	-0.14		0.05
rain_1h	-0.02	0.07	-0.03	0.04	-0.06	0.18	0.22	-0.09	-0.03	0.05	
	dayang	hrang	elevref	azim	temp	humidity	clouds	pressure	wind_speed	wind_deg	rain_1h

Fonte: elaboração própria (2022).

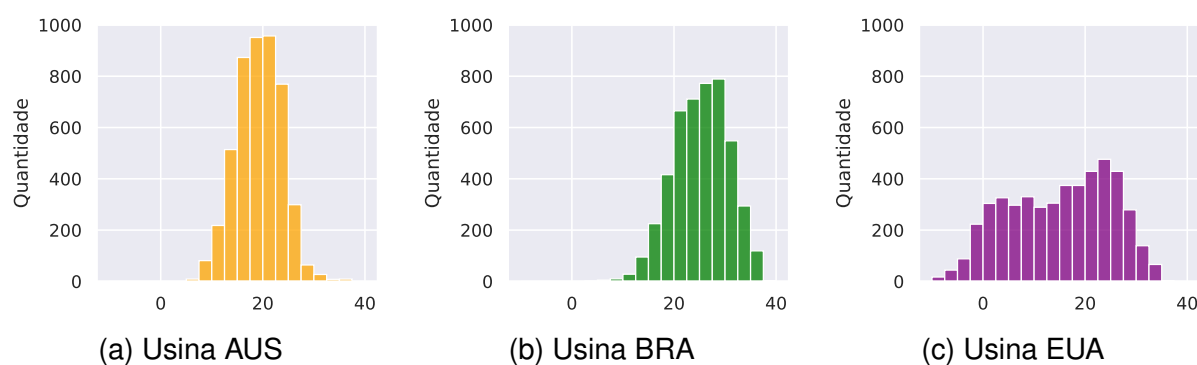
⁵Inclui o índice de neve no caso da Usina EUA, porém com impacto desprezível.

⁶Mais precisamente, chove em 5,7% (AUS), 5,8% (BRA) e 15,6% (EUA) das horas da amostra de desenvolvimento. Para a usina EUA este percentual agrega 1,8% de horas sob neve.

A colinearidade entre os valores dos atributos e a multicolinearidade – quando as correlações envolvem combinações dos atributos – podem prejudicar uma solução clássica baseada em regressão linear, com aumento significativo da variância, mas tem pouco efeito sobre as redes neurais artificiais. O número elevado de parâmetros da RNA de AP, e as operações realizadas entre as camadas torna a saída final uma combinação de muitas funções não lineares envolvendo interações de alta ordem dos preditores originais. Assim, as redes neurais se protegem contra os problemas da multicolinearidade em detrimento da interpretabilidade (DE VEAUX; UNGAR, 1994).

O Apêndice E contém os histogramas de todos atributos utilizados. É notável a semelhança entre as usinas das suas distribuições relacionadas com os atributos de posicionamento solar⁷: ângulo diário solar, ângulo horário solar, ângulo de elevação solar e azimute. Esses atributos são cíclicos e formam sempre a mesma distribuição (para cada usina) quando os dados relativos à mesma época do ano são comparados. Já os histogramas dos atributos meteorológicos devem ser vistos com cautela pois podem representar um "ano fora da curva" para a usina, e não são similares para usinas de regiões diferentes, como no caso do atributo da temperatura ambiente, apresentado pela Figura 21, reproduzida do Apêndice E. Observa-se que a usina AUS possui maior regularidade da temperatura ambiente ao longo do ano, com frequências altas em torno de 20 °C, ao passo que a usina EUA possui uma distribuição mais esparsa (com mínima de -11 °C e máxima de 35 °C).

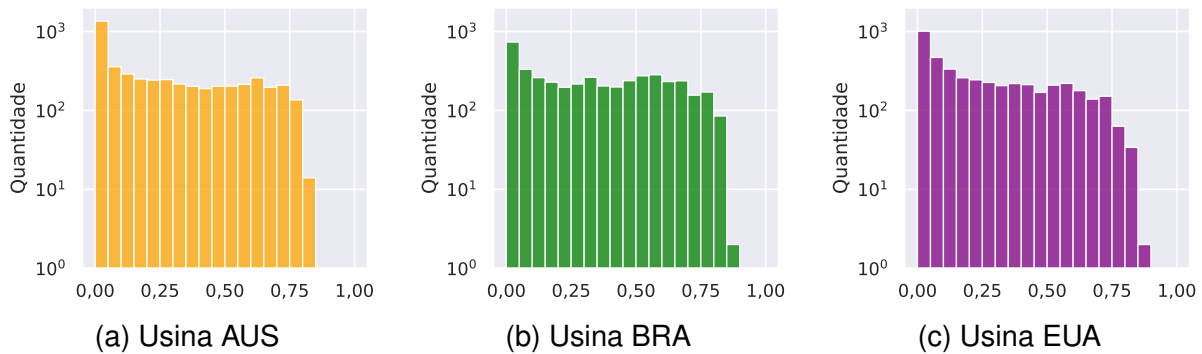
Figura 21 – Histograma do atributo Temperatura Ambiente (°C)



Fonte: elaboração própria (2022).

Do ponto de vista das saídas dos sistemas FV, os histogramas apresentam características semelhantes entre as usinas, como mostra a Figura 22, com frequências mais discrepantes para baixas e altas produções de energia, porém com o DP (em p.u.) muito próximo entre elas, como mostra a Tabela 9.

⁷Considerando o deslocamento de 6 meses entre as usinas de hemisférios diferentes.

Figura 22 – Histograma da Geração de Energia (p.u.)

Fonte: elaboração própria (2022).

Tabela 9 – Estatística de geração das usinas na amostra de desenvolvimento

Unidade	Usina	Méd	DP	Mín	25%	50%	75%	Máx
kWh	AUS	0,6	0,5	0,0	0,1	0,5	1,0	1,7
	BRA	30,4	24,7	-0,4	5,8	28,6	52,2	82,8
	EUA	35,2	33,7	-0,5	4,2	24,3	62,0	124,3
p.u.	AUS	0,28	0,25	0,00	0,04	0,23	0,51	0,82
	BRA	0,32	0,26	-0,00	0,06	0,30	0,55	0,87
	EUA	0,25	0,24	-0,00	0,03	0,17	0,44	0,89

Fonte: elaboração própria (2022).

3.3 Operacionalização do método de previsão proposto

As opções técnicas para a operacionalização do método de previsão são algo realizado pelos programadores em função da estrutura de equipamentos ou serviços de processamento de dados alvo. Pode-se sugerir a implementação do método através de três rotinas apresentadas pela Tabela 10, e detalhadas a seguir.

A Figura 23 ilustra a rotina de atualização dos dados históricos meteorológicos e de medição das usinas necessários para o treinamento da RNA.

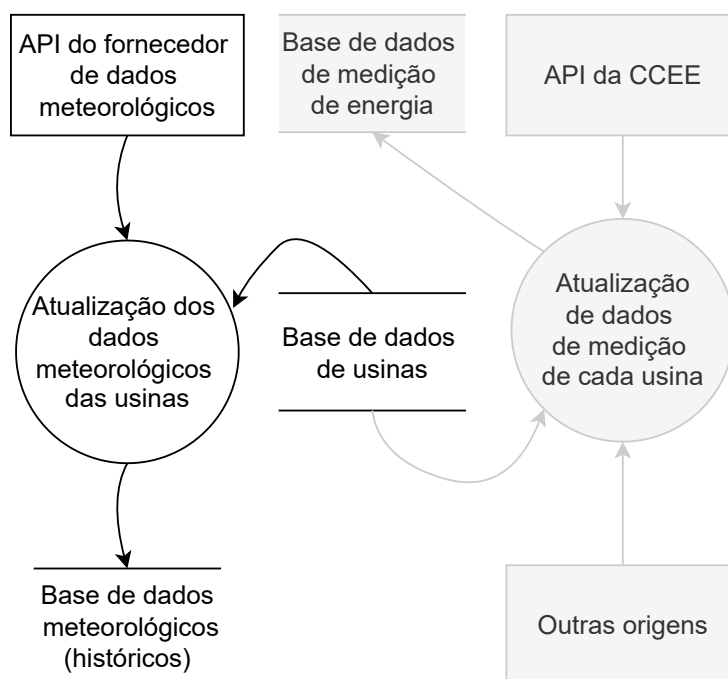
A disponibilidade de dados de medição das usinas é trivial aos sistemas de gerenciamento de plantas virtuais, e pode ser operacionalizado por comunicação direta com o sistema local ou através da Plataforma de Integração da CCEE (CCEE, 2022c). Os dados meteorológicos reais podem ser obtidos, via API, de um fornecedor desse tipo de serviço, como o proposto nesse trabalho (seção 3.2). A segunda rotina – de treinamento da RNA – é apresentada pela Figura 24. Se o treinamento anterior estiver dentro da frequência de atualização definida, o fluxo é interrompido.

As frequências de treinamento sugeridas na Tabela 10 são baseadas em

Tabela 10 – Rotinas para operacionalização da previsão de geração

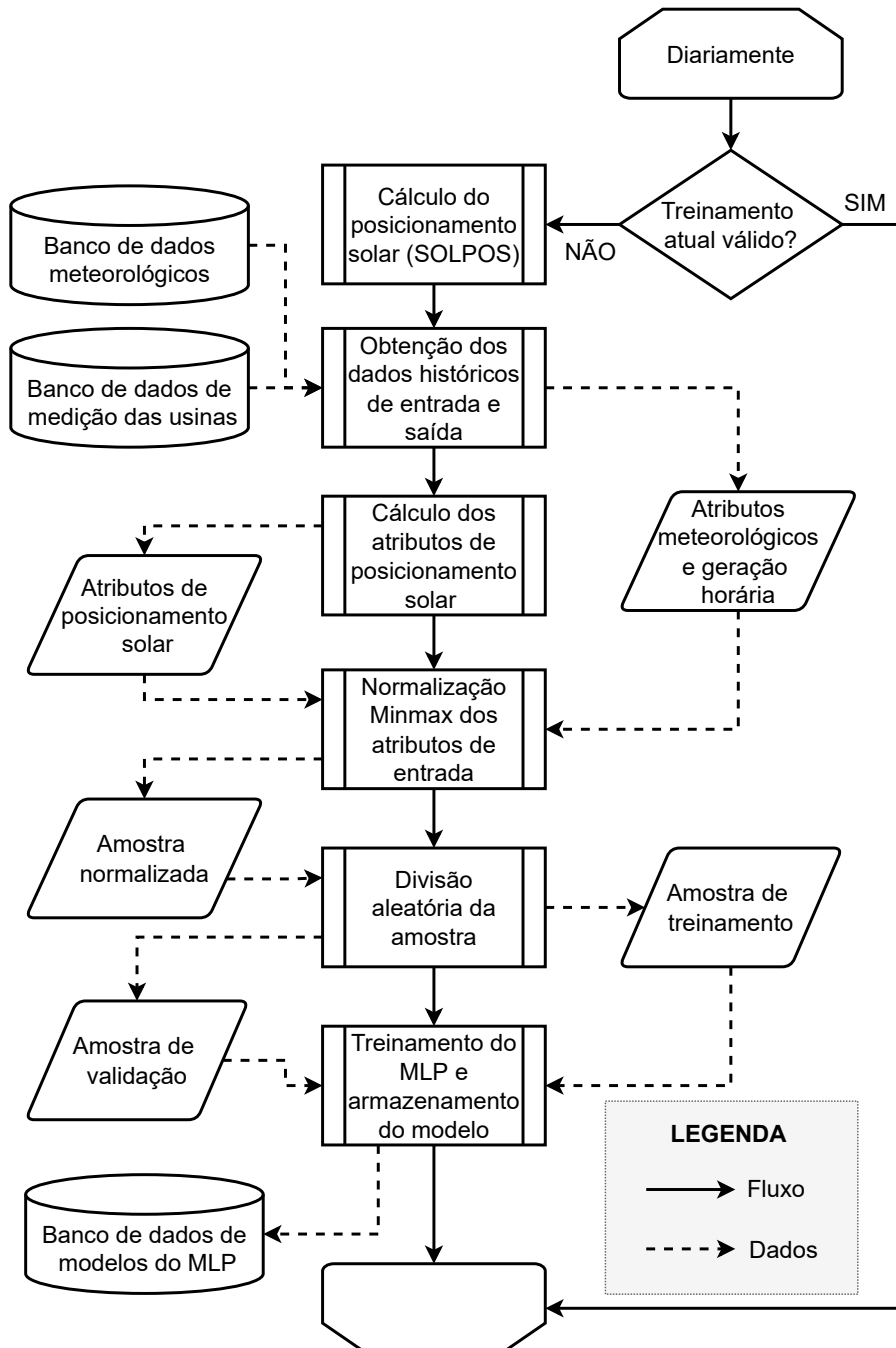
Rotina	Frequência	Comentários
Atualização	Diária	Armazenamento da medição de geração e das condições meteorológicas reais do dia anterior
Treinamento	Diária	Histórico de medição inferior a 6 meses
	Semanal	Histórico de medição inferior a 12 meses
	Mensal	Frequência padrão
Previsão de geração	Horária	Os dados meteorológicos são atualizados nesta frequência pelo fornecedor, uma vez que seus modelos vão se realimentando de informações atualizadas

Fonte: elaboração própria (2022).

Figura 23 – Diagrama de fluxo de dados da rotina de atualização

Fonte: elaboração própria (2022).

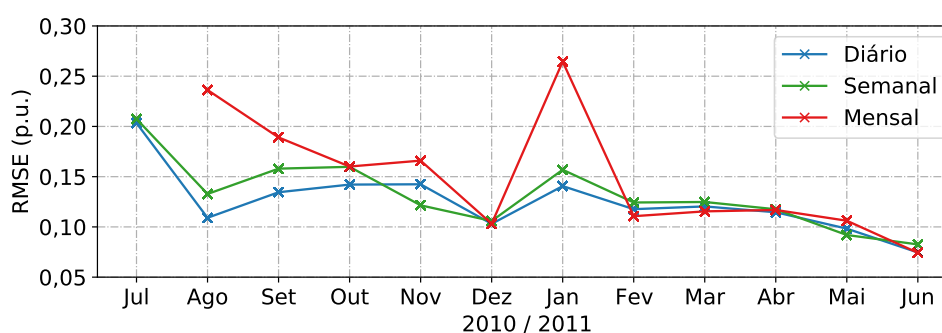
Figura 24 – Fluxograma da rotina de treinamento



Fonte: elaboração própria (2022).

testes realizados a partir da versão final do modelo de previsão⁸, e considerando-se as hipóteses de três diferentes frequências de realização do treinamento – e atualização do modelo –, cujos resultados para as usinas AUS e EUA são apresentados pelas Figuras 25 e 26. O objetivo é reduzir a necessidade de extrapolação pelo modelo, uma vez que as redes neurais não são próprias para isto (XU *et al.*, 2020). Quanto maior a frequência de treinamento, menores as chances de que previsões tenham de ser realizadas a partir de dados de entrada nunca vistos, ou seja, fora do subdomínio conhecido do problema, pois ao surgirem pela primeira vez, passam a integrar a solução mais rapidamente.

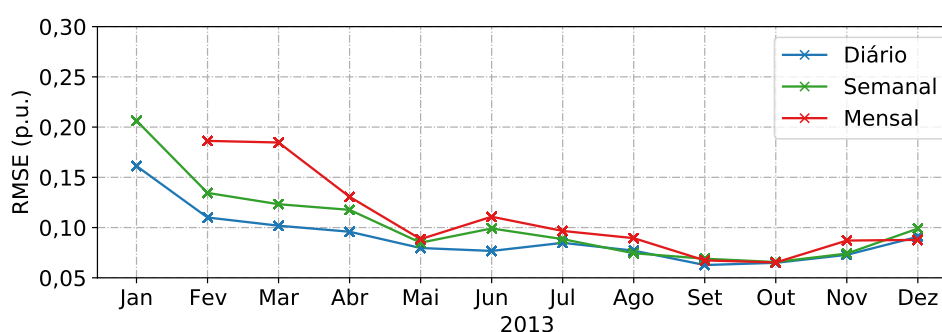
Figura 25 – Esquema de atualização dos dados meteorológicos – Usina AUS



Fonte: elaboração própria (2022).

O segundo dia, na frequência horária, é testado com o treinamento realizado com dados do primeiro dia, e o terceiro dia com dados dos dois primeiros dias. O mesmo vale para as demais frequências (com a adaptação do intervalo). Os resultados de cada mês são provenientes da consolidação dos resultados horários em cada frequência de atualização.

Figura 26 – Esquema de atualização dos dados meteorológicos – Usina EUA



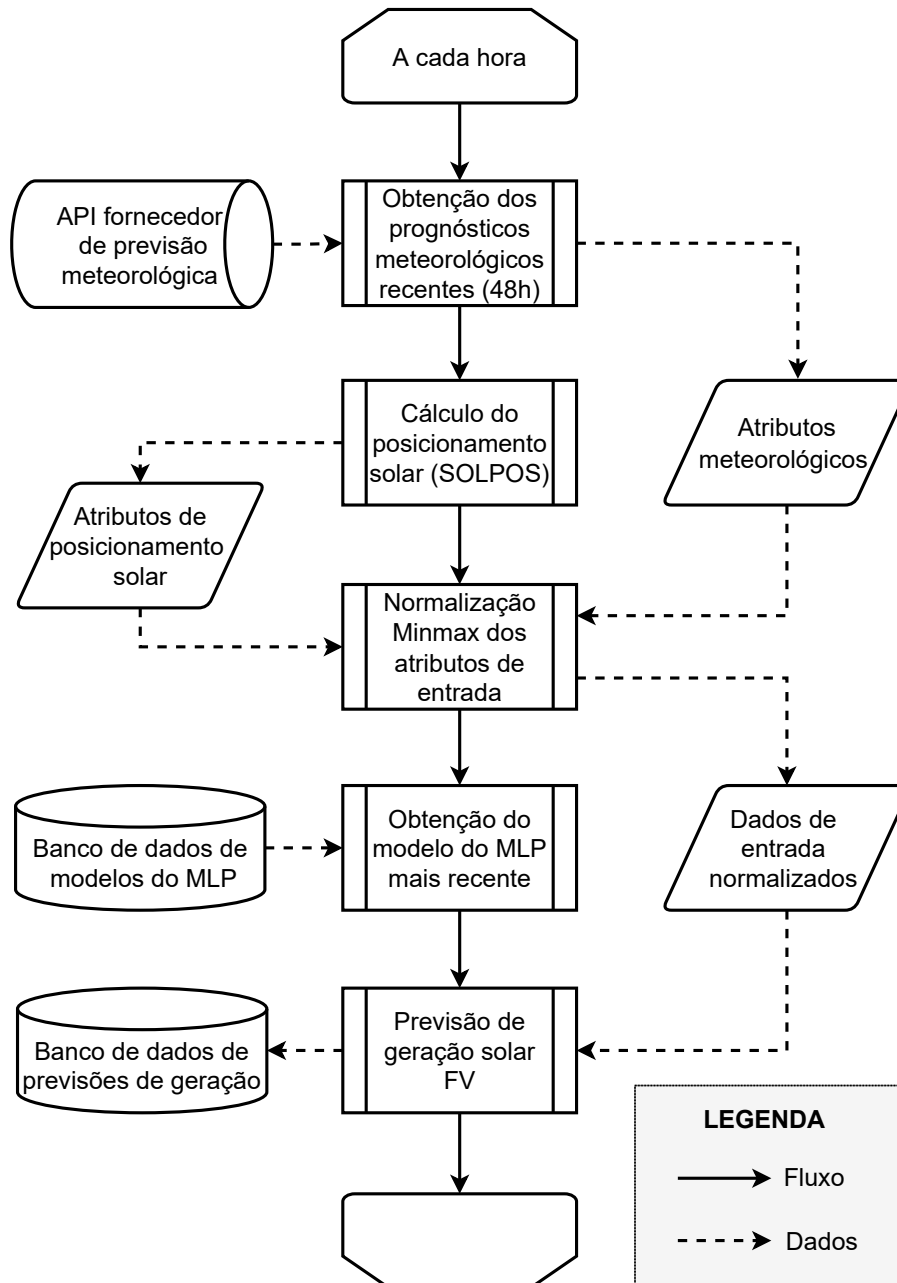
Fonte: elaboração própria (2022).

A rotina de previsão horária é apresentada pela Figura 27. Ela é acionada a cada hora para a previsão da geração solar FV de cada usina, acompanhando a

⁸Apela-se à paciência do leitor, pois esses dados são necessários a essa altura, embora conceitos importantes do modelo utilizado nesses testes sejam apresentados nas seções seguintes.

frequência de atualização dos prognósticos meteorológicos do fornecedor, os quais abrangem as próximas 48 horas, garantindo sempre a cobertura intradiária e do dia seguinte.

Figura 27 – Fluxograma da rotina de previsão

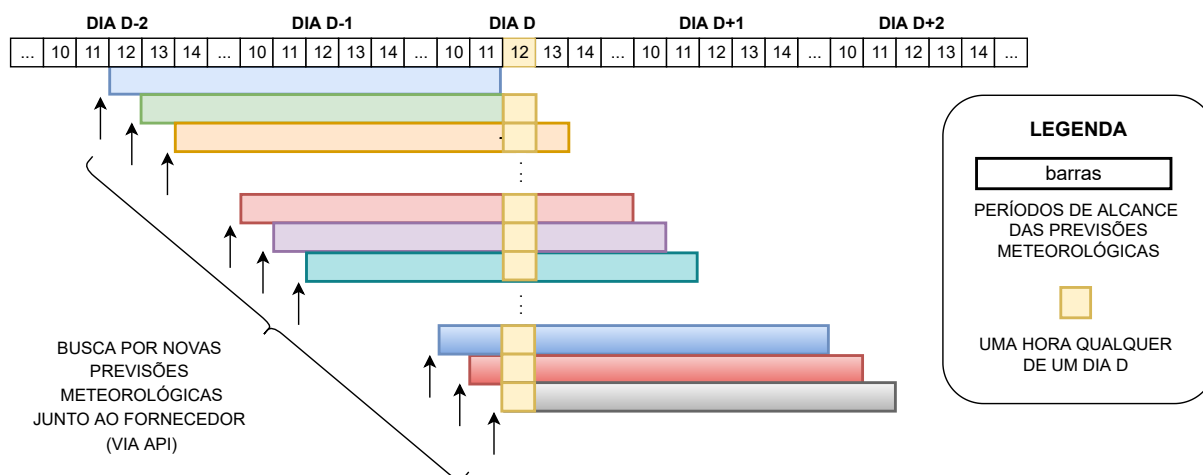


Fonte: elaboração própria (2022).

Os dados do posicionamento solar podem opcionalmente ser calculados uma única vez para cada usina e armazenados, pois são reaproveitáveis. Não é necessário armazenar os prognósticos meteorológicos, pois esses se alteram a cada hora, porém, eles podem ser uma fonte valiosa de informação para avaliação da performance do método e seu aperfeiçoamento futuro. A Figura 28 ilustra o processo iterativo de atualização dos dados meteorológicos que, uma vez alterados, devem gerar

novas previsões. Assim, uma determinada hora de interesse tem inicialmente uma previsão meteorológica de 48 horas de antecedência, que vai se aproximando até ser a primeira hora da previsão, que conta com maior confiabilidade, uma vez que certos fenômenos de entrada do modelo climático – do fornecedor – são menos incertos.

Figura 28 – Esquema de atualização dos dados meteorológicos



Fonte: elaboração própria (2022).

3.4 Justificativa da opção pelo *perceptron* multicamadas

O algoritmo proposto para a solução do problema da previsão de geração FV é o MLP. Como todas as RNAs ele tende a performar melhor que outros algoritmos de aprendizado de máquina tradicionais quando o tamanho da amostra e o porte da rede neural são grandes o suficiente (HORNİK; STINCHCOMBE; WHITE, 1989).

No caso da adesão das usinas às plantas virtuais, é comum que a desagregação de cargas e geradores ainda não esteja implementada. Isso implica na ausência de histórico inicial de medição, o que compromete os métodos baseados em regressão e os algoritmos de aprendizado de máquina em geral. Mas à medida que o tempo passa o volume de informações armazenadas cresce significativamente, considerando que um sistema FV é projetado para, com alguma manutenção e acompanhamento, durar 20 ou 30 anos (MELLIT; KALOGIROU, 2017).

Outro ponto de preocupação relacionado às redes neurais artificiais é a infraestrutura de processamento de dados necessária para o treinamento de um modelo capaz de generalizar o problema. Mas com o avanço do conceito de computação em nuvem e processamento distribuído esses recursos tornam-se cada vez mais acessíveis e, principalmente, não demandam investimento nem representam um custo fixo para os usuários. No caso das plantas virtuais, que dependem de conexão permanente com um número de instalações consumidoras e produtoras remotas, e em geral de

crescimento gradativo, a elasticidade da computação em nuvem é quase mandatória, e dificilmente escapará da tendência de tudo como serviço (*everything as a Service – XaaS*) do mercado contemporâneo (XU; AHOKANGAS; REUTER, 2018).

Embora muitos trabalhos na área de previsão de irradiação solar ou geração solar fotovoltaica utilizem outras arquiteturas de RNA, notadamente aquelas apropriadas para séries temporais (com capacidade de memorização) como as RNNs, ou mais recentemente as CNNs, destaca-se que estas foram relegadas deste estudo em função de dois principais fatores:

- a) acredita-se que uma parcela majoritária da característica série temporal do problema está relacionada com o posicionamento solar e as condições do tempo. Se esta hipótese estiver correta, o posicionamento solar já está sendo calculado e pertence ao rol das entradas, e alguns indicadores das condições do tempo – entradas do modelo – são incorporados diretamente da previsão meteorológica proveniente do fornecedor, via API, e, portanto, já são considerados;
- b) pretende-se construir o conhecimento de forma gradativa. Ao explorar uma RNA de AP convencional são gerados conhecimentos básicos relevantes para o aperfeiçoamento futuro e comparação com outros modelos.

Concluindo, o interesse pelo uso do aprendizado profundo para previsão de geração FV de curto prazo se apresenta como um primeiro passo do autor na área das redes neurais artificiais para solução de problemas relacionados ao setor elétrico.

3.5 Construção do *perceptron* multicamadas de previsão de geração fotovoltaica

A definição do modelo de uma rede neural artificial envolve uma série de opções e hiperparâmetros escolhidos em função do conhecimento científico, da experiência empírica ou de um processo de pesquisa de grade (ou em inglês, *grid search*). Como os recursos computacionais são limitados, não há a pretensão de percorrer todo o domínio do problema. Além disso, é importante que o tempo de execução e a carga de processamento do algoritmo computacional, em produção, seja conhecido, para evitar o risco de *bugs* sistêmicos. Desta forma, optou-se pela exclusão do *grid search* como alternativa em produção.

Para atingir uma estrutura MLP comum a todas usinas participantes do *pool* de uma VPP, procurou-se conhecer o impacto de certas características de sua configuração e ajuste na capacidade de generalização, observando o balanço entre a otimicidade e a praticidade de implementação e manutenção como algoritmo de previsão de um sistema maior, de gerenciamento de plantas virtuais. Para isso foram executados vários experimentos práticos envolvendo diferentes ajustes, os quais

encontram-se reportados nos Apêndices G, H, I, J, K, L e M, e que são referenciados ao longo desta seção. Além disso, as decisões tomadas para a construção da rede neural foram delimitadas pelo arsenal de ferramentas computacionais selecionadas, como mostra a Tabela 11.

Tabela 11 – Principais ferramentas computacionais utilizadas no desenvolvimento

Tipo	Descrição
Hardware	PC Dell XPS 8940 (I7, 16Gb) GPU NVIDIA GeForce RTX 2060 SUPER (compute capability 7.5)
Software	S. O. Linux Ubuntu Desktop 18.04 Docker Engine CE 20.10.10 Docker Image NVIDIA (tensorflow/tensorflow:latest-gpu-jupyter) Python 3.6.9 Jupyter Notebook 6.4.3 TensorFlow 2.6.0 (CPU + GPU) Pandas 1.1.5 Scikit-learn 0.24.2

Fonte: elaboração própria (2022).

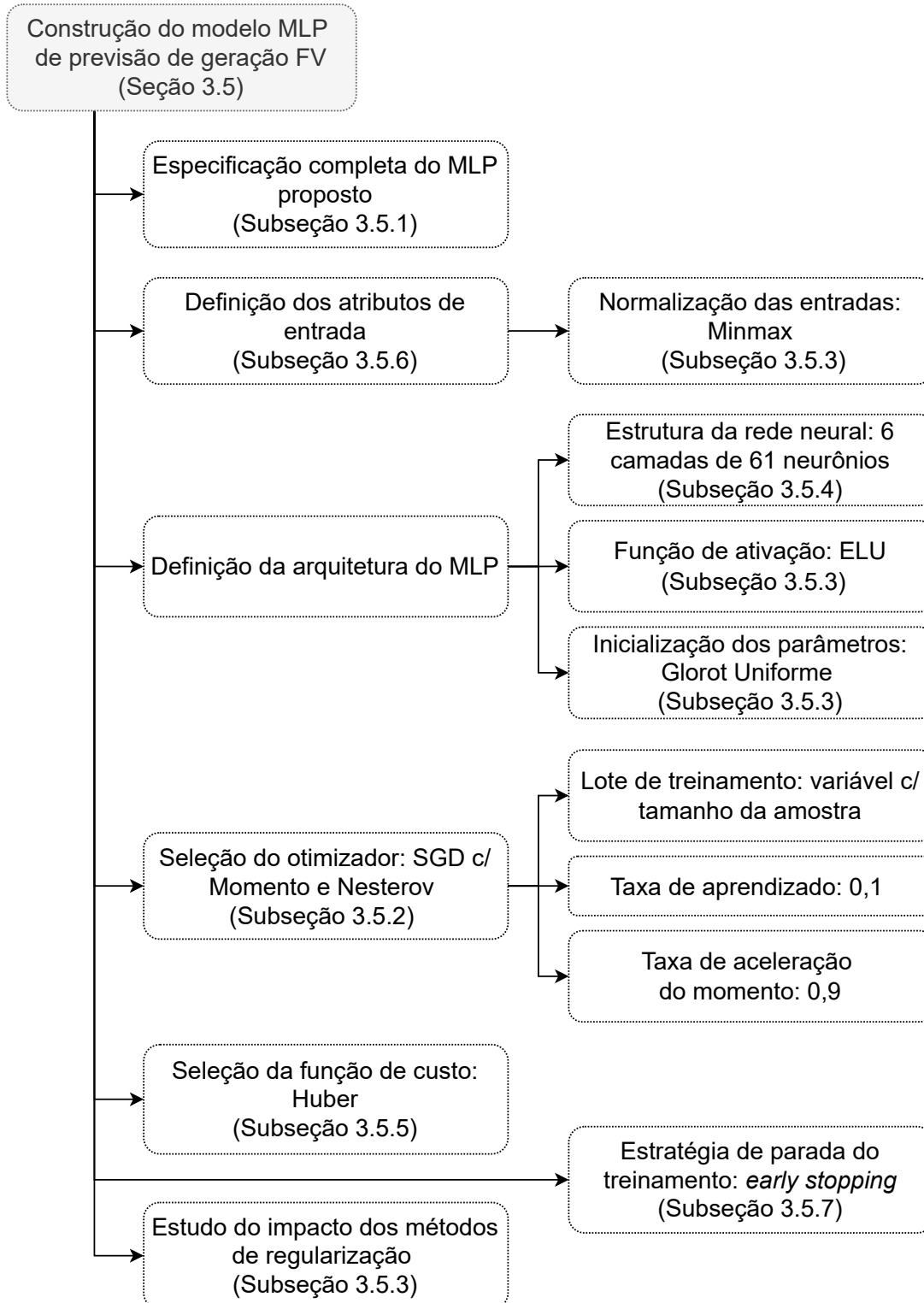
Além dessas ferramentas, foram desenvolvidos códigos em Python para coordenar o tratamento dos dados originais de acordo com as especificações propostas e para a realização dos treinamentos necessários aos experimentos com diferentes ajustes, além do estudo de casos.

A solução em sua forma integral e os motivos das escolhas realizadas são expostos nas subseções seguintes. Para facilitar a compreensão do contexto completo, a Figura 29 apresenta uma árvore dos aspectos considerados na modelagem, uma vez que as decisões tomadas são interdependentes e, portanto, difíceis de encadear de maneira lógica.

3.5.1 Especificação completa do MLP proposto

A especificação do *perceptron* multicamadas de aprendizado profundo sugerido para a previsão de geração solar fotovoltaica, nos termos deste trabalho, é apresentada pela Tabela 12. Essa estrutura pode ser implementada utilizando-se das bibliotecas de código aberto Tensorflow (ABADI *et al.*, 2015) e Keras (CHOLLET, 2015), para a linguagem de programação Python, que oferecem recursos suficientes para a construção, treinamento, previsão, avaliação, armazenamento e recuperação dos modelos.

Figura 29 – Árvore das decisões tomadas para construção do modelo MLP



Fonte: elaboração própria (2022).

Tabela 12 – Especificação da rede neural artificial

Característica	Valor
Entradas (atributos)	<i>azim, elevref, dayang, hrang, temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h</i>
Normalização	Minmax $\{-1, 1\}$ c/ <i>clip</i> ^a
Saída	Geração de energia (em p.u. ^b)
Tipo de rede	<i>perceptron</i> multicamadas (MLP)
Estrutura de camadas	
Entrada	18 neurônios (11 atributos ^c)
Ocultas	6 camadas de 61 neurônios cada
Saída	1 neurônio
Pesos e vieses	
Quantidade	20.131 parâmetros
Método de inicialização	Glorot uniforme c/ vieses zerados
Função de ativação	Unidade linear exponencial (ELU), exceto o neurônio de saída que utiliza a função identidade
Otimizador	Gradiente descendente estocástico (SGD)
Variação	c/ Momento e Nesterov
Taxa de aprendizado	0,1
Taxa de aceleração	0,9
Lote	Variável (mínimo de 8 exemplos) conforme a Equação 13
Função de custo	Huber
Número de execuções	3 (o valor de previsão é a média das execuções)

^a Valores de entrada de teste acima ou abaixo dos limites da normalização são truncados nos valores limites.

^b A base é a capacidade nominal de cada usina.

^c Os atributos de condições do tempo: *temp, pressure, humidity, clouds, wind_speed* e *wind_deg* são pares de valores relativos ao instante inicial e final de cada hora. O atributo de geoposicionamento solar *day_ang* é desdobrado em duas componentes: seno e cosseno.

Fonte: elaboração própria (2022).

3.5.2 Seleção do otimizador

A seleção do algoritmo de otimização para a solução de um problema de AM determina a velocidade de treinamento e a performance final de predição do modelo, embora não exista – até o momento – uma teoria capaz de direcionar essa escolha (CHOI *et al.*, 2019).

Os algoritmos utilizados nas redes neurais artificiais são definidos caso a caso, a partir de testes próprios, ou observando os resultados apresentados na literatura, provenientes da comparação empírica. No entanto, mesmo para os bancos de dados padronizados e difundidos entre os pesquisadores, esses resultados podem ser conflitantes a depender do espaço de hiperparâmetros utilizado.

O arsenal desse trabalho contém os otimizadores de primeira ordem da biblioteca Keras do TensorFlow. Eles são apresentados na Tabela 13.

Wilson *et al.* (2017) sugerem que os métodos que utilizam gradientes adaptativos levam a resultados inferiores aos dos métodos tradicionais – do gradiente descendente ou SGD – e deveriam ser evitados. Seus testes envolvendo classificação binária em uma população linearmente separável com o emprego de um modelo generativo resultou em soluções muito diferentes, e apenas o SGD atingiu o erro zero na amostra de teste, ainda que com piores resultados de treinamento.

Já Schneider, Balles e Hennig (2019) buscaram criar uma forma padronizada de verificação da performance dos otimizadores e a testaram em oito bancos de dados de teste amplamente difundidos, principalmente com redes neurais convolucionais ou recorrentes. Suas conclusões principais foram que não há um algoritmo de otimização com desempenho melhor que os demais em todos os problemas ou cujas métricas de treinamento e validação se mantenham proporcionais entre esses problemas. O desempenho é dependente do problema. Dentre os três algoritmos testados (SGD, SGD c/ momento e Adam), o Adam apresentou os hiperparâmetros de velocidade de aprendizado ótimos mais concentrados – entre 10^{-3} e 10^{-4} em sete dos problemas, ao passo que para os demais foram um pouco mais espalhados.

No entanto, Choi *et al.* (2019) testaram os algoritmos SGD, SGD c/ momento, SGD c/ nesterov, RMSprop, Adam e Nadam considerando um espaço mais amplo dos hiperparâmetros e concluíram que nenhum dos algoritmos generalistas é pior que sua referência mais especializada se esgotadas as possibilidades de ajuste de seus hiperparâmetros. Os resultados de outros trabalhos no sentido inverso estariam contaminados pela premissa inicial – fixação de alguns dos hiperparâmetros em seus valores padrões sugeridos – ou estariam considerando apenas o início do processo iterativo, sem avançar o suficiente no processo de treinamento (número de épocas).

É importante observar que os algoritmos da Tabela 13 criados na última

Tabela 13 – Otimizadores da biblioteca Keras do TensorFlow

Algoritmo	Gradiente adaptativo	Referências	Observação
SGD	Não	(ROBBINS; MONRO, 1951)	Método do gradiente em sua forma mais básica. Possui a convergência mais lenta, mas é muito testado e possui sólida teoria.
SGD c/ momento	Não	(POLYAK, 1964)	Acelera o SGD na direção relevante e amortece oscilações. Basicamente libera o avanço na direção comum entre as iterações e reduz o avanço em nova direção.
SGD c/ nesterov	Não	(SUTSKEVER <i>et al.</i> , 2013)	Age sobre o SGD c/ momento impedindo um salto muito rápido na direção relevante, o que resulta no aumento da capacidade de resposta e do desempenho das RNAs.
Adagrad	Sim	(DUCHI; SINGER, 2011)	Age sobre o SGD adaptando a taxa de aprendizagem aos parâmetros (ou atributos), realizando atualizações maiores para os parâmetros raros e menores para os parâmetros mais frequentes. Dispensa o ajuste da taxa de aprendizagem e é apropriada para dados esparsos.
Adadelta	Sim	(ZEILER, 2012)	Extensão do Adagrad que busca reduzir a agressividade da variação da taxa de aprendizagem, e dispensa a definição de uma taxa inicial.
RMSprop	Sim	(HINTON; SRIVASTAVA; SWERSKY, 2012)	Idêntico ao Adadelta, embora independente. Busca resolver o problema de redução excessiva, em alguns casos, da taxa de aprendizagem pelo Adagrad.
Adam	Sim	(KINGMA; BA, 2014)	Um dos métodos adaptativos mais testados e utilizados na prática. Essencialmente, combina Adadelta ou RMSprop com momento, armazenando tanto a taxa de aprendizagem individual (RMSprop) quanto a média ponderada do momento.
Adamax	Sim	(KINGMA; BA, 2014)	A partir do otimizador Adam, este algoritmo busca maior estabilidade através do equacionamento utilizado.
Nadam	Sim	(DOZAT, 2016)	Combina Adam com Nesterov para obter um passo mais preciso na direção do gradiente.
Ftrl	—	(MCMAHAN <i>et al.</i> , 2013)	Desenvolvido pelo Google para previsão em modelos básicos com espaço de atributos amplo e esparsos.

década foram motivados por desafios mais conectados com os sistemas automatizados de reconhecimento e produção visual e de linguagem, e redes mais complexas que as redes neurais de camadas densamente conectadas. Os problemas de regressão clássica – como o presente – são mais triviais e não oferecem os mesmos desafios.

Assim, dadas as limitações de hardware e tempo para execução de testes envolvendo todos os algoritmos e hiperparâmetros, e sabendo-se que um teste superficial envolvendo hiperparâmetros não otimizados de cada algoritmo conduziria a uma conclusão, ao final, aleatória, arbitrou-se pela utilização do algoritmo tradicional gradiente descendente estocástico (SGD) com momento do tipo nesterov.

Em sua forma estocástica, as iterações parciais de cada época envolvem um lote de exemplos sorteados da amostra de treinamento, em geral sem repetição dos exemplos. Para entender se este hiperparâmetro – o tamanho do lote – pode afetar o resultado global, o Experimento 6 – detalhado no Apêndice F – testou tamanhos de lote entre 8 (oito) e 48 exemplos para a amostra de desenvolvimento⁹, com duas estruturas de rede neural de interesse neste trabalho, antes de um processo de regularização, uma vez que o próprio tamanho do lote de treinamento pode contribuir para a regularização (KALE; SEKHARI; SRIDHARAN, 2021). O experimento mostrou que:

- a) a primeira estrutura, com normalização Minmax e função de ativação unidade linear exponencial (ELU) apresentou resultados similares com os lotes testados. Apesar de alguma variação da distribuição dos valores de validação cruzada, as médias são próximas para a amostra de cada usina;
- b) a segunda opção, com normalização Z-score e função de ativação unidade linear retificada (ReLU) apresentou maior variabilidade entre os resultados, sem uma tendência clara, exceto que lotes pequenos podem ser benéficos.

Como o tempo de execução do treinamento em relação ao número de iterações do algoritmo, dentro de cada época, apresentou complexidade do tipo $O(n)$, durante os testes, o crescimento contínuo do tamanho da base de dados das usinas, com o passar do tempo, trará uma demanda proporcional de tempo de hardware para o treinamento. Para dar estabilidade ao processo de treinamento, sugere-se um lote variável, dado por¹⁰:

$$\text{tamanho do lote} = \max \left(8, 2^{\left(\lfloor \log_2(n^{\text{exemplos}}) \rfloor - 7 \right)} \right), \quad (13)$$

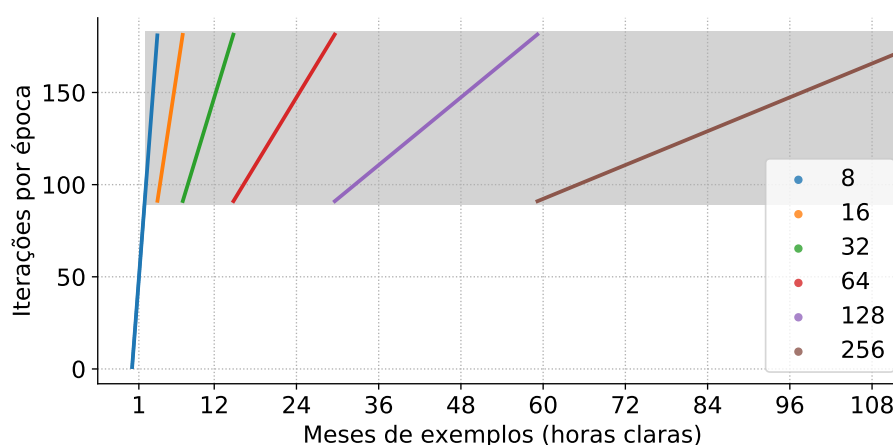
⁹A amostra de desenvolvimento compreende aproximadamente 4700 exemplos de cada usina, relativos a 12 meses de operação.

¹⁰Nessa equação é utilizada uma notação de arredondamento incomum que significa arredondar para o inteiro mais próximo ($\lfloor x \rfloor$).

mas ainda dentro do conceito de lote pequeno (*small batch*), visando evitar a degradação da generalização da solução, com a geração de mínimos locais apertados (*sharp minima*) (KESKAR *et al.*, 2017).

O resultado é um lote variável e controlado – com no mínimo 8 exemplos – que mantém o número de iterações controlado, dentro de uma faixa de iterações por época, como mostra a Figura 30, contribuindo para a estabilidade dos recursos de hardware utilizados, seja quanto ao tempo de processamento (aumenta com o número de iterações) ou uso de memória (aumenta com o tamanho do lote). Há outras formas mais especializadas de lidar com o aumento do tamanho do lote, como o proposto por Hoffer, Hubara e Soudry (2017), em que o ajuste é realizado em tempo de execução, iniciando pequeno e crescendo ao longo do treinamento – complexidade considerada desnecessária. Os lotes foram definidos na potência de dois para melhor eficiência de hardware (CPU/GPU), algo mencionado em fóruns da internet (ATWOOD; SPOLSKY, 2022), e verificado na prática durante os experimentos.

Figura 30 – Variação do tamanho do lote de treinamento por tamanho da amostra



Fonte: elaboração própria (2022).

Finalmente, a taxa de aprendizado foi alterada do valor padrão do pacote Tensorflow/Keras, de 0,01 (um centésimo), para 0,1 (um décimo), com base em um experimento realizado para avaliá-la, conforme o Apêndice M. Com isso, os resultados de validação são próximos, com agilização do treinamento – cerca de 1/3 (um terço) do número de épocas necessárias com ajuste padrão. A taxa de aceleração do momento foi definida no valor padrão sugerido para o algoritmo SGD com momento (SGDM) pelos autores do TensorFlow/Keras (ABADI *et al.*, 2015).

3.5.3 Normalização das entradas, função de ativação e estratégia de inicialização dos parâmetros

A forma de normalização dos valores dos atributos de entrada do modelo, bem como a estratégia de inicialização dos parâmetros (pesos e vieses) e o tipo de

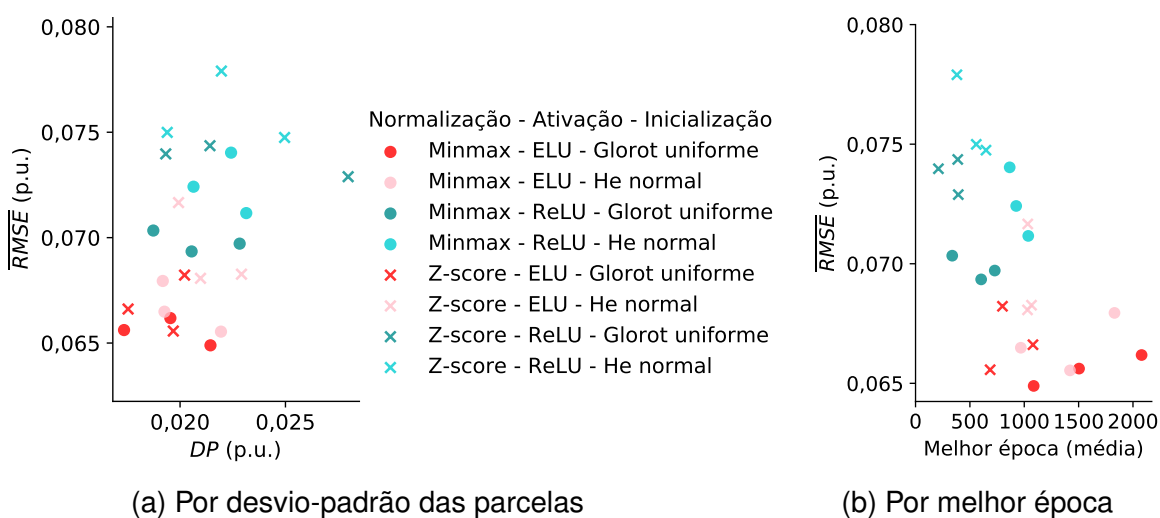
função de ativação dos neurônios da rede neural foram estudados de forma agrupada através de uma série de experimentos práticos envolvendo as três usinas. Os valores dos vieses iniciais foram assumidos zerados.

A definição da estratégia de inicialização dos pesos da rede neural artificial baseou-se no Experimento 1 – detalhado no Apêndice G – e envolveu oito variações das estratégias de normalização de entradas, ativação e inicialização dos pesos. Além disso, para aumentar a representatividade dos resultados, foram testadas três estruturas diferentes, de uma a três camadas ocultas e mesmo número de neurônios em cada camada, todas elas com aproximadamente 2.000 parâmetros totais (vieses e pesos). As condições de teste e os resultados são detalhados no Apêndice G, e as alternativas testadas são:

- função de ativação ReLU vs. ELU;
- inicialização Glorot uniforme vs. He normal;
- normalização de entradas Minmax vs. Z-score.

A Figura 31 apresenta os resultados de validação cruzada da usina EUA. As conclusões são similares entre as usinas. Este experimento mostrou que a inicialização do tipo Glorot uniforme (cores fortes) proporciona melhor generalização que He normal, embora de forma marginal, mas em todos os casos, seja com a normalização de entradas do tipo Minmax ou Z-score, ou ainda com a função de ativação ReLU ou ELU. Por isso, foi selecionada a inicialização do tipo Glorot com distribuição dos pesos uniforme.

Figura 31 – Resultados de validação cruzada da Usina EUA por alternativa



Fonte: elaboração própria (2022).

A função de ativação ReLU já se mostrou apropriada em diversas aplicações práticas e é indicada como ponto de partida em qualquer projeto de rede neural (GÉRON, 2019). Mas para ter uma base de comparação e eventualmente prevenir

o risco da "morte" de neurônios, ou seja, o caso em que os pesos atingem valores negativos e os gradientes permanecem zerados durante o restante do treinamento, optou-se por verificar também a função de ativação ELU.

O resultado, como visto na Figura 31a, é que a ELU (itens avermelhados) generaliza melhor que a ReLU para o problema em estudo, antes da regularização. Porém ela é mais lenta que a ReLU em todas as situações testadas, o que pode ser visto na Figura 31b pela época de início da caracterização de sobreajuste, ou seja, no ponto de inflexão dos valores de validação. Para entender se essas conclusões seriam mantidas após um processo de regularização, foram testados quatro tipos de regularização:

- a) lasso e ridge: Experimento 3, detalhado no Apêndice I;
- b) rede elástica (combinação de lasso e ridge): Experimento 4, detalhado no Apêndice J;
- c) aumento de dados: Experimento 5, detalhado no Apêndice K;

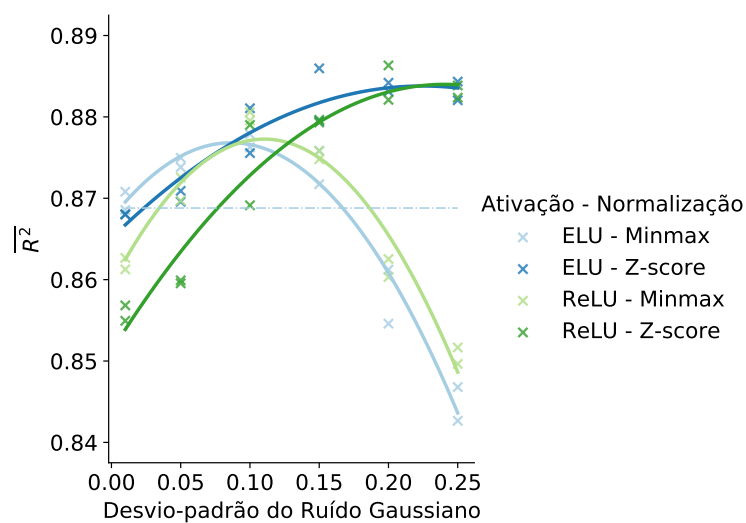
envolvendo as combinações dessas duas funções de ativação com as normalizações de entrada Minmax e Z-Score. Desta vez, uma estrutura de quatro camadas ocultas e 83 neurônios por camada oculta, com cerca de 20 mil parâmetros totais (pesos e vieses), foi a base de comparação. O Experimento 3 testou as regularizações dos tipos lasso e ridge separadamente, o que permitiu concluir que:

- a) a combinação Minmax/ELU não se beneficia razoavelmente da aplicação de regularização de ambos tipos;
- b) quanto pior o resultado antes da regularização, maior o benefício da regularização de ambos tipos;
- c) as usinas BRA e EUA não se beneficiam da aplicação de ambos tipos de regularização em relação à melhor combinação antes da regularização;
- d) a usina AUS é a única que tem um resultado significativamente melhor após a regularização, com a combinação Z-score/ReLU;
- e) a regularização dobra o tempo de execução, mas não altera a ordem de grandeza entre ReLU (mais rápida) e ELU (mais lenta).

Para verificar se a combinação de lasso e ridge poderia trazer oportunidades melhores, o Experimento 4 testou a hipótese de aplicação da rede elástica, limitando-se à combinação Z-score/ReLU, mais rápida, e que trouxe vantagem significativa para ao menos uma das usinas. Os melhores resultados de validação cruzada utilizando-se da rede elástica ultrapassaram os resultados com ridge, mas não se sobressaíram em relação a lasso. Dada a maior complexidade dos hiperparâmetros envolvidos neste tipo de regularização, mostra-se inviável.

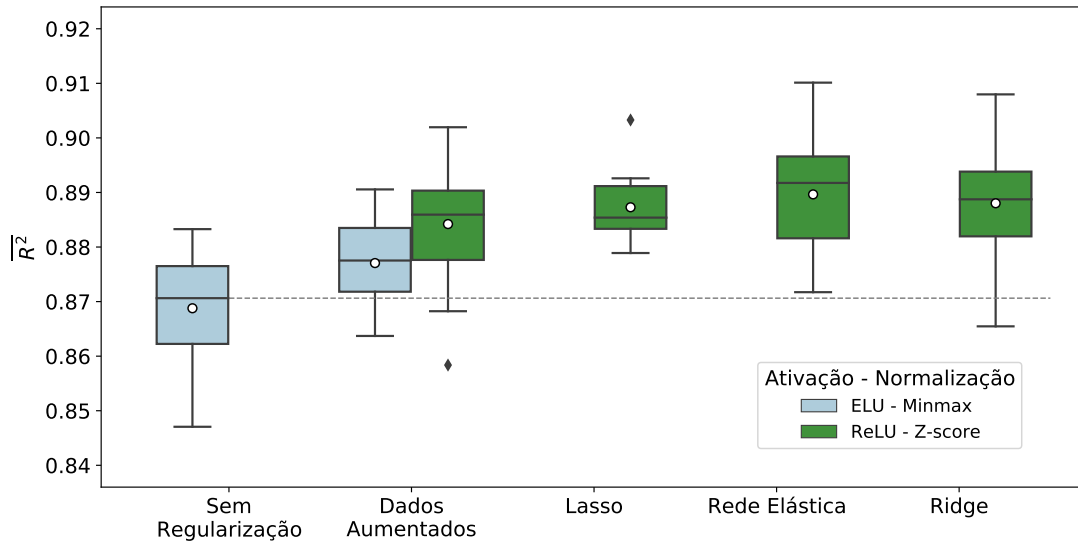
O Experimento 5 buscou verificar se o processo de regularização por dados aumentados através da aplicação de ruído gaussiano é capaz de melhorar a generalização das RNAs estudadas. Foram testadas quatro combinações envolvendo ELU e ReLU com Minmax e Z-score, e para a mesma estrutura dos Experimentos 3 e 4. Apenas os resultados da usina AUS apresentaram benefício deste tipo de regularização, como mostra a Figura 32. A linha horizontal tracejada representa o melhor resultado sem regularização, obtido com ELU/Minmax. As demais usinas tiveram resultado pior para todas combinações. Destaca-se o fato de que os melhores resultados estão associados à normalização de entradas Z-score.

Figura 32 – Resultados de validação cruzada da aplicação de dados aumentados

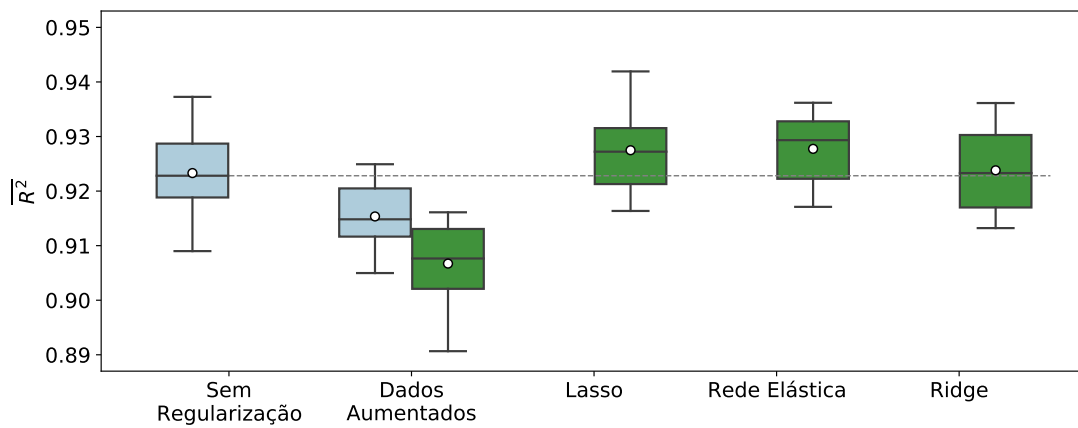


Fonte: elaboração própria (2022).

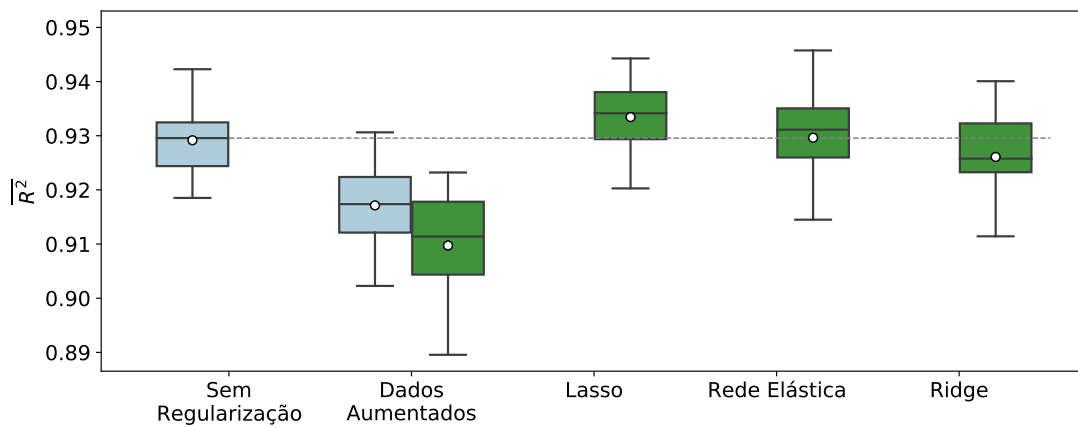
Figura 33 – Resultados de validação cruzada por tipo de regularização



(a) Usina AUS



(b) Usina BRA

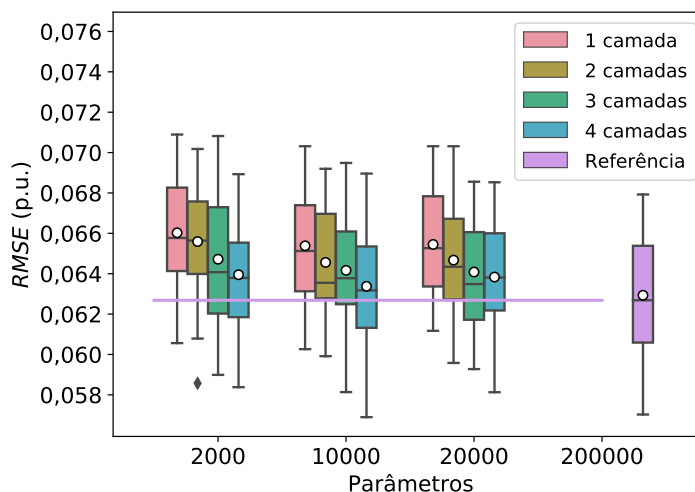


(c) Usina EUA

Fonte: elaboração própria (2022).

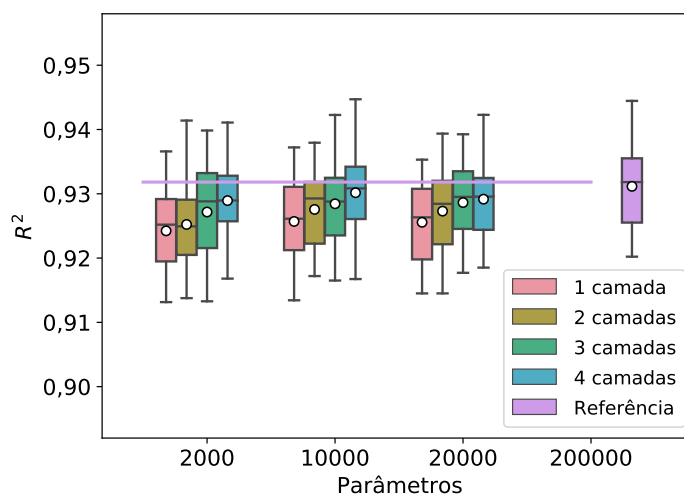
diferentes.

Figura 34 – RMSE de validação cruzada por número de camadas ocultas e de parâmetros – Usina EUA



Fonte: elaboração própria (2022).

Figura 35 – R^2 de validação cruzada por número de camadas ocultas e de parâmetros – Usina EUA



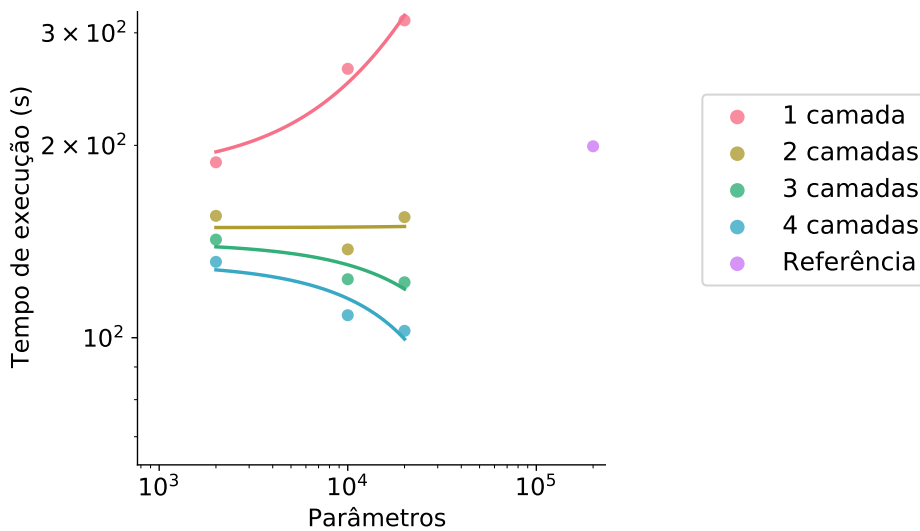
Fonte: elaboração própria (2022).

Verifica-se, tomando o resultado da estrutura de referência como base, que os resultados não melhoram significativamente com o aumento do número de parâmetros, mas sim com o aumento do número de camadas ocultas, mantido igual – durante o experimento – o número de neurônios por camada. Isso converge com alguns achados de Golubeva, Neyshabur e Gur-Ari (2020), embora seus experimentos tenham sido todos sobre bancos de dados de imagens.

Por outro lado, também não há uma piora da dispersão com o aumento do número de parâmetros, algo esperado em outros tipos de algoritmos de aprendizado de máquina. Ao observar as curvas do tempo de execução médio (todas usinas)

para diferentes portes de RNA, até a configuração do sobreajuste, como mostra a Figura 36, percebe-se que com o aumento do número de camadas ocultas, tem-se uma redução do tempo total de execução – uma combinação entre o tempo por época e o número de épocas até a solução – para um mesmo número de parâmetros. Mas o aumento do número de parâmetros só é benéfico se acompanhado do aumento do número de camadas ocultas. Esses resultados confirmam o fato de que o AP visa justamente reduzir o esforço computacional e aumentar a capacidade de representação dos padrões repetitivos existentes nos dados visando a generalização (AGGARWAL, 2018).

Figura 36 – Tempo de execução do treinamento em função do número de camadas ocultas e de parâmetros (todas usinas)



Fonte: elaboração própria (2022).

Aparentemente, a rede de referência atingiu um tamanho superior à capacidade do hardware utilizado de multiplicar entradas e pesos de uma única vez. Com dez vezes o número de parâmetros, e apenas seis camadas ocultas, o número de neurônios por camada¹² ficou excessivo. Ela foi resolvida com a GPU, ao passo que as demais estruturas foram mais rapidamente resolvidas com a CPU (exclusivamente).

Desta forma, pela tendência apresentada, optou-se por uma estrutura ligeiramente mais complexa que as já testadas, de seis camadas ocultas de 61 neurônios cada, totalizando 20.131 parâmetros (pesos e vieses) de treinamento.

3.5.5 Seleção da função de custo

A disponibilidade de irradiação solar varia ao longo do ano, com pico e maior número de horas claras no verão. Essa característica, somada ao eventual azimute dos painéis solares – que altera a simetria entre a geração da manhã e da tarde –

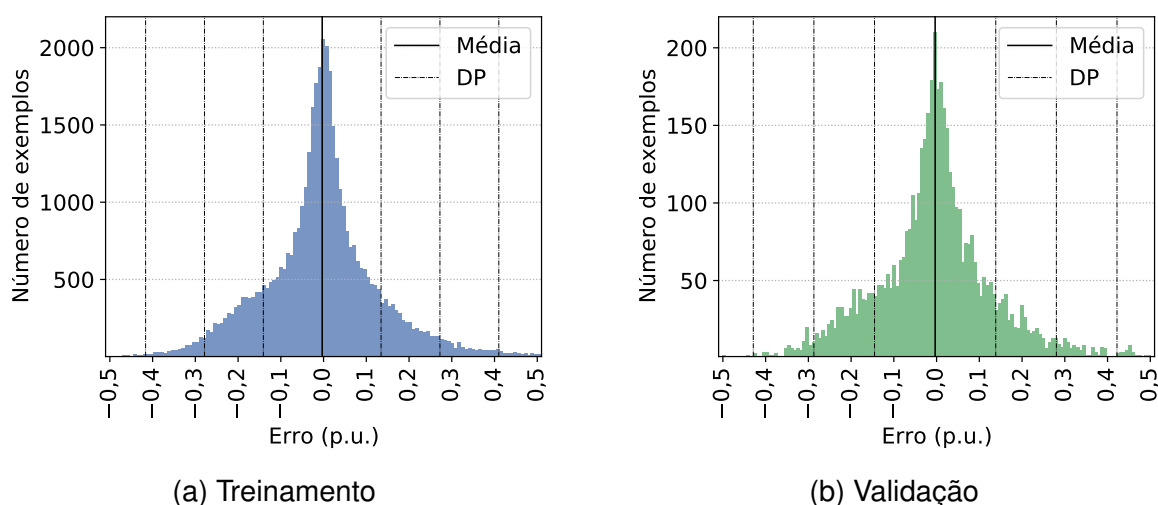
¹²A estrutura de referência é do tipo 19 (entradas) x 198 x 198 x 198 x 198 x 198 x 198 x 1 (saída).

contribuem para a variação da curva de geração de energia da usina. Finalmente, fatores como a cobertura de nuvens ou reflexão do terreno e das nuvens completam este cenário de dispersão entre as horas do dia e os dias do ano.

A função de custo erro quadrático médio (MSE) – a mais comum aos problemas de regressão – é uma alternativa interessante porque impõe um peso maior para os erros maiores, de forma que a solução tende ao equilíbrio entre viés e variância – dadas as devidas proporções¹³ – ao longo de todas as horas do ano. Isso é positivo porque, além das previsões alimentarem processos de decisão das plantas virtuais, seus valores são apresentados na tela da interface humana, e neste ponto, parafraseando um famoso provérbio¹⁴, a previsão deve ser boa – ter erros médios reduzidos (ao longo do tempo) – e parecer boa – ter forma similar à real quando apresentada numa interface de gerenciamento.

No entanto, durante o desenvolvimento verificou-se que uma parcela dos exemplos de treinamento desviam de forma significativa do modelo do qual participam. A Figura 37 apresenta o exemplo da usina AUS, a qual possui a maior dispersão dos erros de treinamento da amostra de desenvolvimento, com o modelo final selecionado.

Figura 37 – Histograma dos erros de validação cruzada da Usina AUS



Fonte: elaboração própria (2022).

Esses pontos fora da curva só foram identificados após o treinamento. Três estratégias foram pensadas inicialmente:

- a) identificação e exclusão dos *outliers* com base nos dados de entrada;
- b) substituição da função de custo MSE pela função Huber – que reduz a importância dos erros maiores –; e

¹³Exemplos com características semelhantes, em quantidade maior, podem prevalecer e distorcer o resultado dos exemplos incomuns (baixa ocorrência).

¹⁴"À mulher de César não basta ser honesta, deve parecer honesta".

- c) um processo de treinamento duplo: uma primeira execução seria realizada com a integralidade dos exemplos da amostra, e após a identificação dos pontos fora da curva, uma segunda execução seria realizada sem eles, com validação pela amostra de validação intocada.

Testadas, a estratégia da função Huber, própria para o amortecimento do efeito de pontos fora da curva, melhorou a distribuição dos erros de treinamento mas não teve efeito significativo sobre os erros de validação, apenas marginal, provavelmente em função de que a amostra de validação também contém esses *outliers*. A estratégia de treinamento duplo, menos elegante, trouxe alguns resultados melhores de validação, mas aleatórios entre as usinas e entre diferentes execuções. A busca de uma técnica de identificação de pontos fora da curva com base nas próprias entradas foi descartada pela complexidade adicional e os recursos limitados.

Assim, a função de custo Huber foi a opção selecionada, por precaução, com o parâmetro δ igual a 0,33 p.u., algo acima de dois desvios padrão e acima da média de geração de todas as usinas. Valores menores indexados nas métricas das usinas (média e DP) resultaram insatisfatórios, provavelmente porque no início da otimização os erros são naturalmente elevados, e esses limites impactam demais o caminho de busca da solução.

3.5.6 Definição dos atributos de entrada

A modelagem da rede neural artificial foi iniciada com a perspectiva de se restringir os atributos irrelevantes através da regularização do tipo Lasso, que possui a propriedade de enfraquecer os pesos dos neurônios de menor relevância (tendendo a zerá-los) e fortalecer aqueles que conectam os pesos mais relevantes. Funciona como um filtro dos atributos, embora não garanta a exclusão completa de um determinado atributo. Essa opção decorreu da crença de que a relevância dos atributos pode diferir entre as usinas. Mas os resultados dos experimentos indicam, como visto na Figura 34, que a configuração proposta dispensa – considerando os prós e contras – a necessidade de regularização.

Desta forma, os atributos inicialmente propostos são mantidos (Tabela 6), com uma exceção. Já na fase final de desenvolvimento, observou-se que o atributo do ângulo diário solar (*dayang*) interfere significativamente nos resultados das previsões quando os dados de entrada extrapolam o domínio aprendido pela RNA. Principalmente enquanto a amostra de treinamento não inclui exemplos representativos de todo o ciclo solar anual. Por exemplo, se os dados de treinamento referem-se ao segundo semestre do ano (ângulos entre 180° e 359°), ao tentar prever os primeiros dias de janeiro do ano seguinte, o ângulo cai subitamente para 0° . Em função da normalização Minmax, o treinamento considera valores do atributo *dayang* entre -1 (180°) e 1 (359°), e os

parâmetros de previsão normalizados nas mesmas bases iniciam-se em -3 (0°), muito fora do domínio aprendido.

Não é possível perceber o problema enquanto se trabalha com validação cruzada estratificada porque a normalização leva em conta todos os exemplos, e cada parcela tende a conter a mesma diversidade da amostra completa. A transformação do atributo em duas componentes – seno e cosseno – reduziu as discrepâncias da extração, e melhorou o resultado de validação cruzada (interpolação), como mostra a Tabela 14. A estrutura comparada é a MLP de 4 camadas e cerca de 20.000 parâmetros, com normalização Minmax e função de ativação ELU, utilizada nos experimentos.

Tabela 14 – Resultados com Ângulo Diário Solar vs. suas componentes trigonométricas

Usina	Métrica	Ângulo Diário Solar ($^\circ$)	Seno e Cosseno
AUS	R^2	0,8688	0,8734
BRA	R^2	0,9233	0,9246
EUA	R^2	0,9292	0,9322

Fonte: elaboração própria (2022).

3.5.7 Estratégia de parada do treinamento

A decisão do momento de parada do treinamento é feita através da técnica da parada antecipada (*early stopping*). Com a experiência obtida ao longo dos treinamentos realizados foi possível estabelecer uma paciência (espera) de 200 épocas antes da desistência do treinamento. Especialmente com a taxa de aprendizado maior que o padrão da biblioteca Tensorflow e sem a aplicação de um método de regularização. Assim, quando a RMSE de validação apresenta um resultado mais baixo, reinicia-se a contagem do número de épocas, e se não melhorar dentro de 200 épocas, o processo é interrompido e o melhor resultado até então é recarregado no modelo.

A amostra de validação assume, nessa estratégia, a responsabilidade de representar a diversidade dos novos casos de previsão. Isso nem sempre é a realidade. Então, para aumentar a capacidade de generalização do modelo, são executados três treinamentos e armazenados seus parâmetros. A previsão final é baseada na média desses três modelos, reduzindo assim o risco de uma amostra de validação inapropriada selecionada.

3.6 Conclusão parcial

Nesse capítulo foi apresentado o método de previsão de geração FV proposto. Devido ao relacionamento cruzado entre os diversos aspectos, optou-se pela

apresentação inicial das principais predefinições e bases hipotéticas do método, a origem dos dados de desenvolvimento e, então, sua apresentação e detalhamento.

O processo sugerido de atualização dos treinamentos (diário nos primeiros seis meses, semanal até completar um ano, e mensal após 12 meses) visa compensar o fato de que inicialmente os dados de treinamento são reduzidos (insuficientes) e que crescem com o passar do tempo até atingir uma escala suficiente para o treinamento mensal da RNA.

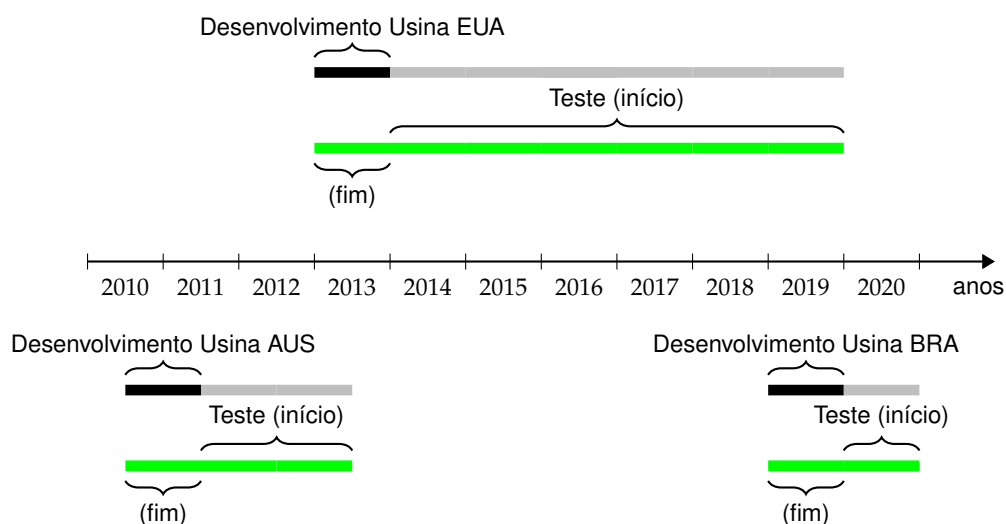
A seguir apresentou-se o modelo MLP proposto de forma completa, para facilitar o engajamento do leitor mais experiente, seguido do detalhamento das decisões tomadas na definição desse modelo, permeado por experimentos práticos e recomendações da literatura. Procurou-se evidenciar o caráter convergente entre os aspectos da estrutura proposta para as três usinas de referência e as razões para o descarte de um processo de regularização, fatos que sugerem a viabilidade de inclusão de novos ativos de geração fotovoltaica na VPP com relativa segurança, a partir de configurações universais.

4 ESTUDO DE CASOS

Neste capítulo são apresentados os resultados da aplicação do método proposto de previsão de geração solar fotovoltaica, conforme o Capítulo 3 (Previsão de Geração Solar Fotovoltaica), às três usinas de referência deste trabalho – AUS, BRA e EUA – apresentadas na seção 3.2 (Prospecção de dados para o desenvolvimento da previsão de geração), cujo detalhamento encontra-se no Apêndice D (Usinas de referência).

A verificação do método de previsão incluiu nos testes os dados utilizados no desenvolvimento do modelo de previsão, para o máximo aproveitamento dos dados disponíveis, mas os períodos iniciais utilizados no desenvolvimento foram deslocados no tempo para o final do período de teste, como mostra a Figura 38, com a expectativa de não comprometer os resultados. Assim, por exemplo, no caso da usina BRA, cujos dados obtidos são limitados em dois anos, os anos de 2019 e 2020 foram considerados temporalmente invertidos, como se um ano como 2020 fosse seguido por outro ano como 2019. Com isso, os dados conhecidos pelo modelo colaboram para a avaliação mas podem ser destacados, bastando para isso um olhar sobre os resultados acumulados até antes do início desses períodos.

Figura 38 – Linha do tempo dos dados de desenvolvimento e teste



Fonte: elaboração própria (2022).

Seguindo o método proposto, uma RNA foi selecionada para o modelo de previsão com as características apresentadas na Tabela 15¹, e um algoritmo foi construído para simular o ciclo de treinamento da RNA ao longo do tempo, como

¹Para a conveniência do leitor, a Tabela 15 reproduz de forma integral a Tabela 12, apresentada na descrição do método (Capítulo 3).

Tabela 15 – Especificação da rede neural artificial de teste

Característica	Valor
Entradas (atributos)	<i>azim, elevref, dayang, hrang, temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h</i>
Normalização	Minmax $\{-1, 1\}$ c/ <i>clip</i> ^a
Saída	Geração de energia (em p.u. ^b)
Tipo de rede	<i>perceptron</i> multicamadas (MLP)
Estrutura de camadas	
Entrada	18 neurônios (11 atributos ^c)
Ocultas	6 camadas de 61 neurônios cada
Saída	1 neurônio
Pesos e vieses	20.131 parâmetros
Método de inicialização	Glorot uniforme c/ vieses zerados
Função de ativação	ELU, exceto o neurônio de saída (função identidade)
Otimizador	Gradiente descendente estocástico (SGD)
Varição	c/ Momento e Nesterov
Taxa de aprendizado	0,1
Taxa de aceleração	0,9
Lote	Variável (mínimo de 8 exemplos) conforme a Equação 13
Função de custo	Huber
Número de execuções	3 (o valor da previsão é a média das execuções)

^a Valores de entrada de teste são truncados nos limites da normalização.

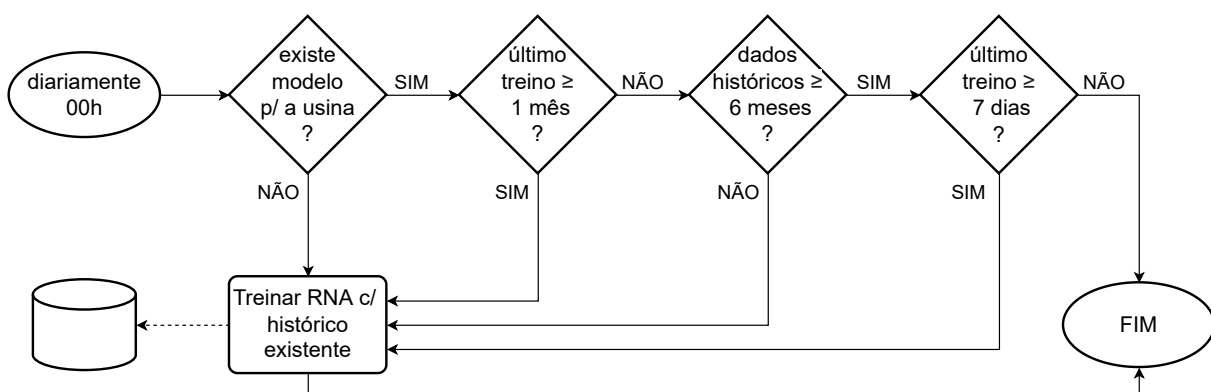
^b A base é a capacidade nominal de cada usina.

^c Os atributos de condições do tempo: *temp, pressure, humidity, clouds, wind_speed* e *wind_deg* são pares de valores relativos ao instante inicial e final de cada hora. O atributo de geoposicionamento solar *day_ang* é desdobrado em duas componentes: seno e cosseno.

Fonte: elaboração própria (2022).

mostra a Figura 39.

Figura 39 – Ciclo de treinamento



Fonte: elaboração própria (2022).

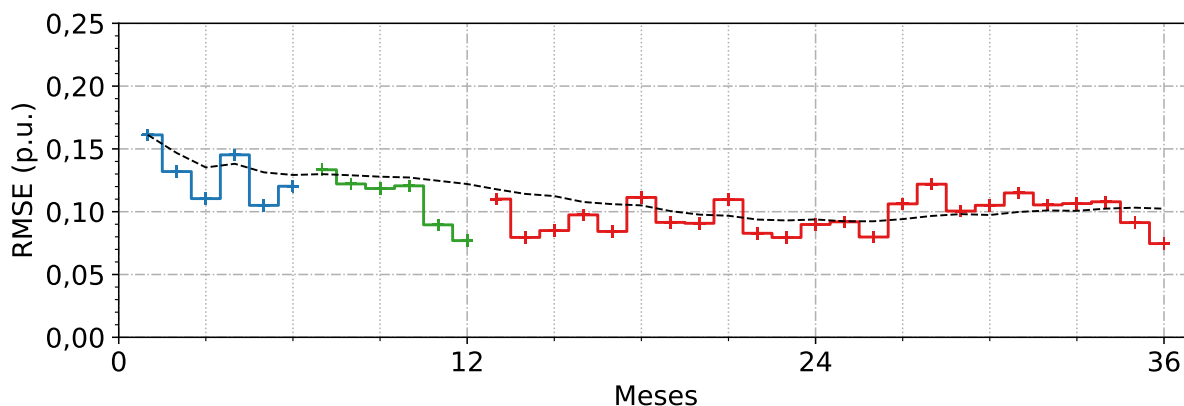
Inicialmente o ciclo de treinamento é diário, até que um volume mínimo e suficiente de dados históricos que permita avançar para um intervalo maior esteja disponível – o que foi definido em seis meses. Na sequência o treinamento passa a ser semanal até completar um ano de dados, o que fornece uma base mínima da sazonalidade da usina. E finalmente, o intervalo regular passa a ser mensal ao longo da vida útil do projeto (presença da usina na VPP).

Em produção, as previsões de geração da VPP são realizadas em intervalo horário, à medida que novas previsões meteorológicas são liberadas pelo fornecedor. Os resultados apresentados neste capítulo não consideram os efeitos dos erros de previsão meteorológica, que serão maiores quanto maior a antecedência da previsão. Busca-se o melhor resultado dada a previsão meteorológica.

4.1 Apresentação dos resultados de teste

Os resultados de teste do modelo desenvolvido aplicado às três usinas de referência são inicialmente apresentados com discretização mensal e identificação da frequência de treinamento, como mostram as Figuras 40 e 41.

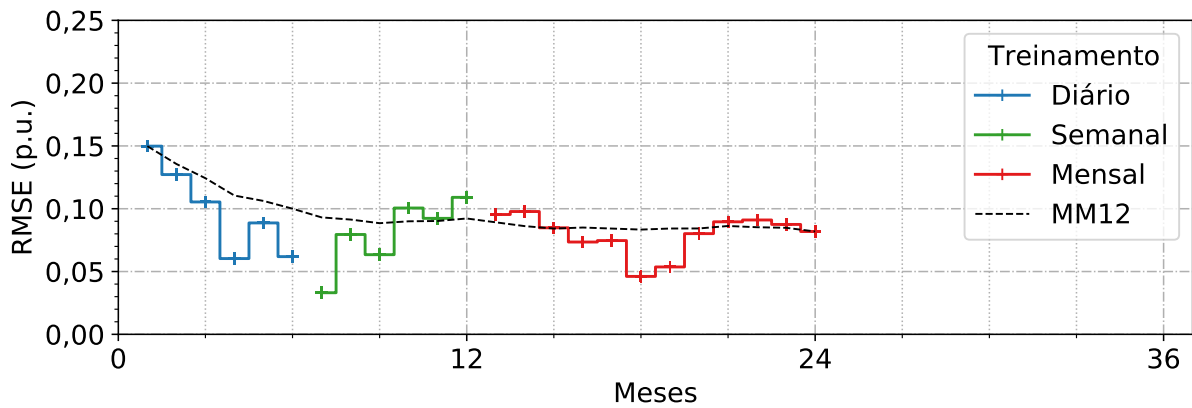
Figura 40 – Resultado dos testes visto pela métrica da RMSE – Usina AUS



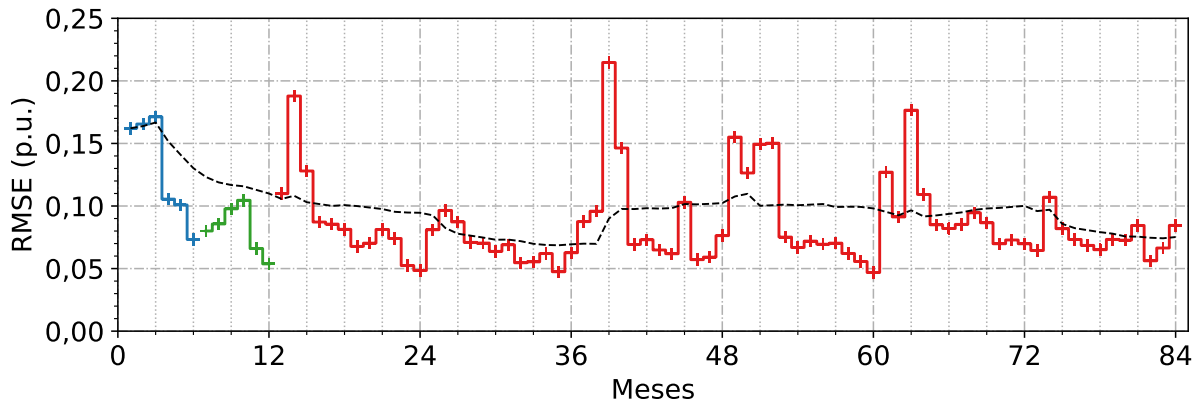
Fonte: elaboração própria (2022).

A curva da média móvel apresentada é de até 12 meses (MM12). Inicialmente os erros são mais elevados, em função do histórico reduzido, mas com a frequência de treinamento diário (linhas azuis), em três ou quatro meses os resultados começam a se tornar comparáveis com alguns daqueles obtidos a partir de amostras de treinamento maiores. A transição para a frequência padrão mensal (linhas vermelhas) se dá pela frequência semanal de treinamento (linhas verdes). Apesar de a RMSE ser uma métrica absoluta, os valores estão apresentados no sistema por unidade, logo, podem ser entendidos como índices relativos à potência nominal das usinas, e comparados entre elas.

Figura 41 – Resultado dos testes visto pela métrica da RMSE



(a) Usina BRA

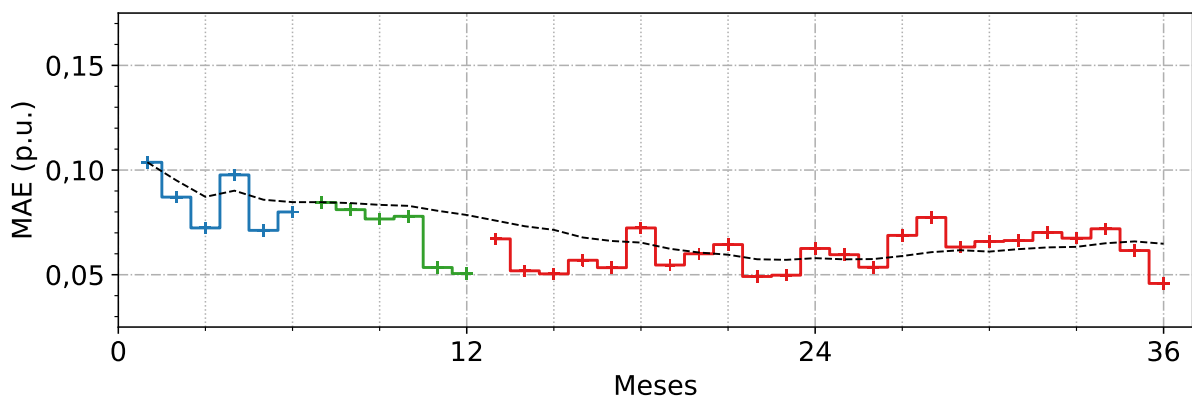


(b) Usina EUA

Fonte: elaboração própria (2022).

Uma visão pela métrica do MAE, apresentada pelas Figuras 42 e 43, permite uma clareza maior sobre os resultados obtidos, uma vez que o MAE é mais intuitivo e menos sensível aos *outliers*.

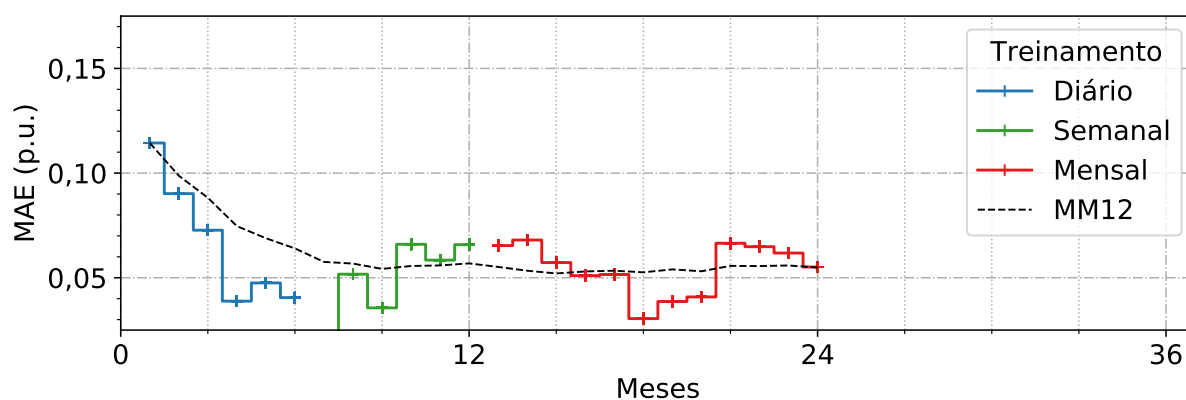
Figura 42 – Resultado dos testes visto pela métrica do MAE – Usina AUS



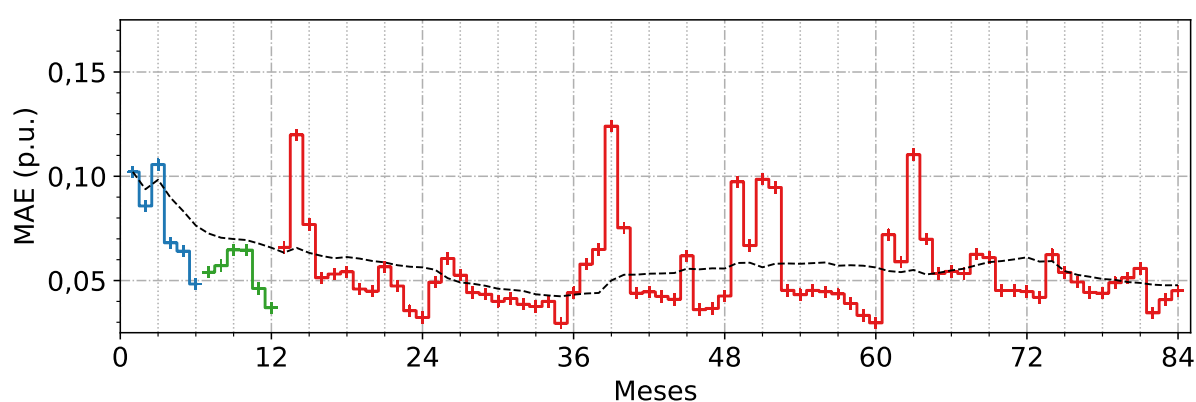
Fonte: elaboração própria (2022).

Para se ter uma ideia da capacidade de representação do comportamento

Figura 43 – Resultado dos testes visto pela métrica do MAE



(a) Usina BRA

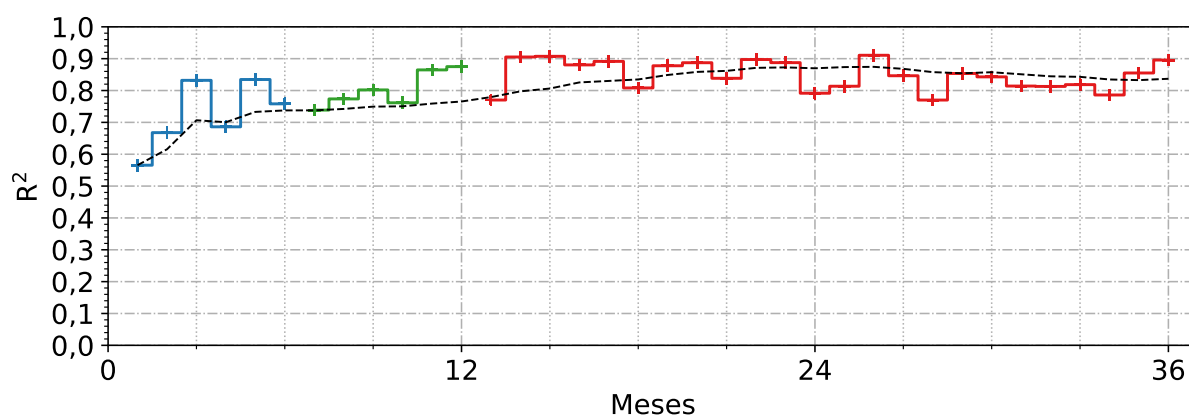


(b) Usina EUA

Fonte: elaboração própria (2022).

real pelo modelo proposto, outra métrica apropriada é o R^2 , apresentada pelas Figuras 44 e 45. Nota-se uma diferença significativa entre os resultados das usinas. A usina

Figura 44 – Resultado dos testes visto pela métrica do R^2 – Usina AUS

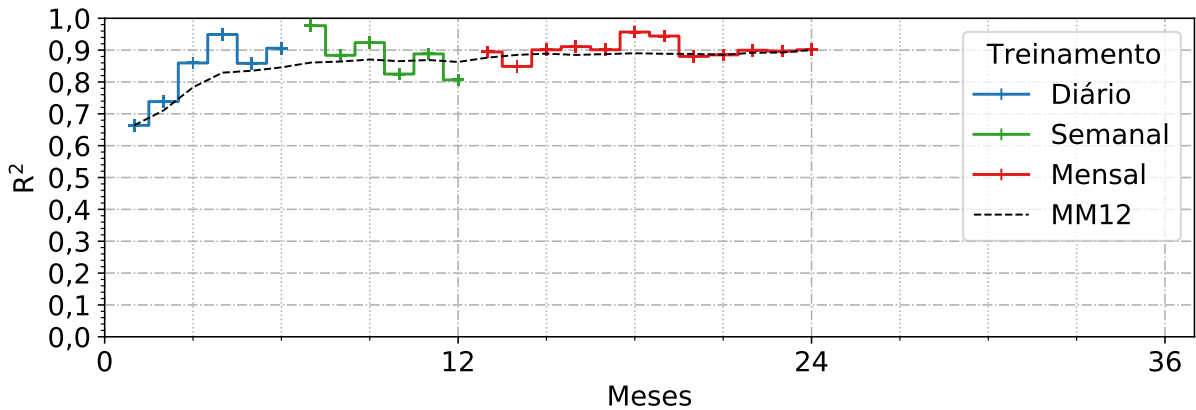


Fonte: elaboração própria (2022).

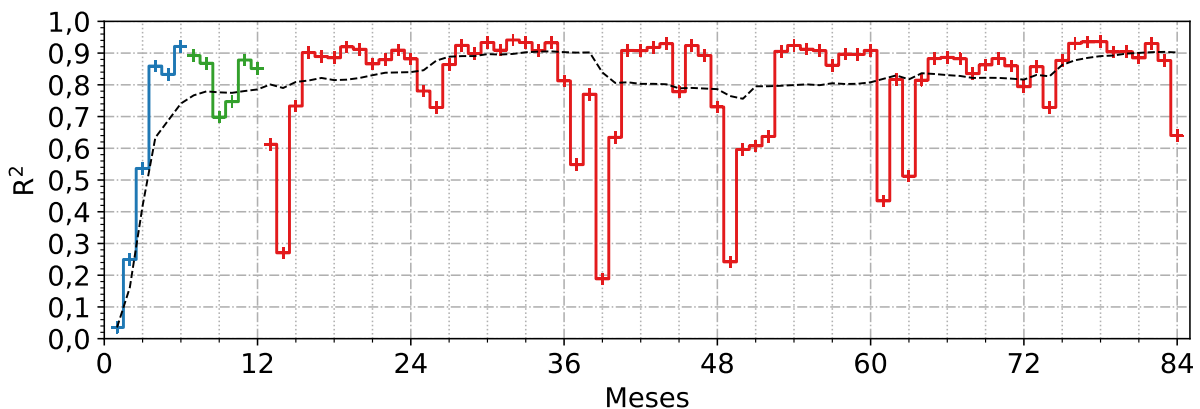
BRA é a melhor representada pelo modelo, o que de certa forma era esperado, à medida que o local de sua instalação, em Bauru/SP, possui cerca de 75% dos dias de

céu aberto ou com até 20% de cobertura de nuvens, uma das variáveis intuitivamente mais importantes para a variabilidade da geração das usinas.

Figura 45 – Resultado dos testes visto pela métrica do R^2



(a) Usina BRA



(b) Usina EUA

Fonte: elaboração própria (2022).

A Tabela 16 apresenta a consolidação dos resultados para as métricas e períodos de teste considerados. No período inicial do teste os resultados são mais sofríveis por conta do volume reduzido de dados para treinamento, como já visto, incapaz de representar o ciclo solar anual. Mas o modelo, que durante o desenvolvimento apresentou resultados melhores de validação cruzada – AUS ($R^2 \approx 0,87$), BRA ($R^2 \approx 0,92$) e EUA ($R^2 \approx 0,93$) – não mantém a capacidade de generalização esperada, principalmente no caso da usina EUA, após 12 meses de teste².

As métricas apresentadas até este ponto fornecem valores objetivos a cerca dos resultados gradativos ou consolidados obtidos, mas as Figuras 46, 47 e 48 apresentam uma visão do significado real dessas previsões, ao plotar as curvas das horas claras de alguns dias das usinas AUS, BRA e EUA, respectivamente, no período de estabilidade da previsão (após os primeiros 12 meses). São dez exemplos de cada

²Em períodos menores a degradação é esperada, pela baixa diversidade no treinamento.

Tabela 16 – Resultados consolidados dos testes

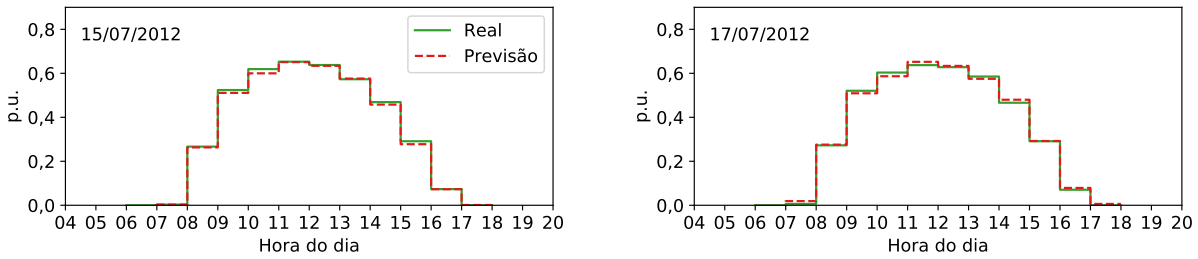
Período	Métrica ^a	Unidade	Usina AUS	Usina BRA	Usina EUA
Até 6 meses	Amostra	meses	6	6	6
	RMSE	p.u.	0,1292	0,1000	0,1302
	nRMSE	%	45,9	32,6	55,2
	MAE	p.u.	0,0847	0,0640	0,0764
	nMAE	%	30,0	20,8	32,4
	MBE	p.u.	-0,0114	-0,0058	-0,0015
	nMBE	%	-4,0	-1,9	-0,6
	R ²	adimensional	0,7376	0,8458	0,7411
6 a 12 meses	Amostra	meses	6	6	6
	RMSE	p.u.	0,1140	0,0849	0,0842
	nRMSE	%	40,6	28,9	37,5
	MAE	p.u.	0,0721	0,0507	0,0548
	nMAE	%	25,7	17,3	24,4
	MBE	p.u.	0,0029	0,0033	0,0030
	nMBE	%	1,0	1,1	1,3
	R ²	adimensional	0,7950	0,8787	0,8489
Após 12 meses	Amostra	meses	24	12	72
	RMSE	p.u.	0,0982	0,0819	0,0908
	nRMSE	%	33,7	25,6	38,4
	MAE	p.u.	0,0613	0,0549	0,0534
	nMAE	%	21,0	17,1	22,6
	MBE	p.u.	0,0083	-0,0149	-0,0113
	nMBE	%	2,8	-4,6	-4,8
	R ²	adimensional	0,8539	0,9006	0,8443

^a As métricas nRMSE, nMAE e nMBE representam a normalização das métricas RMSE, MAE e MBE, respectivamente, pela geração média medida do período.

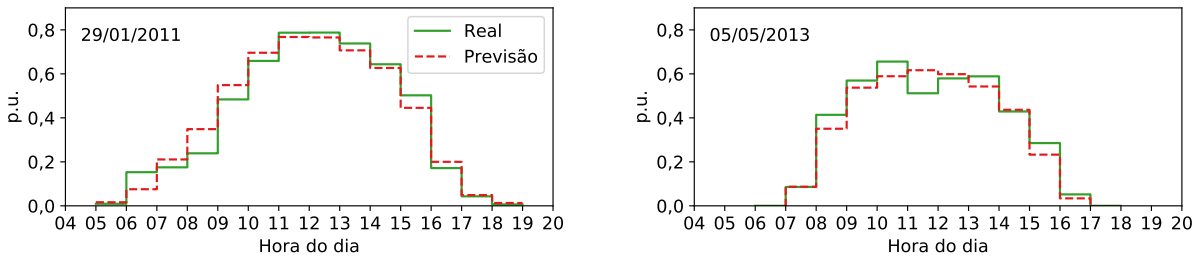
Fonte: elaboração própria (2022).

usina, classificados em cinco grupos de acordo com a posição no ranking crescente da RMSE diária. Ou seja, no caso da usina AUS, por exemplo, a Figura 46a apresenta os dois melhores resultados para um mesmo dia; 25% dos casos possuem previsões piores que essas, porém, iguais ou melhores que as observadas na Figura 46b; e assim sucessivamente, com os piores resultados diários apresentados pela Figura 46e.

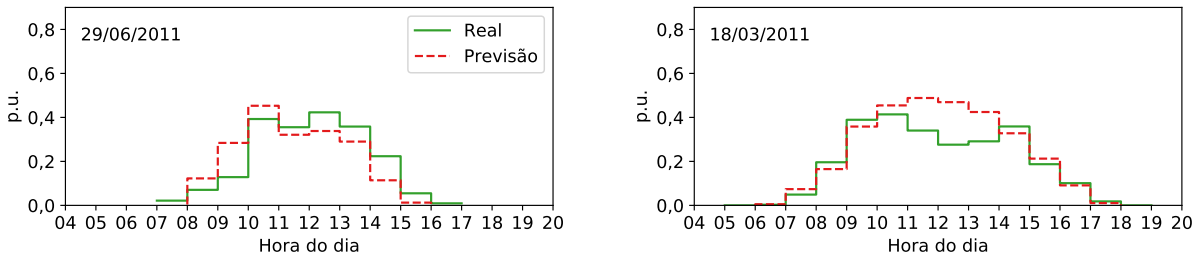
Figura 46 – Curvas selecionadas de geração real vs. prevista – Usina AUS



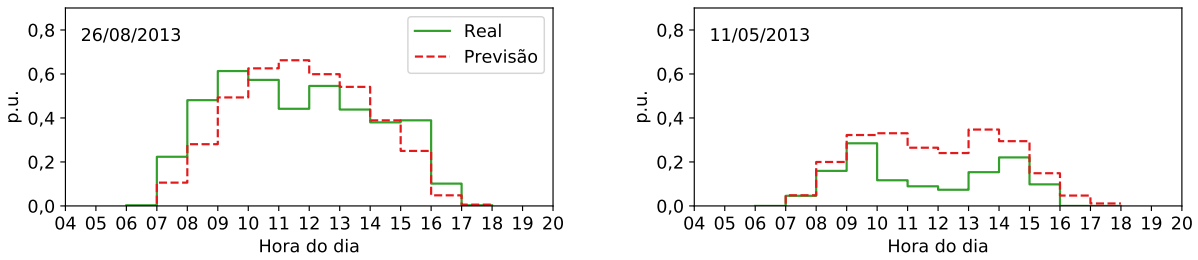
(a) Melhores exemplos com base na RMSE



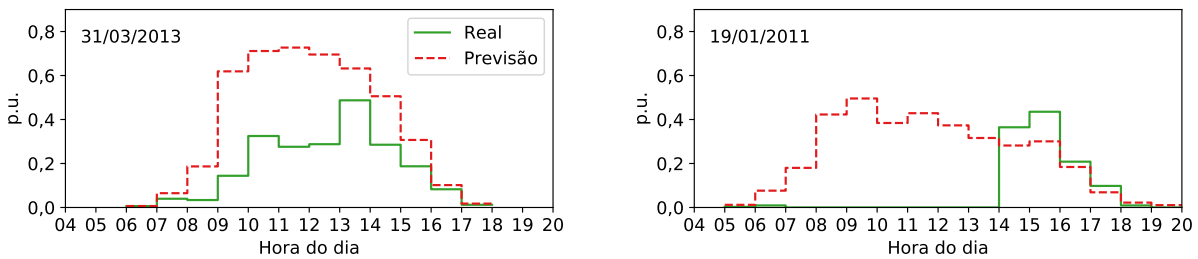
(b) Exemplos do 1º quartil (RMSE)



(c) Exemplos do 2º quartil (RMSE)



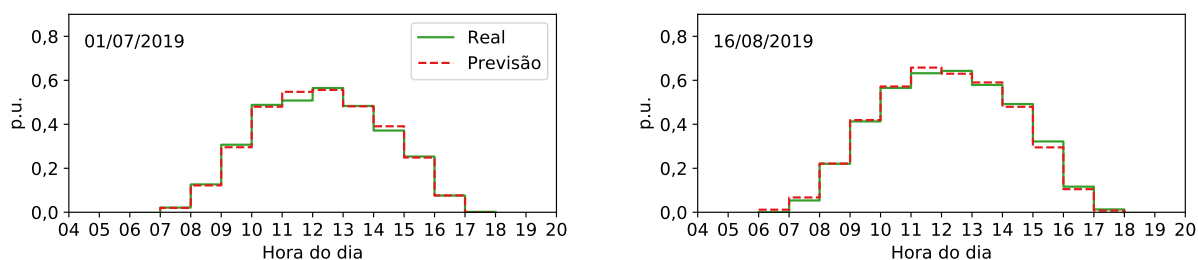
(d) Exemplos do 3º quartil (RMSE)



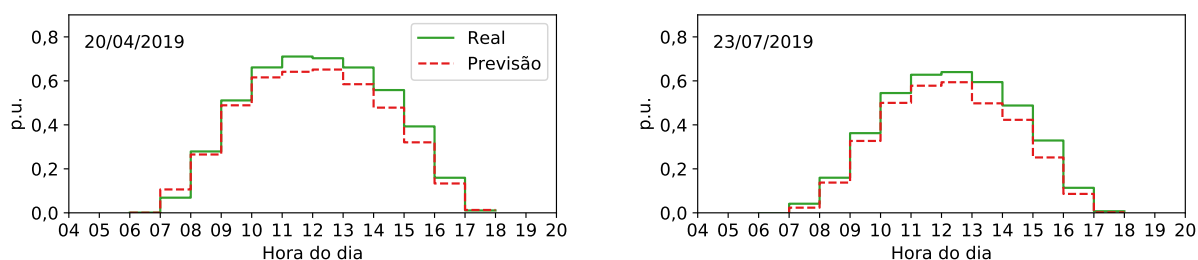
(e) Piores exemplos com base na RMSE

Fonte: elaboração própria (2022).

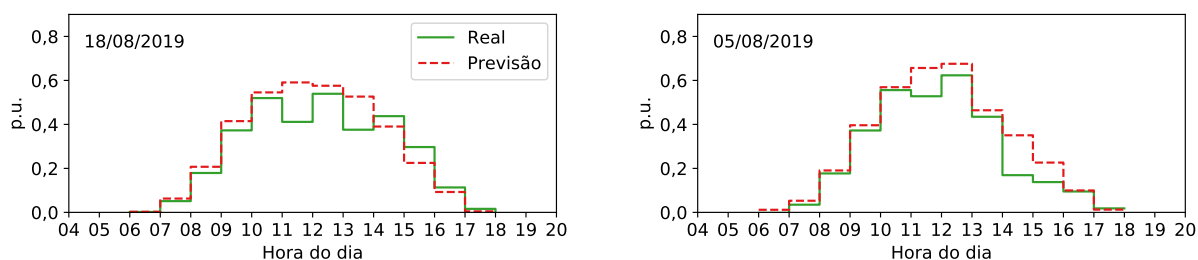
Figura 47 – Curvas seleccionadas de geração real vs. prevista – Usina BRA



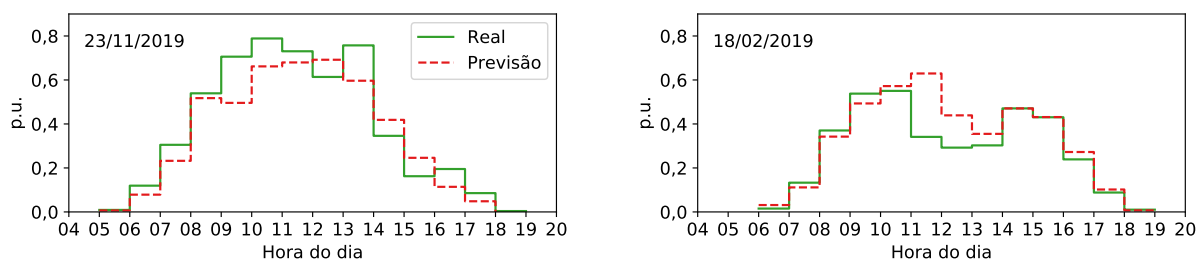
(a) Melhores exemplos com base na RMSE



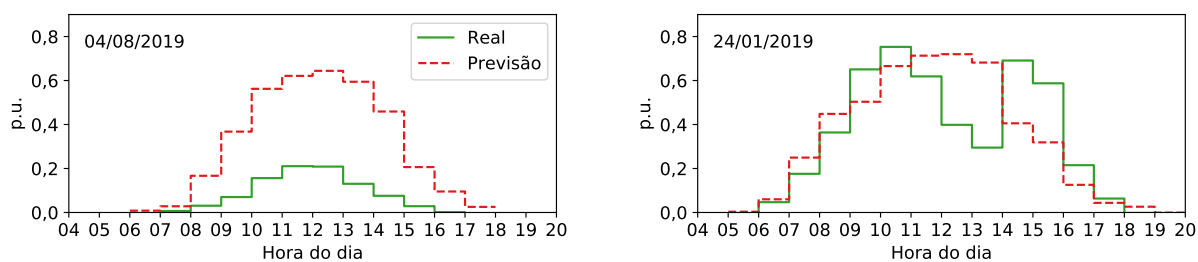
(b) Exemplos do 1º quartil (RMSE)



(c) Exemplos do 2º quartil (RMSE)



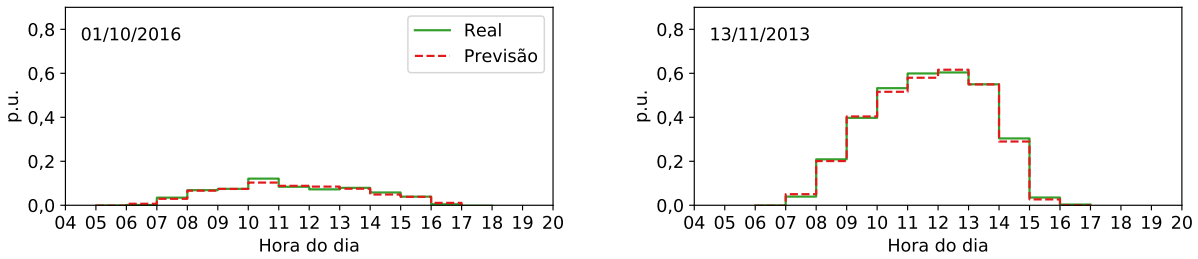
(d) Exemplos do 3º quartil (RMSE)



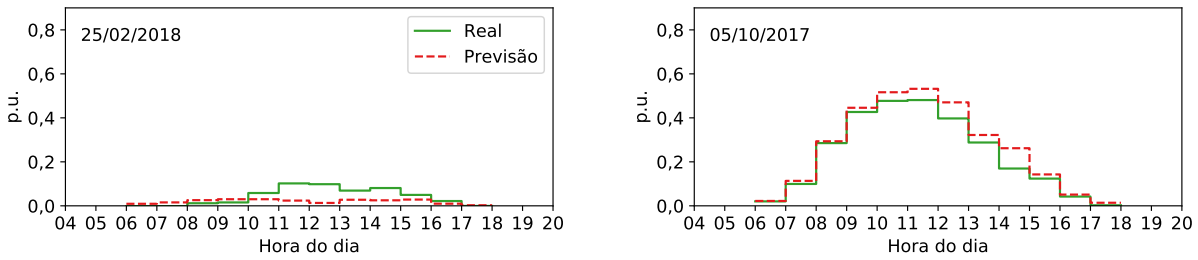
(e) Piores exemplos com base na RMSE

Fonte: elaboração própria (2022).

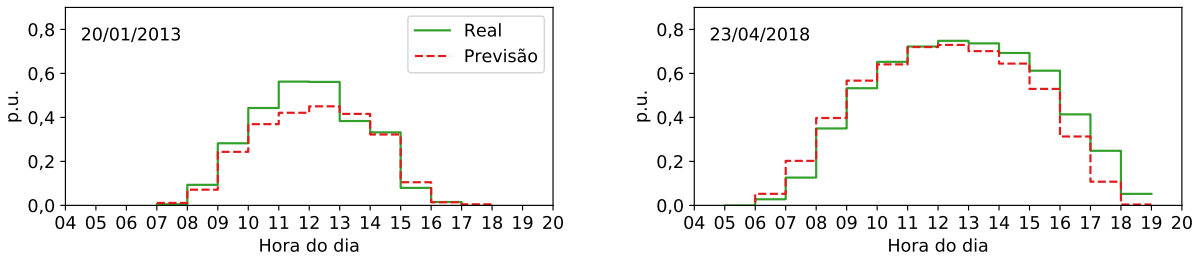
Figura 48 – Curvas selecionadas de geração real vs. prevista – Usina EUA



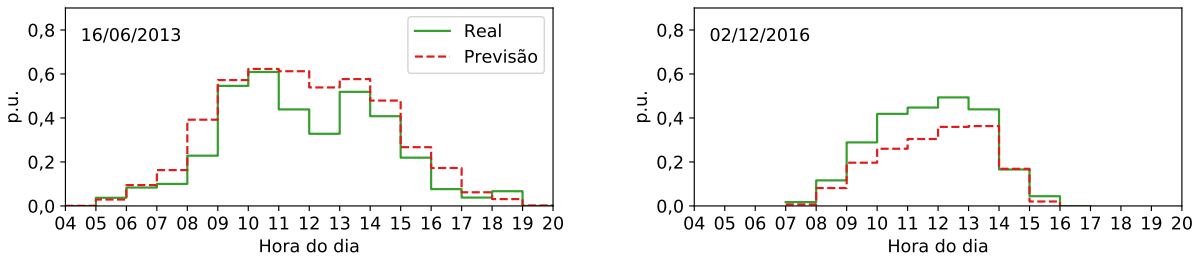
(a) Melhores exemplos com base na RMSE



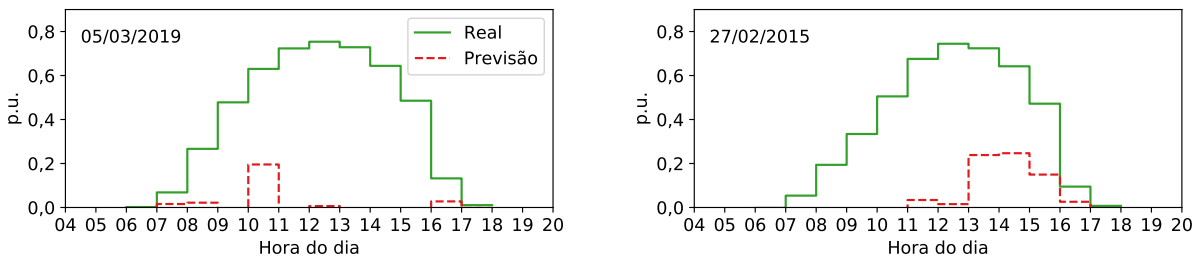
(b) Exemplos do 1º quartil (RMSE)



(c) Exemplos do 2º quartil (RMSE)



(d) Exemplos do 3º quartil (RMSE)



(e) Piores exemplos com base na RMSE

Fonte: elaboração própria (2022).

Finalmente, vale a pena lembrar que uma VPP engloba um *pool* de usinas localizadas em diferentes regiões, e beneficia-se da redução dos erros de previsão – agregada – de sua carteira em relação aos valores individuais das usinas (NAVAL; YUSTA, 2021). Para ilustrar esse aspecto, a Tabela 17 apresenta os principais resultados de previsão individuais e combinados das usinas para os primeiros 12 meses do período estável (após ao menos 12 meses de histórico para treinamento), supondo que todas estivessem conectadas a uma mesma rede elétrica; o fuso-horário é considerado comum, ou seja, as horas locais são mantidas e equiparadas; e os ângulos diários solares são ajustados para equiparação entre os hemisférios Norte e Sul. Evidencia-se que, com o agrupamento de usinas, as métricas de qualidade das previsões agregadas são melhores, com destaque para o R^2 , o qual denota uma redução maior da variância das previsões agregadas do que da variância da geração real agregada, dada sua formulação (Equação 8).

Tabela 17 – Resultados do *pool* de usinas para 12 meses de previsão

Métrica ^a	Unidade	Usina AUS	Usina BRA	Usina EUA	VPP
Potência Nominal	kWp	2,04	95	140,14	237,18
MAE	p.u.	0,0579	0,0549	0,0563	0,0353
nMAE	%	19,4	17,1	22,9	15,7
R^2	adimensional	0,8698	0,9006	0,8400	0,9295

^a A métrica nMAE representa a normalização da métrica MAE pela geração média do período.

Fonte: elaboração própria (2022).

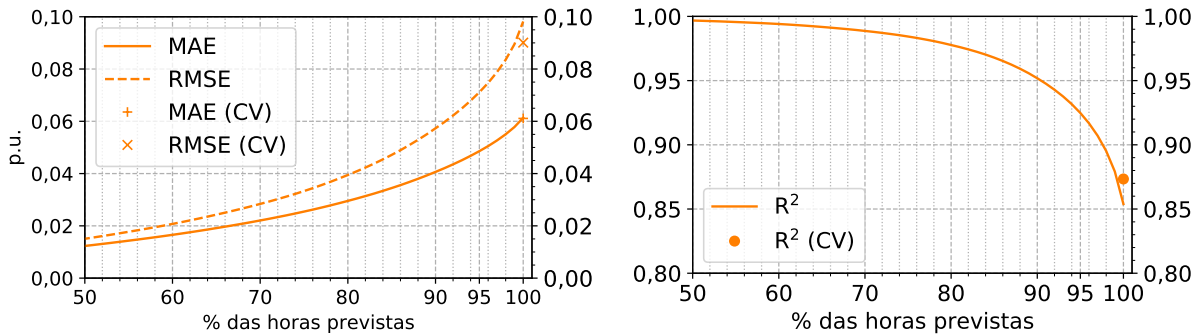
4.2 Análise dos resultados de teste

Para o planejamento das microrredes e plantas virtuais, a proporção de boas previsões pode ser tão ou mais útil que uma boa média delas. A VPP está sujeita a uma série de incertezas, que incluem o preço dos ativos energéticos, o consumo das unidades consumidoras, o gerenciamento das fontes de geração despacháveis (combustível, manutenções corretivas etc.), dentre outros fatores. Ter a prevalência de acertos simultâneos desses aspectos pode trazer resultados mais interessantes que uma boa média, ou em outras palavras, a mediana pode ser mais importante que a média dos erros.

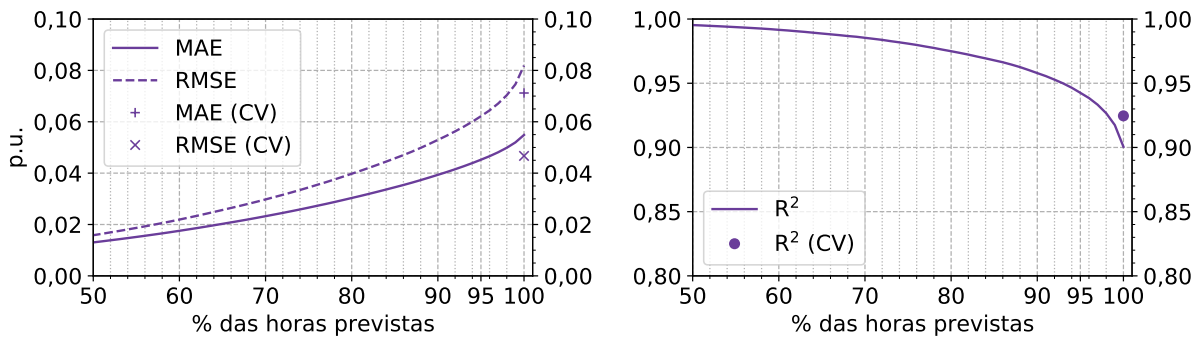
A Figura 49 mostra uma visão acumulada do MAE, RMSE e R^2 em função da % das horas previstas para as usinas AUS, BRA e EUA em ordem crescente dos erros absolutos. Há ainda uma indicação dos erros equivalentes observados durante o desenvolvimento, em validação cruzada (CV). Nessa visão, pode-se perceber que um número relativamente pequeno, da ordem de 2% das horas, interfere significativamente

nos resultados médios das usinas AUS e BRA. Para essas usinas, pelo lado do copo meio cheio, excluía essa parcela, os resultados são aderentes, indicando boa generalização.

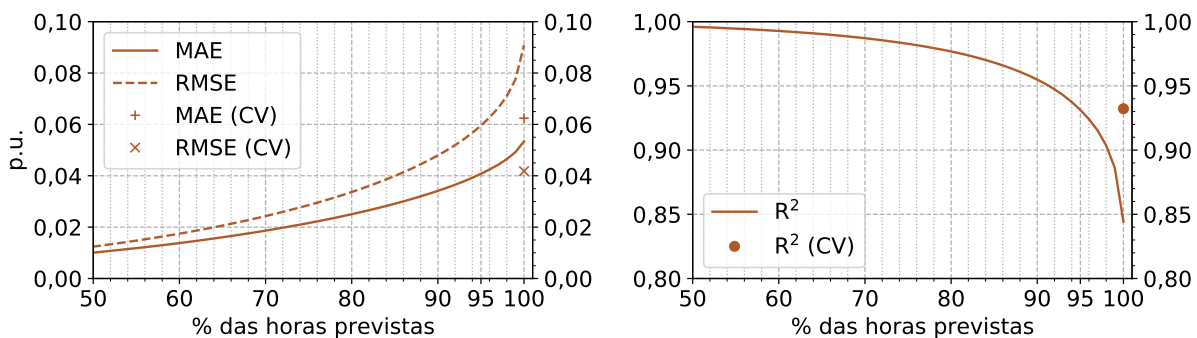
Figura 49 – Resultado dos testes por ordem do erro absoluto



(a) Usina AUS



(b) Usina BRA



(c) Usina EUA

Fonte: elaboração própria (2022).

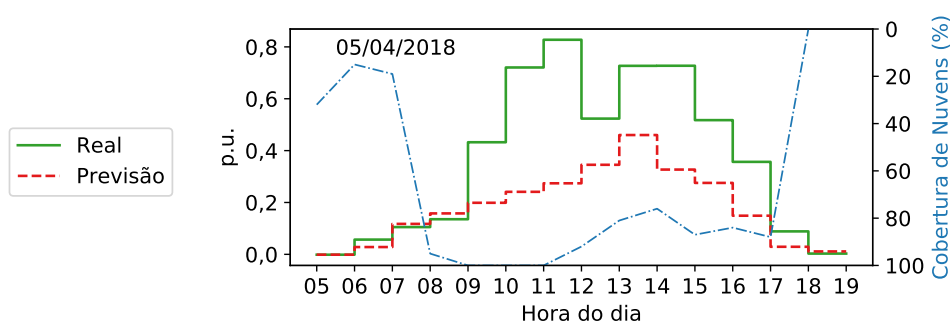
Esse não é um contorcionismo com os dados, pois, avaliando-se alguns dos pontos de maior desvio da usina EUA – mais dissonante – foi possível perceber, mais facilmente, algumas exceções para as quais não se pode esperar resposta deste – e talvez de qualquer outro – modelo, porque ele não foi preparado para isto, como por exemplo:

- a) um período contíguo de 15 dias entre 22/03/2017 (9h) e 05/04/2017 (13h) em que a geração medida nas horas claras foi equivalente à noturna

(valor fixo e negativo, devido ao consumo do inversor). A investigação dos índices de neve e da temperatura ambiente, antes e durante o evento, bem como o retorno súbito à normalidade, indicam uma maior probabilidade de defeito técnico do sistema, uma vez que houveram vários intervalos de céu semi-aberto e temperatura ambiente acima de 10°C;

- b) no dia 05/04/2018 houve descolamento entre a previsão de geração e o valor medido, como mostra a Figura 50. No entanto, a previsão é mais aderente ao comportamento da cobertura de nuvens daquele dia, visto que com 100% de cobertura de nuvens entre 9h e 11h não se espera uma geração média em torno de 0,6 p.u. Isso lembra que os dados meteorológicos não são reais (medidos), mas calculados (interpolados) pelo fornecedor a partir de dados reais das estações meteorológicas próximas e imagens de satélite. Nessa data, podem estar incorretos.

Figura 50 – Descolamento da geração real vs. cobertura de nuvens - Usina EUA

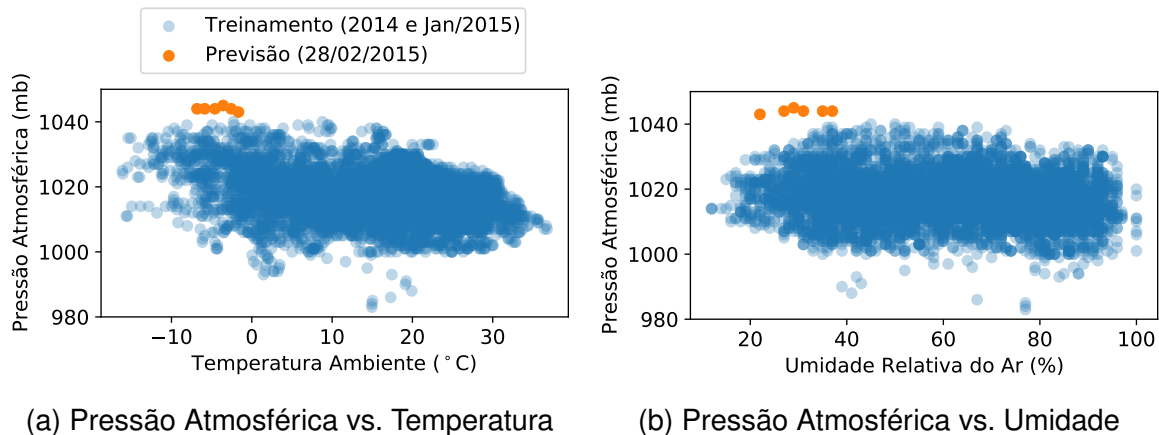


Fonte: elaboração própria (2022).

Além desses eventos causarem baixa performance de previsão durante sua ocorrência, eles interferem no aprendizado da rede neural para as previsões futuras, o que indica a necessidade de um processo de validação dos dados de entrada para treinamento que considere dados de indisponibilidade da usina, por exemplo, o que não se dispõe para o presente trabalho.

Durante o estudo de caso, dois tipos de falha ou incapacidade do método foram encontrados e motivaram uma alteração da metodologia. Por exemplo, no dia 28/02/2015 foram observados valores de previsão de geração superiores (até 20%) à própria potência nominal da usina EUA³. A Figura 51 mostra como a pressão atmosférica daquele dia atingiu picos (1043 mb a 1045 mb) não experimentados na amostra de treinamento.

³Durante o desenvolvimento optou-se por não limitar a saída da RNA em função de ausência de indicadores – ocorrências – do problema, e também porque não há garantia de que a potência nominal conhecida (base do sistema por unidade) será sempre a correta.

Figura 51 – Dados selecionados de treinamento e previsão – Usina EUA

Fonte: elaboração própria (2022).

A avaliação das condições gerais mostraram – por convicção do autor – que o problema foi causado pelo histórico de treinamento insuficiente. No dia seguinte, com pressão atmosférica abaixo de 1040 mb, temperaturas e umidades relativas do ar semelhantes, os valores de geração previstos são normais. Este é um caso em que mesmo uma frequência maior de treinamento não seria capaz de lidar, uma vez que a pressão atmosférica não passou de 1030 mb até o dia anterior, naquele mês.

O mesmo problema de insuficiência de histórico, o que demanda a extrapolação pela RNA, foi observado para o dia 05/06/2013 às 10h, onde o índice de chuva foi de 46,99 mm (pico de todo o banco de dados) e resultou numa previsão de geração da ordem de 1,48 p.u. – acima da unidade (potência nominal) – para uma geração real de 0,676 p.u. Anote-se que apesar de seis anos de histórico de treinamento da usina EUA. Este caso e o anterior foram resolvidos a partir da truncagem dos valores normalizados dos atributos de entrada entre -1 e 1 , com benefício marginal aos resultados globais.

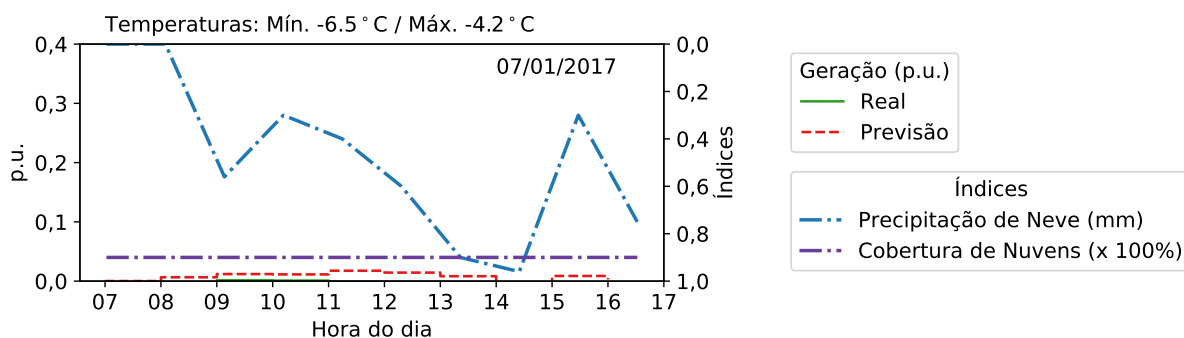
Outro problema – ilustrado pela Figura 52 – parece ser causado pela ausência de memória das ocorrências passadas. No dia 07/01/2017 houve precipitação de neve e o céu estava bem encoberto; a previsão foi aderente à realidade naquele dia. No entanto, apesar da abertura do céu, a geração se manteve restrita pelos próximos dois dias, o que pode ter decorrido da obstrução dos painéis solares pela neve, uma vez que a temperatura ambiente permaneceu muito baixa ao longo de todo o período.

Para eventualmente identificar características não aprendidas pelo modelo da RNA, os erros de previsão foram plotados para o período estável das previsões – após os primeiros 12 meses⁴ – em função das estações do ano⁵ e de alguns atributos⁶: hora do dia (em substituição ao ângulo horário solar), azimute, temperatura ambiente,

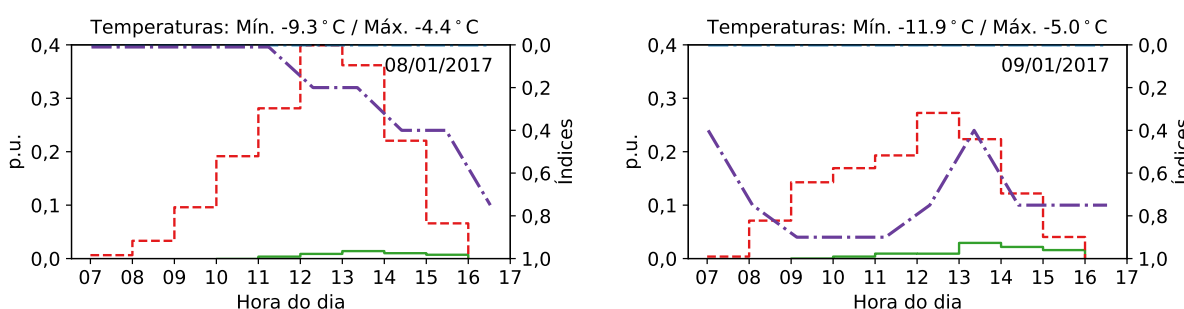
⁴Os dados integram dois, um e seis anos de dados, respectivamente, das usinas AUS, BRA e EUA.

⁵As estações do ano estão relacionadas com o ângulo diário solar, mas são mais intuitivas.

⁶Algumas escalas dos atributos foram adaptadas para melhor intuição do leitor.

Figura 52 – Caso de sombreamento pela queda de neve – Usina EUA

(a) Dia de queda de neve



(b) 1º dia seguinte

(c) 2º dia seguinte

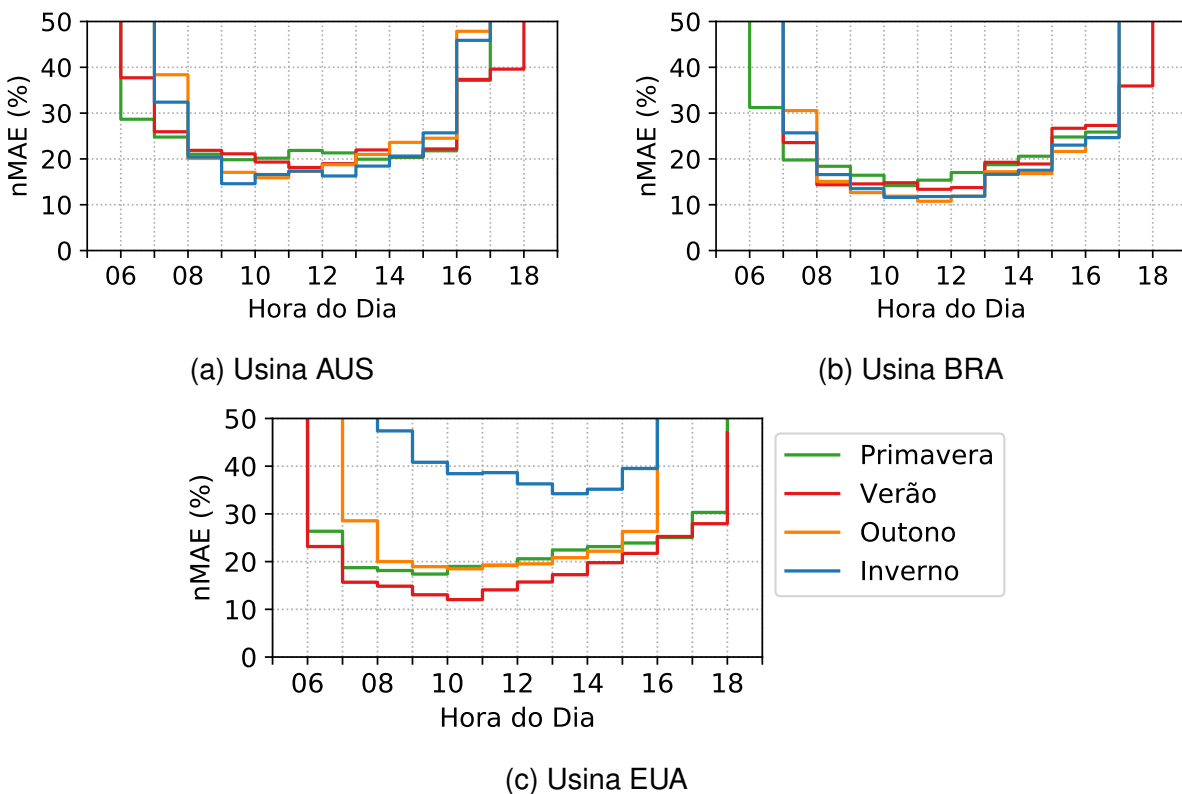
Fonte: elaboração própria (2022).

pressão atmosférica, umidade relativa do ar, cobertura de nuvens, velocidade dos ventos, direção dos ventos, índice de chuva e índice de neve. Antes de apresentá-los, é interessante observar, a partir da experiência ao longo dos testes, que as métricas utilizadas neste trabalho podem variar – para a mesma estrutura RNA – em função da diversidade da amostra de treinamento, pois isso afeta a capacidade de representação de novos eventos. Essa diversidade está relacionada, em geral, com o número de exemplos e sua distribuição ao longo do ciclo solar diário (horas do dia) e anual (estações do ano). Além disso, um sistema de previsão que englobe as 24 horas do dia quase certamente apresentará RMSE, MAE e MBE muito abaixo daqueles vistos neste trabalho – onde apenas as horas claras são tratadas –, por efeito de diluição. A RNA aprenderia facilmente o ponto de corte a partir do qual os valores de geração de energia são sempre iguais (zerados, ou negativos e fixos) e os erros seriam baixos nas horas escuras. Mas também, se as horas forem restritas às de maior insolação, tem-se o mesmo efeito de melhoria de algumas métricas de previsão normalizadas, em função de uma geração média maior.

A Figura 53 ilustra esse aspecto ao apresentar as curvas do MAE normalizado (nMAE) por estação do ano e hora do dia. Observa-se que, proporcionalmente à produção média de cada hora as previsões são mais assertivas para as horas de maior geração de energia. Como a otimização do modelo é baseada no MSE, erros absolutos

iguais têm o mesmo peso – valor ou custo – em qualquer hora do dia, independente da geração média, o que desequilibra os valores dos erros normalizados entre as horas. Ressalta-se ainda a forte influência da reflexão na geração total da usina nas horas iniciais e finais do dia, ocasiões de pouca ou nenhuma irradiação direta sobre os painéis solares. Acredita-se que os atributos utilizados neste trabalho não sejam suficientes para representar a variabilidade do efeito da reflexão sobre a geração de energia das usinas.

Figura 53 – Resultado da previsão por estação do ano e hora do dia

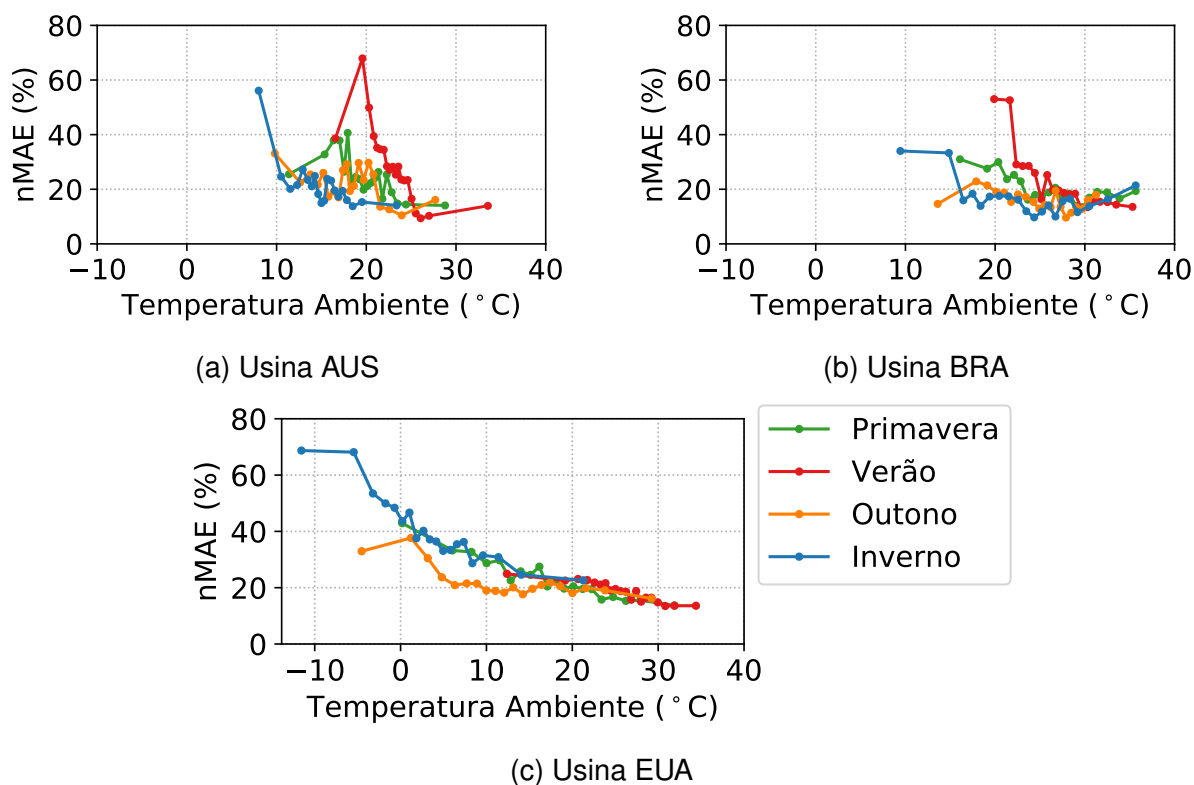


Fonte: elaboração própria (2022).

As Figuras 54 a 59 apresentam o nMAE atributo a atributo, e em função de 20 grupos baseados em percentis equidistantes. A exceção foi para o atributo Índice de Chuva e a informação complementar sobre o Índice de Neve, que foram tratados em 50 grupos. Portanto, em geral, cada ponto representa 5% dos dados (ou 2% para as exceções) e está posicionado na média da subamostra de dados. Pontos mais distanciados no eixo das abscissas evidenciam faixas mais largas ou ausência de registros intermediários.

O modelo é menos assertivo para temperaturas mais baixas (Figura 54). Isso pode ser – em parte – decorrência do desbalanço entre os números de exemplos de treinamento nessa e outras faixas de temperatura. Outro aspecto é que as temperaturas mais altas em cada dia coincidem com as horas de maior produção de energia, com efeito sobre a normalização e, portanto, melhores erros relativos.

Figura 54 – Resultado da previsão por estação do ano e temperatura ambiente

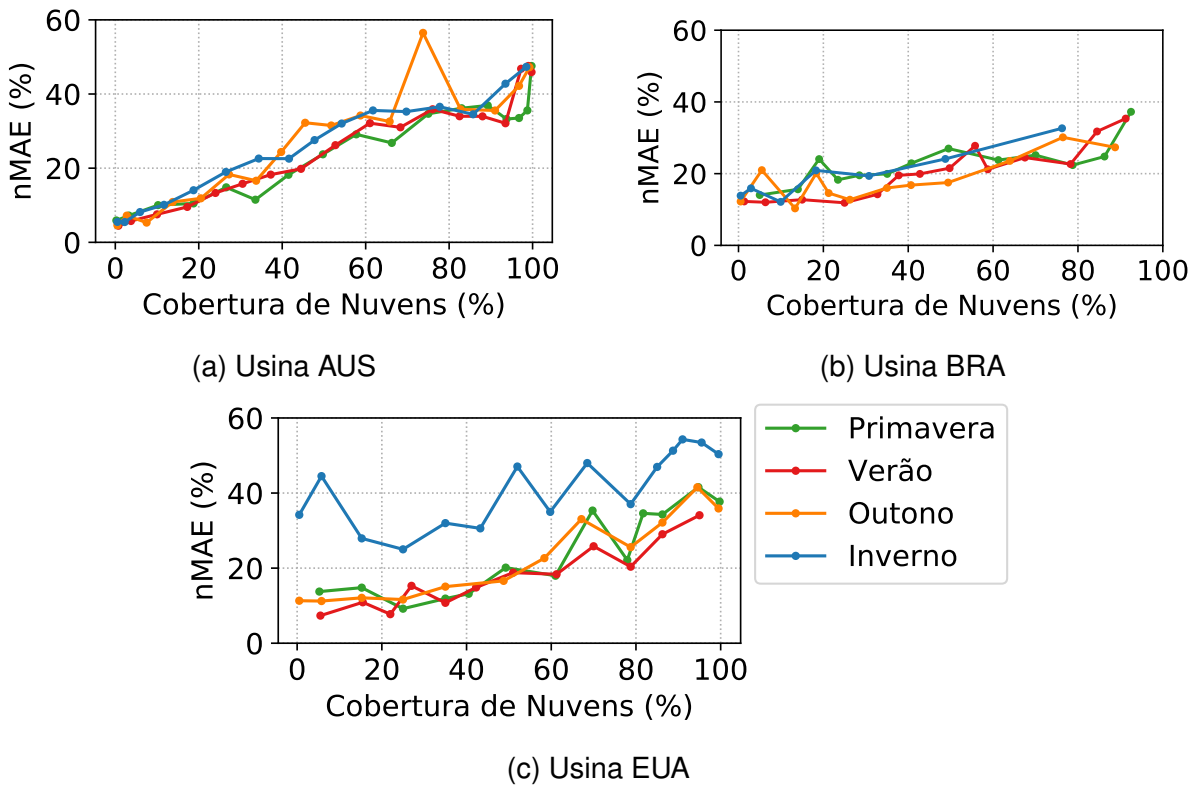


Fonte: elaboração própria (2022).

A cobertura de nuvens (Figura 55) é o atributo de interpretação mais clara, quanto maior o percentual de cobertura de nuvens, maior a dificuldade de representação do modelo; algo intuitivo, uma vez que para duas ocasiões de mesma cobertura de nuvens, mantidos os demais atributos iguais, essas nuvens podem ter diferentes composições e estar distribuídas em diferentes altitudes e posições na abóboda celeste, capazes de afetar a irradiação solar direta e difusa de formas diferentes.

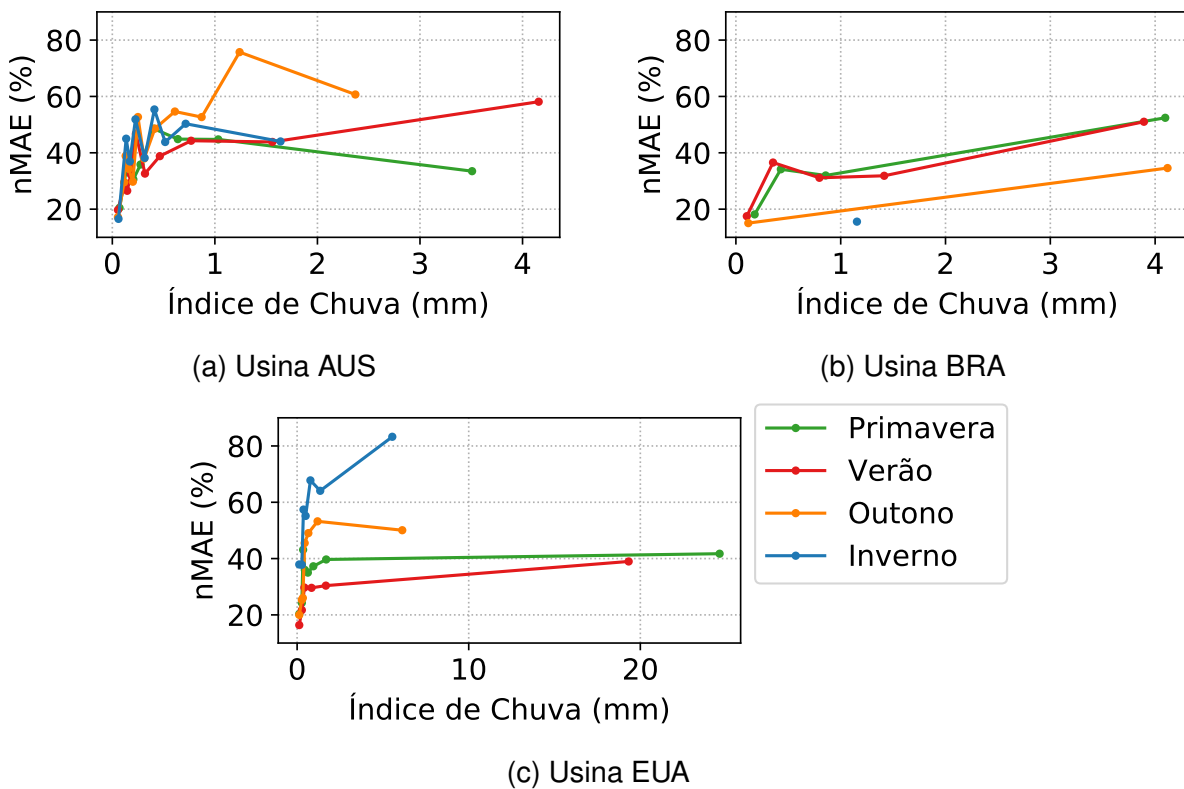
As chuvas (Figura 56), principalmente de outono e inverno – quando são menos comuns – estão associadas a resultados mais sofríveis. É realmente intuitivo como a instabilidade nesses dias não pode ser representada por um atributo simples de volume de água. E talvez, este seja um dos atributos meteorológicos mais difíceis de estimar por interpolação de dados das estações meteorológicas e imagens de satélite. O mesmo pode ser dito quanto à neve na usina EUA (Figura 57).

Figura 55 – Resultado da previsão por estação do ano e cobertura de nuvens



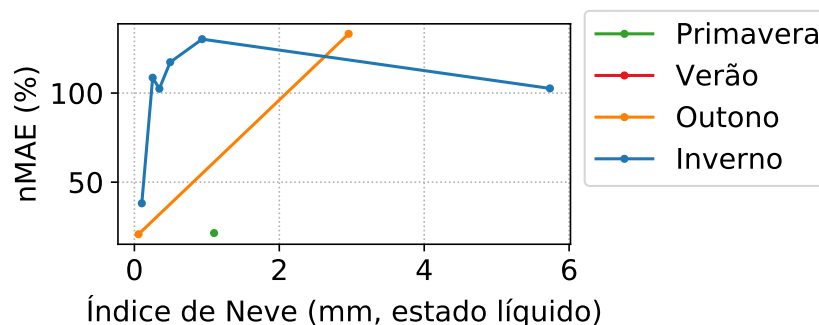
Fonte: elaboração própria (2022).

Figura 56 – Resultado da previsão por estação do ano e índice de chuva



Fonte: elaboração própria (2022).

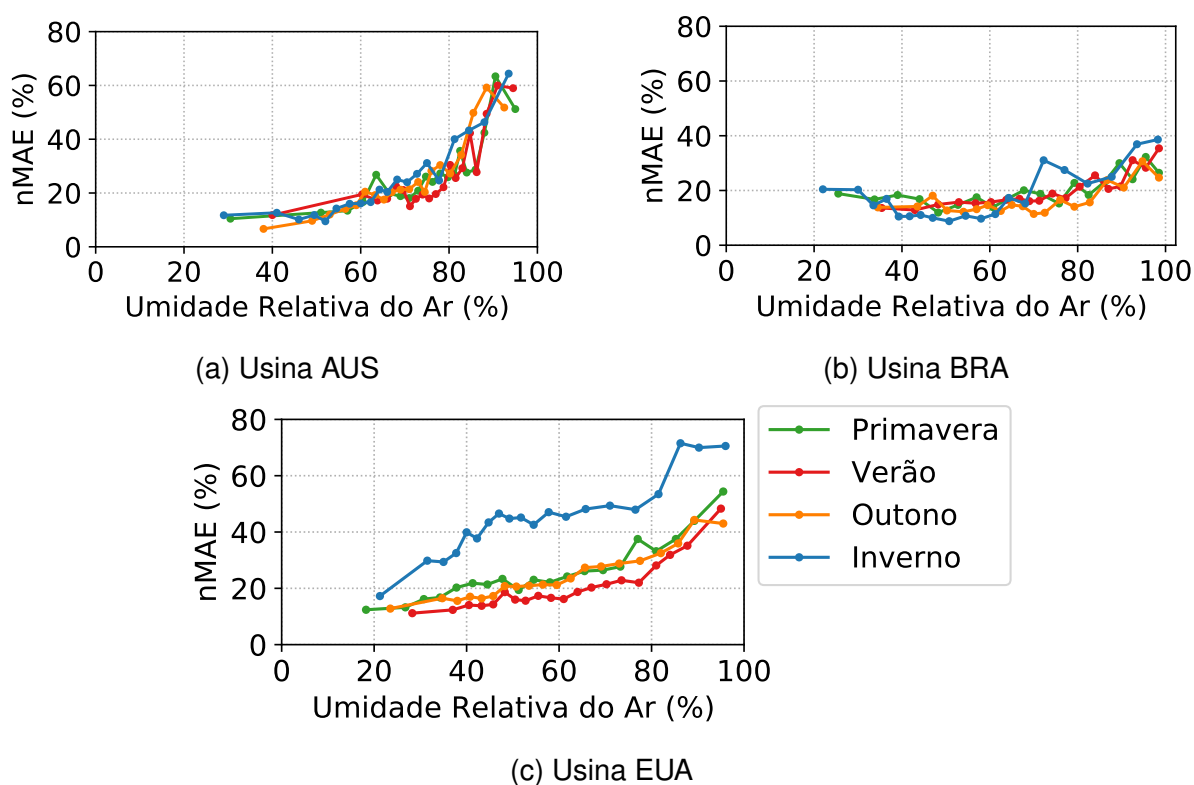
Figura 57 – Resultado da previsão por estação do ano e índice de neve – Usina EUA



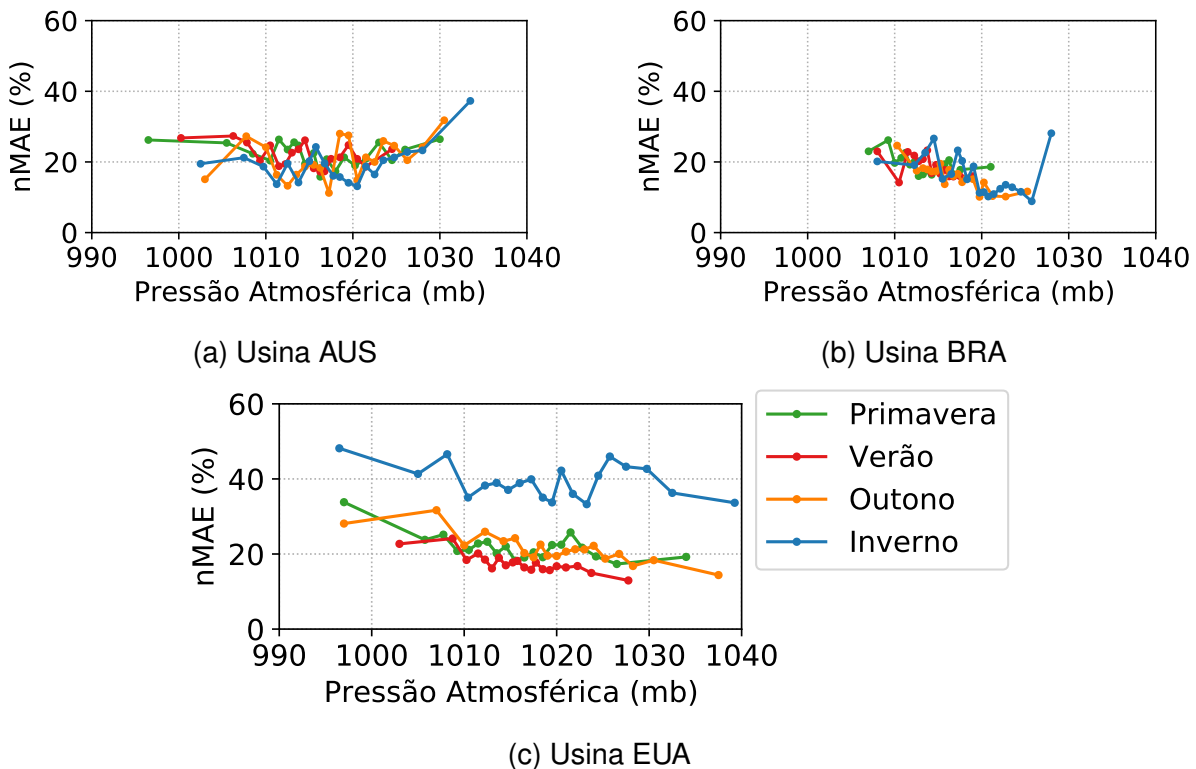
Fonte: elaboração própria (2022).

A umidade relativa do ar (Figura 58) é um atributo muito correlacionado – inversamente – com a temperatura ambiente, como visto na Figura 20, o que provavelmente justifica a tendência de erros maiores para umidades relativas do ar maiores. Ademais, eventos de umidade relativa do ar elevadas estão associados à presença de chuvas ou ao aumento da nebulosidade, dois fatores que, como visto, dificultam a representação pelo modelo. No caso da pressão atmosférica (Figura 59), não há uma tendência única para as três usinas. Alguns erros relativos maiores nos extremos podem estar associados a um número menor de observações.

Figura 58 – Resultado da previsão por estação do ano e umidade relativa do ar



Fonte: elaboração própria (2022).

Figura 59 – Resultado da previsão por estação do ano e pressão atmosférica

Fonte: elaboração própria (2022).

Finalmente, os atributos de azimute, velocidade e direção dos ventos não apresentaram qualquer tendência identificável pelo autor, razão pela qual não são apresentados.

4.3 Considerações sobre outros métodos

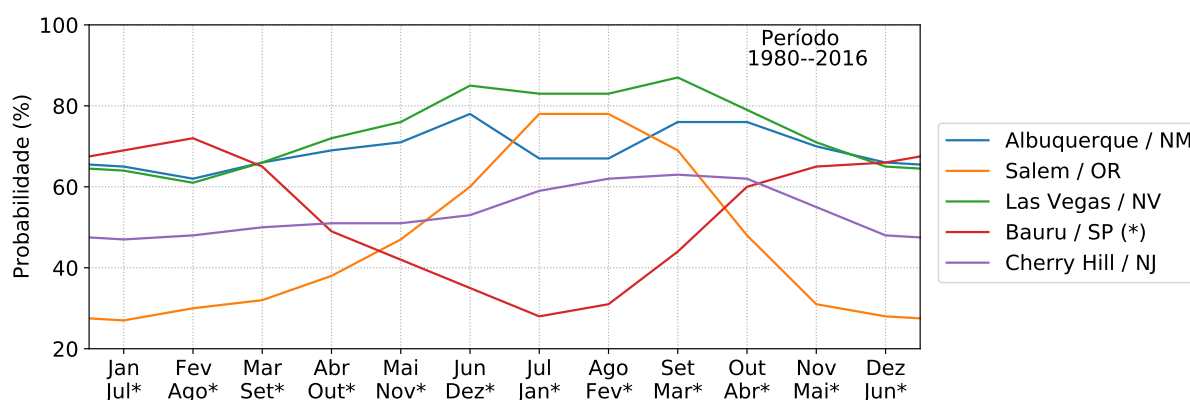
As tentativas de comparar as métricas obtidas com aquelas vistas em outros trabalhos similares, envolvendo o MLP na solução, encontraram alguns obstáculos relacionados à diferença entre os sistemas FVs, à indisponibilidade de dados de entrada – atributos exóticos –, à especificação incompleta dos métodos sugeridos, ao entendimento de que uma solução é pouco elegante⁷ (complexa ou híbrida demais) ou pela limitação dos recursos deste autor.

Em geral, não é viável comparar métodos diferentes a partir de resultados de usinas diferentes, e principalmente, de diferentes locais de instalação ou intervalos de tempo. A Figura 60 ilustra a nebulosidade típica entre os locais das usinas utilizadas neste trabalho: Bauru/SP e Cherry Hill/NJ; e em alguns outros trabalhos:

⁷Uma das características mais notórias das RNAs é sua capacidade de encontrar padrões desconhecidos nos dados, independentemente do conhecimento dos fenômenos naturais envolvidos no problema. Se o método é muito especializado, envolve um número excessivo de ferramentas, tratamentos pré e pós treinamento ou de segmentações do domínio do problema, parece pouco interessante para este autor, pois denota uma escolha ruim de configuração de RNA para a solução do problema.

Albuquerque/NM (THEOCHARIDES *et al.*, 2020), Salem/OR (HUANG *et al.*, 2018) e Las Vegas/NV (LIU *et al.*, 2015) – nos EUA. Isso afeta o erro médio⁸, como pode ser visto na seção 4.2. A chance de céu aberto durante o inverno em Salem no Óregon (EUA) é baixa, o mesmo pode ser dito sobre o verão em Bauru. Se a base de dados for inferior a 12 meses, a qualidade dos resultados será influenciada pelo período disponível. Métodos validados com dados das usinas em Albuquerque ou Las Vegas têm maior chance de apresentar bons resultados que com as demais.

Figura 60 – Chance de céu aberto a parcialmente nublado



Fonte: adaptado de Weather Spark (2022).

Nem sempre os atributos de entrada se referem a um conjunto de dados disponível para o propósito deste trabalho. Theocharides *et al.* (2020) sugerem uma metodologia híbrida utilizando o MLP e um pós-processamento dos resultados para a previsão do dia seguinte, porém, a partir, dentre outras variáveis, da irradiação global horizontal histórica e prevista, dados estes nem sempre disponíveis (e seguramente não para este trabalho). A irradiação solar também é proposta por Hanifulkhair *et al.* (2020) em conjunto com a temperatura dos painéis solares (além da temperatura ambiente), ambos indisponíveis, dentre outros atributos. Liu *et al.* (2015) usam o MLP e variáveis meteorológicas similares a algumas das utilizadas neste trabalho para a previsão do dia seguinte. A diferença estrutural com o modelo aqui proposto, além dos hiperparâmetros diferentes (não informados por aqueles autores), está, principalmente, no fato de que as previsões são feitas para todas as horas do dia seguinte de uma vez, além da proposta de uma variável de entrada para um índice de aerossóis presentes na atmosfera.

Huang *et al.* (2018) aplicaram o MLP com a função de ativação Huber, como no presente trabalho, para a previsão direta do dia seguinte, tendo como entrada a geração dos últimos sete dias (hora a hora), e utilizando-se da mesma fonte de dados de origem da usina EUA. Porém os autores excluíram parte dos registros de treinamento (2,7 meses em 5 anos) e, principalmente, de validação (dois em 12 meses). Não é

⁸Ao menos no modelo aqui proposto, mas parece razoável supor a trivialidade deste fato.

possível saber se parte dos maiores ou menores erros foram – inadvertidamente – excluídos, e como já visto, uma pequena parte das horas é responsável por degradação considerável das métricas com a metodologia aqui proposta. A partir de uma simulação realizada com reprodução parcial do método desses autores, uma vez que os ajustes originais não são conhecidos, e seguindo a dinâmica de aumento gradual do histórico de dados, foram obtidos valores do R^2 ao redor de 0,70 após os primeiros doze meses para as usinas de referência deste trabalho. Neste ponto, vale lembrar que os resultados dos testes apresentados para o método proposto devem ser combinados com a qualidade dos prognósticos meteorológicos, hora a hora, durante a operação, em função do esquema de atualização do fornecedor (Figura 28), o que não foi possível verificar, devido à indisponibilidade – para este trabalho – de dados concomitantes de previsão e medição de energia.

Recomendam-se outros dois cuidados a se tomar ao avaliar os resultados dos trabalhos acadêmicos. Primeiro que as métricas de avaliação são afetadas pela abrangência das horas do dia na amostra. Janelas fixas (pré-definidas) costumam incluir horas escuras (no final da madrugada ou início da noite, dependendo da estação do ano), fáceis de prever. Como visto, é mais difícil prever – em termos relativos – a geração FV das primeiras e últimas horas claras do dia, mas os erros são menores – em termos absolutos. Portanto, dependendo da métrica, será beneficiada ou prejudicada pela abertura ou fechamento do escopo de horas de previsão. Segundo que alguns trabalhos testam os resultados finais do modelo a partir de uma subamostra de exemplos retirados aleatoriamente da amostra completa. Nesse caso, a proporção de ocorrências de extrapolação é mais reduzida que no teste aqui conduzido, em que os dados de treinamento são contíguos, independentes e sazonalmente desproporcionais.

4.3.1 Comparação com Floresta Aleatória

Para mensurar os resultados obtidos da previsão fotovoltaica proposta, optou-se pela aplicação de florestas aleatórias (RFs). O Tensorflow (ABADI *et al.*, 2015) possui um pacote chamado `tensorflow_decision_forests` que oferece um conjunto de algoritmos de última geração para treinar, exibir e interpretar modelos de floresta de decisão. O objetivo foi ter uma ordem de grandeza, não houve a pretensão formal de comparar os métodos. A aplicação foi realizada com todos os ajustes padrão (como uma caixa preta), conforme a Tabela 18, e o processo de treinamento seguiu o ciclo proposto pela Figura 39, com previsões horárias, de forma a tornar os resultados finais comparáveis.

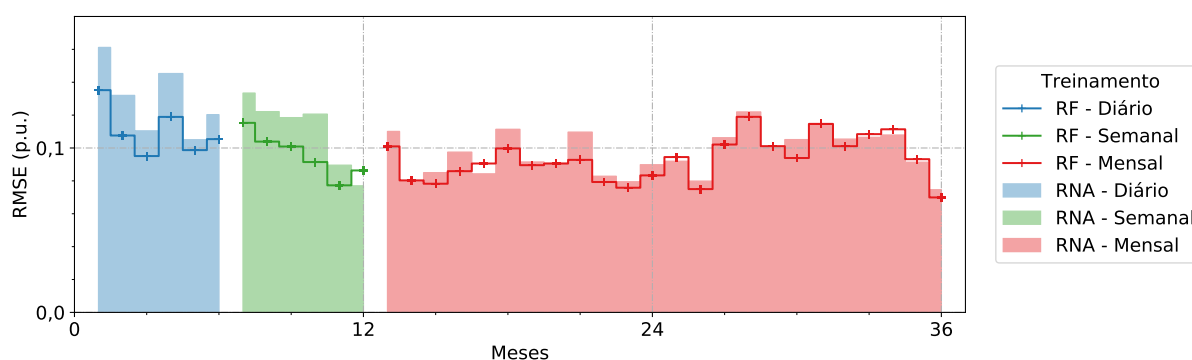
As RFs apresentaram resultados melhores quando os dados eram insuficientes, ou seja, no início do ciclo de treinamento (diário e semanal) quando são demandadas extrapolações, como mostram as Figuras 61 e 62. Além disso, as RFs

Tabela 18 – Configuração padrão da Floresta Aleatória do pacote Tensorflow

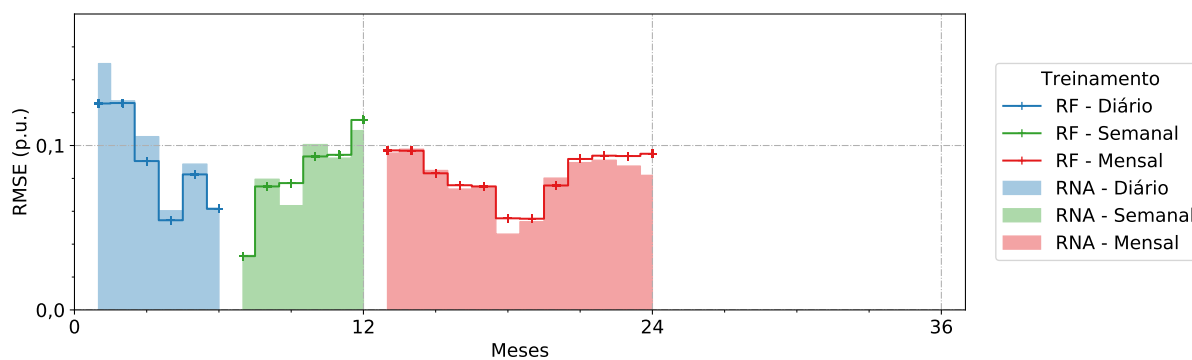
Aspecto	Valor
Modelo	tfdf.keras.RandomForestModel
Tarefa	tfdf.keras.Task.REGRESSION
Função de custo	MSE
Nº de árvores	300
Validação	validação <i>out-of-bag</i> (OOB)

Fonte: elaboração própria (2022).

Figura 61 – Comparativo RNA vs. RF da RMSE



(a) Usina AUS

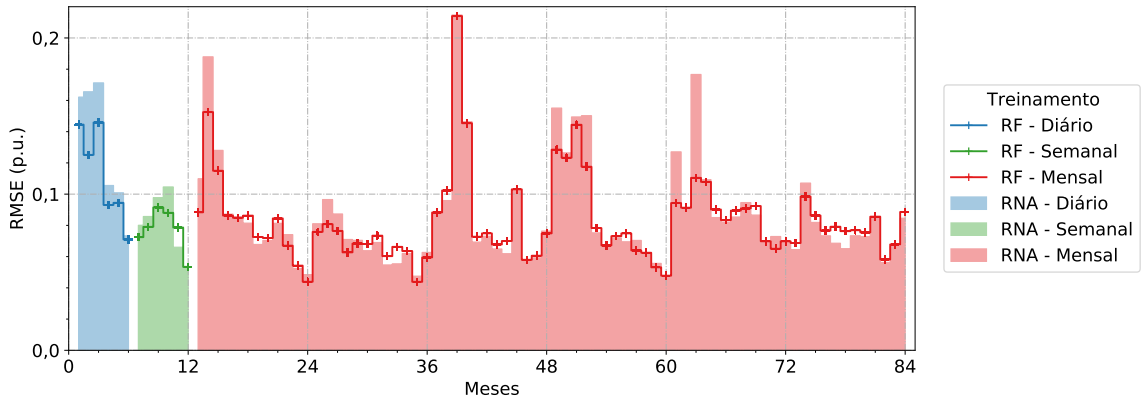


(b) Usina BRA

Fonte: elaboração própria (2022).

conseguiram lidar melhor com as características do inverno rigoroso (janeiro a março) nos EUA onde há ocorrência de neve.

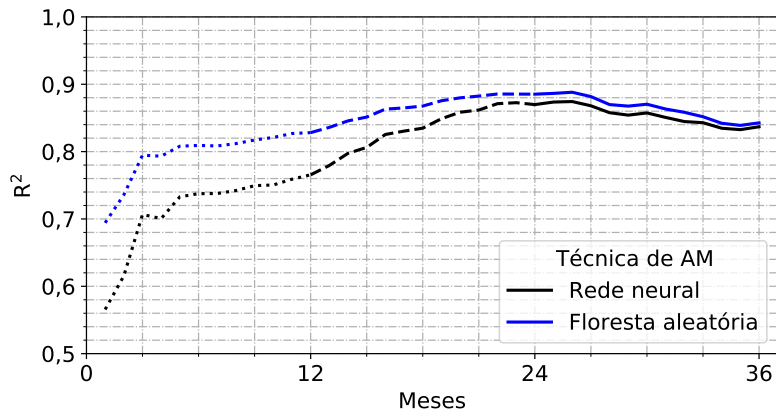
Figura 62 – Comparativo RNA vs. RF da RMSE da Usina EUA



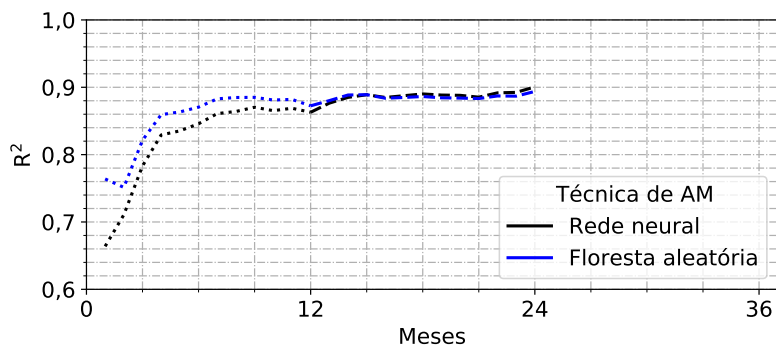
Fonte: elaboração própria (2022).

Dois aspectos interessantes podem ainda ser mencionados. As RNAs parecem se aproximar dos resultados apresentados pelas RFs conforme aumenta a disponibilidade de dados de entrada, como mostram as médias móveis de 12 meses das Figuras 63 e 64. Inicialmente, nos primeiros 12 meses (curvas pontilhadas), a

Figura 63 – Comparativo das curvas R^2 (MM12) entre RNA e RF



(a) Usina AUS

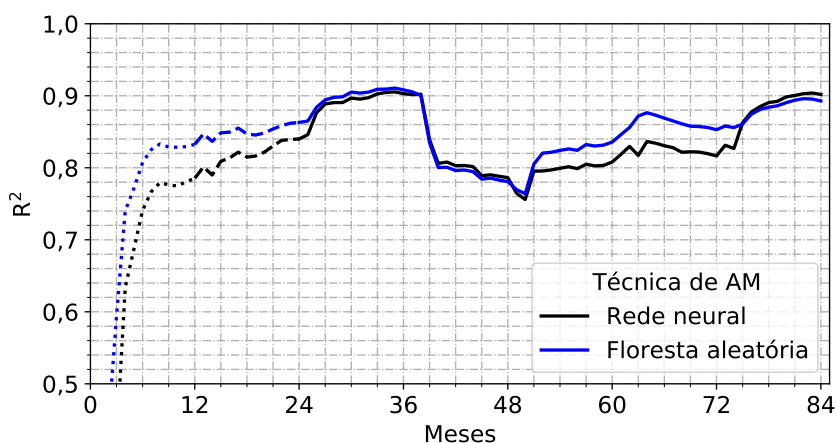


(b) Usina BRA

Fonte: elaboração própria (2022).

desvantagem das redes neurais é evidente. A partir do 2º ano, curvas tracejadas, começa a substituição desses valores piores, e já se pode observar uma melhora. No final do 3º ano das usinas AUS e EUA os resultados já são similares.

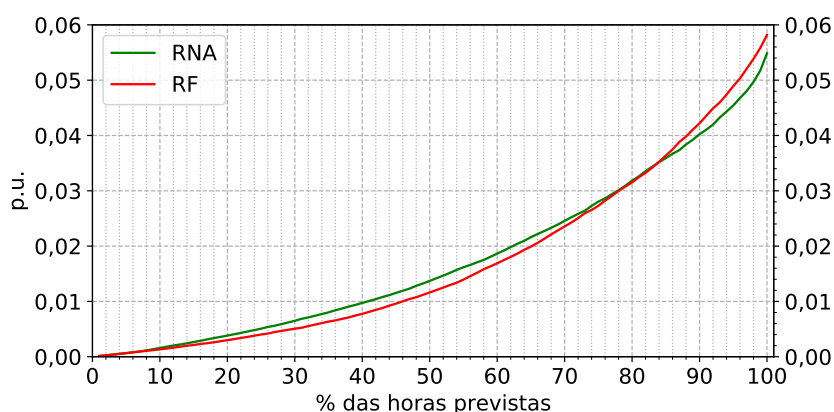
Figura 64 – Comparativo das curvas R^2 (MM12) entre RNA e RF – Usina EUA



Fonte: elaboração própria (2022).

O outro aspecto é que no período de estabilidade das previsões (a partir do 2º ano), as RNAs parecem se sair melhor com os casos mais complexos, se for aceitável associar os erros menores às horas de menor complexidade de representação pelos métodos. A Figura 65 compara o MAE da usina BRA por ambos métodos. Pode-se perceber a vantagem das RFs até 80% dos dados, quando as RNAs começam a performar melhor. Com pequenas variações, a ideia é similar entre as usinas.

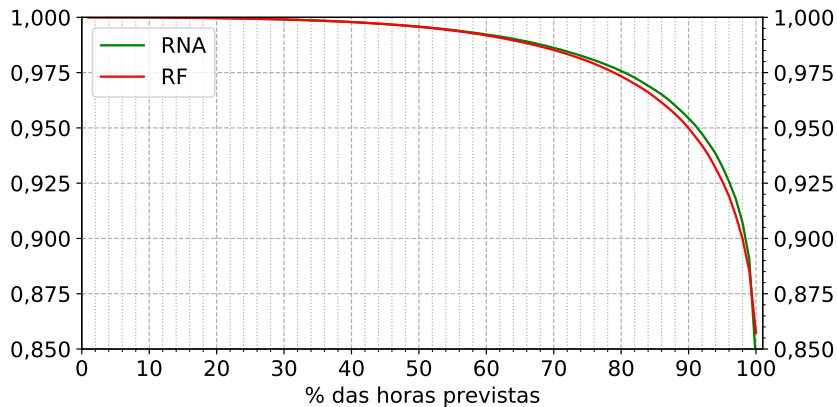
Figura 65 – Comparativo do MAE acumulado entre RNA e RF – Usina BRA



Fonte: elaboração própria (2022).

E comparando-se pela métrica do R^2 , as RNAs mantêm uma curva de melhor desempenho até ao menos 98% dos casos, para todas usinas. A Figura 66 ilustra esse aspecto com o exemplo da usina EUA, aplicável às demais.

A Tabela 19 mostra alguns nichos em que as RNAs conseguiram melhores valores do MAE que as RFs. Vale lembrar que a coordenada geográfica da usina AUS

Figura 66 – Comparativo do R² acumulado entre RNA e RF – Usina EUA

Fonte: elaboração própria (2022).

é estimada, visto que não se conhece o endereço exato mas apenas a região, o que pode ter comprometido os dados meteorológicos obtidos, e que é a única das usinas de referência com margem de melhora por regularização (descartada). Com exceção do caso da usina EUA que possui muitos *outliers* no inverno, as demais possuem baixa nebulosidade e baixo índice pluviométrico nessa estação do ano (é mais fácil de prever).

Tabela 19 – Estações c/ melhor resultado para a RNA por usina – a partir do 2º ano

Estação	AUS	BRA	EUA
Primavera		X	X
Verão		X	X
Outono		X	X
Inverno			

Fonte: elaboração própria (2022).

Pode-se concluir que o MLP é mais sensível ao volume de dados de treinamento e aos pontos fora da curva, mas performa melhor nos casos complexos (de interpolação). Embora ambos métodos sejam inapropriados para extrapolações, o MLP parece sofrer mais, provavelmente em função da aritmética envolvida com os valores absolutos de entrada durante o processo de previsão, ao contrário da RF que utiliza apenas classificações durante a previsão, ou seja, qualquer valor de saída do modelo da RF, será algum possível, visto durante o treinamento.

4.4 Conclusão parcial

Da aplicação do método proposto às três usinas de referência desse trabalho (AUS, BRA e EUA), em uma lógica de inclusão de usinas sem histórico, o que deverá

ser comum no caso da agregação de pequenas usinas às VPPs, e com dados nunca vistos, segregados do desenvolvimento, percebeu-se que após cerca de nove meses uma frequência mensal de treinamento da rede neural é suficiente.

A usina com menor cobertura de nuvens média ao longo do ano – usina BRA – apresentou métricas melhores, o que era intuitivamente esperado, e boa convergência com o resultado de validação cruzada observado durante o desenvolvimento.

A análise de alguns erros maiores observados na amostra da usina EUA mostrou que existem *outliers* entre os dados de treinamento e de teste, mas ao observar os resultados acumulados pode-se perceber que uma parcela significativa das horas apresentou previsões melhores, e que as métricas são muito degradadas por um número pequeno desses *outliers*.

Na forma atual, o MLP não apresentou vantagem sobre o método de referência de RF, principalmente em função da complexidade maior e nem tanto pela pequena defasagem das métricas globais. Porém, o avanço para inclusão de algum tipo de memória de curto prazo, para lidar com eventos relacionados a defeitos técnicos, indisponibilidade e sombreamento, pode colaborar para uma maior capacidade de representação da RNA.

A capacidade intrínseca de redução da variabilidade das previsões por agregação foi evidenciada, mesmo com apenas duas usinas significativas (a usina AUS é relativamente pequena), e sugere o potencial de previsões assertivas para a carteira da VPP, com o crescimento do número de usinas distribuídas geograficamente.

5 CONCLUSÃO

Este trabalho investigou a abordagem de previsões de geração solar fotovoltaica baseada em prognósticos meteorológicos disponíveis a qualquer interessado, adquiridos via API – sem a necessidade de investimentos em estações meteorológicas e integrações mais sofisticadas –, voltadas para o pequeno porte e considerando o contexto das VPPs difundidas.

Uma forma de cálculo do geoposicionamento solar foi investigada e incorporada ao trabalho. E uma fonte de dados meteorológicos é sugerida. Ela recebe informações de sensores meteorológicos de estações ao redor do mundo, além de imagens de satélites, e as processa com o uso de algoritmos próprios de IA para fornecer previsões das condições do tempo em qualquer coordenada da Terra para as próximas 48 horas, como: temperatura ambiente, umidade relativa do ar, pressão atmosférica, cobertura de nuvens, velocidade e direção dos ventos, índice de chuvas e neve; de forma regular para uma série de usos que vão da agricultura e geração de energia ao combate aos desastres naturais.

Desconsiderando-se outras técnicas qualificadas para resolução do problema, a adoção das RNAs de AP pelo autor procura resolver a previsão de geração solar fotovoltaica ao mesmo tempo em que se constrói conhecimento sobre essa técnica – a partir do zero – e experiência para a multiplicação dos usos dessa ferramenta de IA.

O MLP, a princípio, pareceu uma estrutura simples e de fácil aplicação, sobretudo porque as bibliotecas de código aberto disponíveis para a linguagem Python, especialmente a ferramenta testada – o Tensorflow –, facilitam muito a obtenção de algum resultado imediato. Claramente, deve haver um número de situações em que as configurações padrão dessas ferramentas trazem bons resultados (aceitáveis) sem a necessidade de qualquer ajuste.

Mas com o objetivo de atingir um método de previsão de geração solar FV tão universal quanto possível, adaptável a novos participantes da VPP, os dados de três usinas de diferentes regiões (Brasil, EUA e Austrália) foram obtidos e buscou-se verificar, passo a passo na construção do modelo, para cada hiperparâmetro, a possibilidade de definir uma configuração padrão. Foram sete experimentos principais e duas dezenas de sondagens não relatadas – quando as hipóteses não se confirmaram – para manutenção de uma extensão aceitável deste trabalho.

Talvez as palavras que melhor exprimam os cuidados a serem tomados durante a aplicação das redes neurais sejam duas: generalização e validação. Assim

como outros algoritmos clássicos de AM, as RNAs tendem ao sobreajuste, ou seja, a reproduzir perfeitamente as saídas a partir de entradas conhecidas, porém, nesse caso, perdem a capacidade de prever as saídas a partir de dados de entrada nunca vistos. Para lidar com isso empregou-se a validação cruzada, não sem um elevado custo computacional. E a maior parte dos dados obtidos foram segregados e reservados para os testes de validação final do estudo de caso.

Uma das dúvidas iniciais, o porte da rede neural, foi esclarecida com suporte da literatura e a confirmação de um dos experimentos: não vale a pena aumentar seu porte de forma indiscriminada, deve-se preferir o crescimento concomitante do número de neurônios e camadas ocultas, com preferência para este último aspecto, pois são as funções de ativação de camadas encadeadas que permitem às RNAs transformar as entradas – de um espaço linearmente não-separável – em saídas de um espaço linearmente separável. E o aumento do número de variáveis (parâmetros da rede neural) – durante o treinamento – não causa prejuízo à variância dos resultados, algo comum aos métodos lineares de AM.

É preciso registrar que muitas das decisões ao longo do processo de experimentação e seleção de configurações dos hiperparâmetros da rede demandaram um cuidado especial com a variabilidade natural. E um dos aprendizados que ajudaram a conter as expectativas e a manter a motivação veio da bibliografia: não existe metodologia capaz de prever aleatoriedades. A incerteza relacionada com a seleção da parcela de validação, ou com o processo de *back propagation* do algoritmo otimizador – SGD – pode ser tratada por validação cruzada ou pela repetição de execuções. Mas alguns dados históricos das condições de tempo, obtidos do fornecedor, e resultantes de cálculos e interpolações das medidas obtidas dos sensores das estações meteorológicas próximas das usinas, não tiveram como ser tratadas, sequer medidas.

Surpreenderam positivamente dois fatos, porque corroboram a hipótese de um sistema único, capaz de absorver novas adesões à VPPs. São eles:

- a) os ajustes dos hiperparâmetros do MLP respondem de forma muito similar entre as usinas, apesar de elas estarem sujeitas a condições do tempo diferentes, em termos de nebulosidade, temperaturas e chuvas ao longo do ano;
- b) o conjunto formado pela otimização SGDM, a função de ativação ELU e a normalização de entradas Minmax desobriga a tomada de ações de regularização para o tipo de problema. Isso reduz significativamente a complexidade da implementação ao eliminar a necessidade de *gridsearch*.

Os resultados obtidos dos estudos de caso não possuem um paralelo para

comparação direta na bibliografia, em função de algumas diferenças com outros trabalhos, que utilizam outros métodos, dados de medição de estações meteorológicas das usinas, dados exóticos indisponíveis ao autor, outras métricas ou discretizações menores, ou referem-se apenas às primeiras horas do horizonte, por exemplo.

Mas da aplicação própria do algoritmo de RF pode-se verificar que as RNAs:

- a) têm potencial para performar melhor quando os dados históricos de treinamento são bem diversos. Diante de *outliers* causados pela obstrução dos painéis solares (neve), incorreções dos dados meteorológicos ou ainda por insuficiência de dados para o treinamento (quando ainda não há histórico do ciclo solar anual), foram mais sensíveis que as RFs;
- b) não devem apresentar vantagem significativa sobre o algoritmo RF na forma atual: MLP;
- c) algumas sondagens mostraram um potencial significativo com a inclusão de alguma capacidade de memória, típica das RNN. Pois se a percepção de sombreamento, por exemplo, for incluída no modelo, os resultados têm um salto.

A agregação dos resultados individuais das usinas, como se estivessem no mesmo fuso-horário e hemisfério, e fossem participantes de uma mesma VPP, apresentou um resultado bem melhor, e com apenas três usinas, tendo uma delas (AUS) volume desprezível frente às demais (BRA e EUA). Isso mostra a capacidade intrínseca da agregação para a redução da variabilidade da previsão, seja através do MLP proposto, de uma RF ou outro método.

Finalmente, acredita-se ter contribuído com uma base de conhecimento para o enfrentamento dos desafios relacionados à previsão de geração solar fotovoltaica de pequeno porte voltada às microrredes e plantas virtuais através de redes neurais. Os testes realizados foram condizentes com a forma como a previsão seria operacionalizada nesses sistemas e geraram vários *insights* e dúvidas para investigação futura. Agora, os esforços podem ser direcionados para a solução dos limites identificados, e os avanços – na mesma linha de pesquisa – passam a ter uma base de comparação mais sólida (*benchmark*).

5.1 Sugestões para trabalhos futuros

Esse desafio não se encerra com esse trabalho, pois é apenas o embrião de um projeto de previsão e monitoramento de geração solar fotovoltaica mais amplo. A ideia é construir um ecossistema de algoritmos inteligentes capaz de lidar com os aspectos da geração solar fotovoltaica de forma integrada, desde o controle e monitoramento até o diagnóstico de falhas, a proposição de ajustes técnicos e a

integração com as cargas e o armazenamento.

Durante o desenvolvimento foi possível perceber como os dados "falam" sobre os sistemas. Por exemplo, uma das primeiras visões foi que seria possível calcular a eficiência do sistema com base na segregação dos dias de céu aberto, ao correlacionar a saída do sistema com dados teóricos de disponibilidade de irradiação solar. Embora não se tenha avançado a uma conclusão, os indícios foram fortes.

Os dados climáticos, de geoposicionamento solar e a saída dos sistemas pode permitir estimar o reposicionamento ótimo dos painéis solares e detectar erros do projeto original. Seria viável um mercado de retrabalho das instalações? Ou revenda para uso em condições mais favoráveis?

A geração média da usina norte-americana em céu aberto – cujo banco de dados era grande o suficiente para essa observação – apresenta uma degradação anual. O que se pode fazer com essa informação? Como comprová-la e desambiguá-la das variações climáticas? Seria possível comparar a performance de tecnologias ou fabricantes?

Outro valor a ser extraído dos dados é a detecção de sombras e outras obstruções como a neve. Na usina norte-americana, aparentemente, porque não é possível afirmar, os painéis não são limpos após a ocorrência de neve. Mesmo sendo o telhado plano e transitável entre os painéis – aparentemente, por análise de imagem de satélite. Evidentemente há que se considerar as questões econômico-operacionais e de segurança, mas uma VPP competitiva pode ter vantagens em acionar algum tipo de serviço nesse sentido.

Ao menos os sistemas FV das usinas EUA e BRA, mas provavelmente também o da usina AUS, não possuem esquema de desligamento noturno dos inversores, o que reflete em consumo de energia. Estima-se uma perda de 1 MWh/ano (EUA) e 1,4 MWh/ano (BRA). Pode valer a pena estudar esse efeito em massa, para a proposição de normas regulatórias.

Deseja-se incluir dados específicos das usinas, como latitude, altitude, número e característica dos strings com posicionamento diferentes para criar um algoritmo capaz de compartilhar as informações presentes nos dados, entre as usinas. Um exemplo limite disso é o caso de duas usinas em edifícios vizinhos. Elas são sujeitas a condições parecidas, logo, a mais antiga pode fornecer informações importantes para o prognóstico de geração da mais jovem delas.

As plantas virtuais de energia e as microrredes possuem outros tipos de REDs, como as cargas e os geradores eólicos, por exemplo, que também demandam prognósticos futuros. Quão similar são esses contextos? É possível comunizar parcialmente as metodologias adotadas? Além disso, seria útil e inovador mensurar a

sensibilidade dos otimizadores das plantas virtuais e microrredes aos prognósticos de carga e geração, para ponderar o esforço viável em seu aperfeiçoamento. No caso das VPPs, estudar o impacto da agregação de usinas para a operação. Uma das sugestões na literatura para lidar com a variabilidade dos sistemas FV é a difusão geográfica. Assim, uma boa qualidade de previsão consolidada pode ser obtida a partir de previsões individuais mais incertas. O mesmo vale para cargas e outros REDs.

Vale a pena estudar a qualidade das previsões dos fornecedores de dados meteorológicos. Mas também se os dados históricos informados por esses serviços são assertivos. Para isto seria necessário obter dados da estação meteorológica de uma ou mais usinas e compará-los com as informações desses serviços.

Para a previsão de geração fotovoltaica proposta, sugere-se compará-la com uma rede neural recorrente, para tratar nuances como: o sombreamento extemporâneo e suas consequências de curto prazo; defeitos técnicos – parada de funcionamento – da eletrônica de potência; e degradação do sistema ao longo do tempo. Além disso, estudar a hipótese de discretização das entradas e saídas da rede neural, tornando-a um problema de classificação, em substituição à regressão. Seria essa técnica suficiente para alcançar e ultrapassar os resultados das RFs, que se mostraram menos sensíveis aos *outliers*? Outro aspecto técnico relacionado com as redes neurais desse estudo seria a avaliação de esquemas de treinamento. Como o volume de dados de treinamento pode ser otimizado? Vale a pena treinar com mais frequência e dados reduzidos, mais recentes?

As redes neurais artificiais do tipo *feedforward* com gradiente descendente e *back propagation* são deficientes quando uma extrapolação é demandada. Seria interessante estudar as condições que eventualmente podem tornar a RNA capaz de desempenhar melhor para a previsão de geração solar fotovoltaica.

E finalmente, no que diz respeito à implementação computacional do treinamento, sugere-se a investigação de meios de paralelização e uso eficiente da CPU/GPU e dos softwares disponíveis, principalmente com vistas ao crescimento do número de unidades da VPP a serem previstas; e a análise de viabilidade da aplicação de serviços comerciais de computação em nuvem (*serverless*).

REFERÊNCIAS

ABADI, M. *et al.* **TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems**. 2015. Disponível em: <www.tensorflow.org>. Acesso em: 16 jun 2022. Citado 4 vezes nas páginas 83, 89, 122 e 222.

AES BRASIL. **Audiência Pública do Programa de Pesquisa e Desenvolvimento da ANEEL**. 2021. Disponível em: <<https://www.aesbrasil.com.br/sites/default/files/2022-03/Termo%20de%20audiencia%20publica%202021.pdf>>. Acesso em: 10 jun 2022. Citado na página 32.

AGGARWAL, C. C. **Neural Networks and Deep Learning: a Textbook**. 1. ed. Cham: Springer International Publishing, 2018. Citado 8 vezes nas páginas 54, 56, 57, 58, 96, 151, 156 e 159.

AKHTER, M. N. *et al.* Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. **IET Renewable Power Generation**, Institution of Engineering and Technology, v. 13, n. 7, p. 1009–1023, 2019. Citado 5 vezes nas páginas 23, 24, 37, 47 e 48.

ALKHAYAT, G.; MEHMOOD, R. A review and taxonomy of wind and solar energy forecasting methods based on deep learning. **Energy and AI**, Elsevier B.V., v. 4, jun 2021. Citado 2 vezes nas páginas 48 e 50.

ANEEL. **Consulta Pública nº 040/2021**. Proposta de modelo para Contrato de Geração Distribuída oriunda de Chamada Pública, nos termos do Decreto nº 5.163, de 2004, e proposta de critérios para contratar energia com o objetivo de reduzir despesas de operação e manutenção ou postergar investimentos. 2021. Citado 2 vezes nas páginas 31 e 32.

ANEEL. **Consulta pública nº 080/2021**. Obter subsídios para o aprimoramento do Programa de Resposta da Demanda, de que trata a Resolução Normativa nº 792/2017. 2021. Disponível em: <<https://antigo.aneel.gov.br/web/guest/consultas-publicas>>. Acesso em: 16 jul 2022. Citado na página 31.

ANEEL. **Matriz Elétrica Brasileira**. 2022. Disponível em: <<https://dadosabertos.aneel.gov.br/dataset/siga-sistema-de-informacoes-de-geracao-da-aneel>>. Acesso em: 1 jun 2022. Citado na página 24.

APPELS, R. *et al.* The effect of dust deposition on photovoltaic modules. In: **Conference Record of the IEEE Photovoltaic Specialists Conference**. [S.l.: s.n.], 2012. p. 1886–1889. Citado na página 63.

ATWOOD, J.; SPOLSKY, J. **StackExchange**. 2022. Disponível em: <<https://stackoverflow.com/>>. Acesso em: 16 jun 2022. Citado 2 vezes nas páginas 89 e 151.

AUSGRID. **Solar home electricity data - notes**. [S.l.], 2014. Disponível em: <<http://www.ausgrid.com.au/Common/About>>. Citado 3 vezes nas páginas 30, 64 e 67.

BECK, H. E. *et al.* Present and future köppen-geiger climate classification maps at 1-km resolution. **Scientific Data**, Nature Publishing Groups, v. 5, 2018. Citado na página 65.

BLAGA, R. *et al.* A current perspective on the accuracy of incoming solar energy forecasting. **Progress in Energy and Combustion Science**, Elsevier Ltd, v. 70, p. 119–144, jan 2019. Citado na página 51.

BLUMTHALER, M.; AMBACH, W.; ELLINGER, R. Increase in solar UV radiation with altitude. **Journal of Photochemistry and Photobiology B: Biology**, v. 39, p. 130–134, 1997. Citado na página 42.

BRASIL. **Lei nº 10.848, de 15 de março de 2004**. Dispõe sobre a comercialização de energia elétrica, altera as Leis nºs 5.655, de 20 de maio de 1971, 8.631, de 4 de março de 1993, 9.074, de 7 de julho de 1995, 9.427, de 26 de dezembro de 1996, 9.478, de 6 de agosto de 1997, 9.648, de 27 de maio de 1998, 9.991, de 24 de julho de 2000, 10.438, de 26 de abril de 2002, e dá outras providências. Brasília, DF: Presidência da República, 2004. Citado na página 31.

BRASIL. **Decreto nº 5.163, de 30 de julho de 2004**. Regulamenta a comercialização de energia elétrica, o processo de outorga de concessões e de autorizações de geração de energia elétrica, e dá outras providências. Brasília, DF: Presidência da República, 2004. Citado na página 31.

BRASIL. **Smart Grid**. Brasília: Ministério de Minas e Energia, 2010. Citado na página 27.

BRASIL. **Instruções para Solicitação de Cadastramento e Habilitação Técnica com vistas à participação nos Leilões de Energia Elétrica**. Brasília: Empresa de Pesquisa Energética, 2014. Citado na página 27.

BRASIL. **Lei nº 14.120 de 1º de março de 2021**. Altera a Lei nº 9.991, de 24 de julho de 2000, a Lei nº 5.655, de 20 de maio de 1971, a Lei nº 9.427, de 26 de dezembro de 1996, a Lei nº 10.438, de 26 de abril de 2002, a Lei nº 10.848, de 15 de março de 2004, a Lei nº 12.111, de 9 de dezembro de 2009, a Lei nº 12.783, de 11 de janeiro de 2013, a Lei nº 13.203, de 8 de dezembro de 2015, e o Decreto-Lei nº 1.383, de 26 de dezembro de 1974; transfere para a União as ações de titularidade da Comissão Nacional de Energia Nuclear (CNEN) representativas do capital social da Indústrias Nucleares do Brasil S.A. (INB) e da Nuclebrás Equipamentos Pesados S.A. (Nuclep); e dá outras providências. Brasília: Presidência da República, 2021. Citado na página 31.

BRASIL. **Projeto de Lei nº 1.539 de 2021**. Altera a Política Nacional sobre Mudança do Clima (Lei 12.187, de 29 de dezembro de 2009) para estabelecer nova meta de compromisso nacional voluntário e seu depósito junto à Convenção-Quadro das Nações Unidas sobre Mudanças Climáticas. Brasília: Câmara dos Deputados, 2021. Citado na página 27.

BRASIL. **Projeto de Lei nº 414 de 2021**. Altera as Leis nº 9.074, de 7 de julho de 1995, nº 9.427, de 26 de dezembro de 1996, nº 9.991, de 24 de julho de 2000, nº 10.438, de 26 de abril de 2002, nº 10.848, de 15 de março de 2004, nº 12.783, de 11 de janeiro de 2013, nº 13.203, de 8 de dezembro de 2015, nº 5.655, de 20 de maio de 1971, e nº 12.111, de 9 de dezembro de 2009, para aprimorar o modelo regulatório e comercial do setor elétrico com vistas à expansão do mercado livre, e dá

outras providências. Brasília: Câmara dos Deputados, 2021. Disponível em: <<https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2270036>>. Acesso em: 10 jun 2022. Citado 3 vezes nas páginas 25, 31 e 34.

BRASIL. **Lei nº 14.300 de 6 de janeiro de 2022**. Institui o marco legal da microgeração e minigeração distribuída, o Sistema de Compensação de Energia Elétrica (SCEE) e o Programa de Energia Renovável Social (PERS); altera as Leis nºs 10.848, de 15 de março de 2004, e 9.427, de 26 de dezembro de 1996; e dá outras providências. Brasília: Presidência da República, 2022. Citado na página 64.

BRASIL. **Lei nº 1.224 de 2022**. Dispõe acerca da atividade de armazenamento de energia no âmbito do Sistema Interligado Nacional (SIN). Brasília: Câmara dos Deputados, 2022. Citado 2 vezes nas páginas 31 e 32.

CASARIN, R. **WEG fornecerá sistema de armazenamento de energia para FAB**. 2021. Citado na página 27.

CCEE. Procedimentos de Comercialização. **Submódulo 1.6 – Comercialização Varejista – Versão 4**. 2022. Disponível em: https://www.ccee.org.br/documents/80415/919498/1.6%20-%20Comercializa%C3%A7%C3%A3o%20varejista_v3.0.pdf/48e3d144-de9d-c543-ccce-15febf48c5ae. Acesso em: 16 jul 2022. Citado na página 31.

CCEE. **Conceitos de Preços**. 2022. Disponível em: <https://www.ccee.org.br/web/guest/precos/conceitos-precos>. Acesso em: 7 jun 2022. Citado 2 vezes nas páginas 26 e 33.

CCEE. **Plataforma de Integração**. 2022. Disponível em: <https://www.ccee.org.br/documentos/plataforma-de-integracao>. Acesso em: 6 mai 2022. Citado 2 vezes nas páginas 61 e 76.

CHOI, D. *et al.* On Empirical Comparisons of Optimizers for Deep Learning. **CoRR**, v. abs/1910.05446, out. 2019. Citado na página 86.

CHOLLET, F. **Keras**. 2015. Disponível em: <https://github.com/fchollet/keras>. Acesso em: 17 nov 2021. Citado 2 vezes nas páginas 83 e 87.

CLEVERT, D. A.; UNTERTHINER, T.; HOCHREITER, S. Fast and accurate deep network learning by exponential linear units (ELUs). **4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings**, p. 1–14, 2016. Citado 2 vezes nas páginas 58 e 157.

DANTAS, S. G.; POMPERMAYER, F. M. **Viabilidade econômica de sistemas fotovoltaicos no Brasil e possíveis efeitos no Setor Elétrico**. [S.l.], 2018. Citado 3 vezes nas páginas 24, 26 e 27.

DE VEAUX, R. D.; UNGAR, L. H. Multicollinearity: A tale of two nonparametric regressions. **Selecting Models from Data**, v. 89, 1994. Citado na página 75.

DEKKing, F. M. *et al.* **A Modern Introduction to Probability and Statistics**. London: Springer London, 2005. 10–27 p. (Springer Texts in Statistics). Citado 6 vezes nas páginas 93, 196, 204, 216, 218 e 222.

DING, M.; WANG, L.; BI, R. An ANN-based approach for forecasting the power output of photovoltaic system. **Procedia Environmental Sciences**, Elsevier B.V., v. 11, p. 1308–1315, 2011. Não citado.

DING, B. *et al.* The dependence of precipitation types on surface elevation and meteorological conditions and its parameterization. **Journal of Hydrology**, Elsevier B.V., v. 513, p. 154–163, maio 2014. Citado na página 68.

DOZAT, T. **Incorporating Nesterov Momentum into Adam**. [S.l.], 2016. Citado na página 87.

DUCHAUD, J. L. *et al.* Trade-off between precision and resolution of a solar power forecasting algorithm for micro-grid optimal control. **Energies**, MDPI AG, v. 13, n. 14, jul 2020. Citado na página 61.

DUCHI, J.; SINGER, Y. **Adaptive Subgradient Methods for Online Learning and Stochastic Optimization** * Elad Hazan. [S.l.], 2011. v. 12, 2121–2159 p. Citado 2 vezes nas páginas 87 e 154.

EMARATI, M.; KEYNIA, F; RASHIDINEJAD, M. **A two-stage stochastic programming framework for risk-based day-ahead operation of a virtual power plant**. *International Transactions on Electrical Energy Systems*, 2020. v. 30, ed. 3. Citado na página 64.

GANDHI, O. *et al.* Review of power system impacts at high PV penetration Part I: Factors limiting PV penetration. **Solar Energy**, Elsevier Ltd, v. 210, p. 181–201, nov 2020. Citado na página 24.

GENDZH, N. **Comparing the weather forecasting quality and accuracy of Europe and America**. [S.l.], 2014. Citado na página 35.

GÉRON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. [S.l.], 2019. Citado 9 vezes nas páginas 56, 57, 58, 90, 150, 152, 153, 159 e 162.

GIANCRISTOFARO, R. A.; SALMASO, L. **Model performance analysis and model validation in logistic regression**. [S.l.], 2003. v. 2. Citado na página 52.

GLOROT, X.; BENGIO, Y. **Understanding the difficulty of training deep feedforward neural networks**. [S.l.], 2010. Citado na página 190.

GOLUBEVA, A.; NEYSHABUR, B.; GUR-ARI, G. Are wider nets better given the same number of parameters? v., 2020. Citado 2 vezes nas páginas 95 e 151.

GOOGLE MAPS. **5 Executive Campus - Cherry Hill**. 2021. Disponível em: <<https://www.google.com/maps/place/5+Executive+Campus,+Cherry+Hill,+NJ+08002,+USA>>. Acesso em: 16 jun 2022. Citado na página 178.

GOOGLE MAPS. **AES Tietê Energia**. 2021. Disponível em: <<https://www.google.com/maps/place/AES+Tiet%C3%AA+Energia>>. Acesso em: 16 jun 2022. Citado na página 176.

GOOGLE MAPS. **The Entrance North/NSW**. 2021. Disponível em: <<https://www.google.com/maps/place/The+Entrance+North+NSW+2261,+Australia>>. Acesso em: 16 jun 2022. Citado na página 173.

HAMILTON, J.; NEGNEVITSKY, M.; WANG, X. The potential of variable speed diesel application in increasing renewable energy source penetration. In: **Energy Procedia**. Melbourne: Elsevier Ltd, 2019. v. 160, p. 558–565. Citado na página 34.

HANIFULKHAIR *et al.* One Day Ahead Prediction of PV Power Plant for Energy Management System Using Neural Network. In: **International Seminar on Intelligent Technology and Its Applications**. IEEE, 2020. p. 107–112. Citado na página 121.

HE, K. *et al.* Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. fev 2015. Citado na página 190.

HINTON, G.; SRIVASTAVA, N.; SWERSKY, K. **Neural Networks for Machine Learning Lecture 6a Overview of mini--batch gradient descent**. [S.l.], 2012. Citado 2 vezes nas páginas 87 e 154.

HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. **Neural Computation**, MIT Press Journals, v. 9, n. 8, p. 1735–1780, nov 1997. Citado na página 60.

HOFFER, E.; HUBARA, I.; SOUDRY, D. Train longer, generalize better: Closing the generalization gap in large batch training of neural networks. **Advances in Neural Information Processing Systems**, v. 2017-Decem, p. 1732–1742, 2017. Citado na página 89.

HONSBURG, C. B.; BOWDEN, S. G. **Photovoltaics Education Website**. 2022. Disponível em: <<http://pveducation.org>>. Acesso em: 16 jun 2022. Citado na página 44.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural Networks**, v. 2, n. 5, p. 359–366, 1989. Citado 2 vezes nas páginas 81 e 151.

HUANG, C. *et al.* Day-ahead forecasting of hourly photovoltaic power based on robust multilayer perception. **Sustainability**, v. 10, n. 12, p. 4863, MDPI, 2018. Citado na página 121.

HURWITZ, J.; KIRSCH, D. **Machine Learning for dummies**. 1. ed. Hoboken: John Wiley and Sons, 2018. Citado na página 54.

JAMES, G. *et al.* **An Introduction to Statistical Learning**. New York, NY: Springer New York, 2013. v. 103. (Springer Texts in Statistics, v. 103). Citado 3 vezes nas páginas 51, 53 e 190.

JEBLI, I.; BELOUADHA, F. Z.; KABBAJ, M. I. The forecasting of solar energy based on Machine Learning. In: **2020 9th International Conference on Electrical and Information Technologies, ICEIT 2020**. Oxford: Institute of Electrical and Electronics Engineers Inc., 2020. Citado 4 vezes nas páginas 24, 46, 55 e 56.

KALE, S.; SEKHARI, A.; SRIDHARAN, K. SGD: The Role of Implicit Regularization, Batch-size and Multiple-epochs. **arXiv**, jul 2021. Citado na página 88.

KESKAR, N. S. *et al.* On large-batch training for deep learning: Generalization gap and sharp minima. **5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings**, p. 1–16, 2017. Citado 4 vezes nas páginas 89, 152, 153 e 154.

KIM, G. G. *et al.* Prediction Model for PV Performance with Correlation Analysis of Environmental Variables. **IEEE Journal of Photovoltaics**, IEEE Electron Devices Society, v. 9, n. 3, p. 832–841, maio 2019. Citado na página 47.

KINGMA, D. P.; BA, J. Adam: A Method for Stochastic Optimization. **arXiv**, dez 2014. Citado 2 vezes nas páginas 87 e 154.

KIRKERUD, J. G.; NAGEL, N. O.; BOLKESJØ, T. F. The role of demand response in the future renewable northern European energy system. **Energy**, Elsevier Ltd, v. 235, nov 2021. Citado na página 33.

KOHAVI, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In: . [S.l.]: International Joint Conference on Artificial Intelligence (IJCAI), 1995. Citado na página 53.

KOSTYLEV, V.; PAVLOVSKI, A. Solar Power Forecasting Performance – Towards Industry Standards, 2011. Citado na página 50.

KRSTAJIC, D. *et al.* Cross-validation pitfalls when selecting and assessing regression and classification models. **Journal of Cheminformatics**, Gas Turbine Society of Japan, v. 6, n. 1, mar 2014. Citado na página 52.

KUHN, M.; JOHNSON, K. **Applied Predictive Modeling**. New York, NY: Springer New York, 2013. Citado 2 vezes nas páginas 53 e 150.

KUMAR, D. S. *et al.* Solar irradiance resource and forecasting: a comprehensive review. **IET Renewable Power Generation**, John Wiley and Sons Inc, v. 14, n. 10, p. 1641–1656, jul 2020. Citado 4 vezes nas páginas 25, 41, 46 e 48.

LECUN, Y. *et al.* Gradient-Based Learning Applied to Document Recognition. **proc. OF THE IEEE**, 1998. Citado na página 59.

LECUN, Y. A. *et al.* Efficient backprop. 2012. Citado na página 153.

LEE, H. Y.; KO, S. K.; LEE, B. T. Forecast uncertainty-based performance degradation diagnosis of solar PV systems. **IET Renewable Power Generation**, Institution of Engineering and Technology, v. 14, n. 6, p. 1087–1094, abr 2020. Citado 2 vezes nas páginas 24 e 45.

LEITES, T. N. *et al.* Regulatory advances and improvements in Latin America: a Case study of VPP markets in Brazil. In: . Lucerne, CH: European Grid Service Markets, 2018. Citado na página 31.

LI VIGNI, V. *et al.* Proof of concept of an irradiance estimation system for reconfigurable photovoltaic arrays. **Energies**, MDPI AG, v. 8, n. 7, p. 6641–6657, 2015. Citado na página 45.

LIMA, M. A. F. *et al.* Portfolio theory applied to solar and wind resources forecast. **IET Renewable Power Generation**, Institution of Engineering and Technology, v. 11, n. 7, p. 973–978, jun 2017. Citado na página 25.

LIU, J. *et al.* An Improved Photovoltaic Power Forecasting Model With the Assistance of Aerosol Index Data. **IEEE Transactions on Sustainable Energy**, Institute of Electrical and Electronics Engineers Inc., v. 6, n. 2, p. 434–442, abr 2015. Citado 2 vezes nas páginas 47 e 121.

MAKOHIN, D. G. *et al.* Design and Implementation of a Flexible Microgrid Controller through Mixed Integer Linear Programming Optimization. In: **2018 9th IEEE International Symposium on Power Electronics for Distributed Generation Systems, PEDG 2018**. Charlotte: Institute of Electrical and Electronics Engineers Inc., 2018. Citado na página 32.

MASSAOUDI, M. *et al.* Convergence of Photovoltaic Power Forecasting and Deep Learning: State-of-Art Review. **IEEE Access**, Institute of Electrical and Electronics Engineers Inc., v. 9, p. 136593–136615, 2021. Citado 2 vezes nas páginas 59 e 149.

MCMAHAN, H. B. *et al.* Ad click prediction. In: **Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining**. New York, NY, USA: ACM, 2013. p. 1222–1230. Citado na página 87.

MELLIT, A.; KALOGIROU, S. A. **McEvoy's handbook of photovoltaics**. 3. ed. Chennai: Academic Press, 2017. Citado 6 vezes nas páginas 26, 36, 37, 39, 50 e 81.

MICHELI, L. *et al.* Extracting and Generating PV Soiling Profiles for Analysis, Forecasting, and Cleaning Optimization. **IEEE Journal of Photovoltaics**, IEEE Electron Devices Society, v. 10, n. 1, p. 197–205, jan 2020. Citado na página 45.

MOLINARO, A. M.; SIMON, R.; PFEIFFER, R. M. Prediction error estimation: A comparison of resampling methods. **Bioinformatics**, v. 21, n. 15, p. 3301–3307, ago 2005. Citado na página 53.

MONCADA, A.; RICHARDSON, W.; VEGA-AVILA, R. Deep learning to forecast solar irradiance using a six-month UTSA Skylmager dataset. **Energies**, v. 11, n. 8, p. 1–16, 2018. Citado na página 55.

MONTENEGRO, S. **Entidades divulgam manifesto pela aprovação urgente do PL 414**. 2022. Disponível em: <<https://www.canalenergia.com.br/noticias/53211249/entidades-divulgam-manifesto-pela-aprovacao-urgente-do-pl-414>>. Acesso em: 3 jun 2022. Citado na página 25.

MOUSTRIS, K. P. *et al.* One day-ahead forecasting of mean hourly global solar irradiation for energy management systems purposes using artificial neural network modeling. In: **Mediterranean Conference on Power Generation**. Belgrade: IET, 2016. p. 1–6. Citado na página 51.

MURPHY, K. P. **Machine learning : a probabilistic perspective**. 1. ed. Cambridge: MIT Press, 2012. 1067 p. Citado 6 vezes nas páginas 52, 155, 159, 160, 161 e 190.

NAVAL, N.; YUSTA, J. **Virtual power plant models and electricity markets - A review**. Renewable and Sustainable Energy Reviews. v. 149. 2021. Citado 2 vezes nas páginas 31 e 111.

NAIR, V.; HINTON, G. E. Rectified Linear Units Improve Restricted Boltzmann Machines. In: **International Conference on Machine Learning**. Haifa: University of Toronto, 2010. v. 33, n. 5, p. 807–814. Citado na página 157.

NAITZAT, G.; ZHITNIKOV, A.; LIM, L. H. Topology of deep neural networks. **Journal of Machine Learning Research**, v. 21, p. 1–40, 2020. Citado 2 vezes nas páginas 149 e 157.

NESTEROV, Y. **Springer Optimization and Its Applications: Lectures on Convex Optimization Second Edition**. 2. ed. Cham: Springer, 2018. v. 137. Citado na página 150.

NESTEROV, Y. E. A method of solving a convex programming problem with convergence rate $O(1/k^2)$. **Soviet Math. Dokl.**, v. 27, n. 2, p. 372–376, 1983. Citado na página 154.

NREL. **Developer Network**. 2021. Citado 3 vezes nas páginas 64, 165 e 177.

NREL. **Solar Position and Intensity (SOLPOS) Calculator**. 2021. Disponível em: <<https://www.nrel.gov/grid/solar-resource/solpos.html>>. Acesso em: 1 mai 2021. Citado 4 vezes nas páginas 63, 68, 71 e 72.

OLIVEIRA, D. B. S. de *et al.* Analysis of the feasibility of Demand Response on HVAC systems in Brazil: Application for the hotel sector. In: **2019 IEEE PES Conference on Innovative Smart Grid Technologies, ISGT Latin America 2019**. [S.l.]: Institute of Electrical and Electronics Engineers Inc., 2019. Citado na página 33.

OPENWEATHER. **OpenWeather mobile app**. 2020. Disponível em: <openweathermap.org>. Acesso em: 1 dez 2020. Citado 3 vezes nas páginas 67, 69 e 171.

OSPINA, J.; NEWAZ, A.; FARUQUE, M. O. Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model. **IET Renewable Power Generation**, Institution of Engineering and Technology, v. 13, n. 7, p. 1087–1095, 2019. Citado 2 vezes nas páginas 24 e 25.

PEDREGOSA, F. *et al.* Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011. Citado na página 159.

PEREIRA, J. C. **Sistemas Elétricos de Suprimento de Energia para Cargas de Missão Crítica**. [S.l.], 2003. Citado na página 33.

POLYAK, B. Some methods of speeding up the convergence of iteration methods. **USSR Computational Mathematics and Mathematical Physics**, v. 4, n. 5, p. 1–17, jan 1964. Citado 2 vezes nas páginas 87 e 153.

RHINERHART, R. **Regression Modeling for Engineering**. [S.l.: s.n.], 2016. Citado 2 vezes nas páginas 48 e 49.

RIKOS, E. *et al.* Stability and power quality issues in microgrids under weather disturbances. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, Institute of Electrical and Electronics Engineers, v. 1, n. 3, p. 170–179, 2008. Citado na página 23.

ROBBINS, H.; MONRO, S. A stochastic approximation method. **Annals of Mathematical Statistics**, v. 22, p. 400–407, 1951. Citado 2 vezes nas páginas 87 e 152.

RUDER, S. An overview of gradient descent optimization algorithms. set 2016. Citado 2 vezes nas páginas 87 e 154.

RUMELHART, D.; HINTON, G.; WILLIAMS, R. Learning Internal Representations by Error Propagation. In: **Parallel Distributed Processing: Explorations in the Microstructure of Cognition**. [S.l.]: The MIT Press, 1986. Citado 2 vezes nas páginas 60 e 150.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence a modern approach**. [S.l.: s.n.], 2016. Citado 3 vezes nas páginas 57, 155 e 160.

SALOMON, R. Evolutionary algorithms and gradient search: Similarities and differences. **IEEE Transactions on Evolutionary Computation**, v. 2, n. 2, p. 45–55, 1998. Citado na página 150.

SANJARI, M. J.; GOOI, H. B. Probabilistic Forecast of PV Power Generation Based on Higher Order Markov Chain. **IEEE Transactions on Power Systems**, Institute of Electrical and Electronics Engineers Inc., v. 32, n. 4, p. 2942–2952, jul 2017. Citado na página 47.

SANJARI, M. J.; GOOI, H. B.; NAIR, N. K. C. Power generation forecast of hybrid PV-Wind system. **IEEE Transactions on Sustainable Energy**, Institute of Electrical and Electronics Engineers Inc., v. 11, n. 2, p. 703–712, abr 2020. Citado na página 25.

SCHNEIDER, F.; BALLE, L.; HENNIG, P. DeepOBS: A Deep Learning Optimizer Benchmark Suite. mar 2019. Citado na página 86.

SCOLARI, E.; SOSSAN, F.; PAOLONE, M. Photovoltaic-model-based solar irradiance estimators: Performance comparison and application to maximum power forecasting. **IEEE Transactions on Sustainable Energy**, Institute of Electrical and Electronics Engineers Inc., v. 9, n. 1, p. 35–44, jan 2018. Citado na página 47.

SHAKER, H.; ZAREIPOUR, H.; WOOD, D. Estimating Power Generation of Invisible Solar Sites Using Publicly Available Data. **IEEE Transactions on Smart Grid**, Institute of Electrical and Electronics Engineers Inc., v. 7, n. 5, p. 2456–2465, set 2016. Citado na página 25.

SHENG, H. *et al.* Short-Term Solar Power Forecasting Based on Weighted Gaussian Process Regression. **IEEE Transactions on Industrial Electronics**, Institute of Electrical and Electronics Engineers Inc., v. 65, n. 1, p. 300–308, 2018. Citado 2 vezes nas páginas 47 e 67.

SHI, J. *et al.* Forecasting power output of photovoltaic systems based on weather classification and support vector machines. In: **IEEE Transactions on Industry Applications**. [S.l.: s.n.], 2012. v. 48, n. 3, p. 1064–1069. Citado 3 vezes nas páginas 45, 47 e 56.

SHIVASHANKAR, S. *et al.* Mitigating methods of power fluctuation of photovoltaic (PV) sources - A review. **Renewable and Sustainable Energy Reviews**, Elsevier Ltd, v. 59, p. 1170–1184, jun 2016. Citado 2 vezes nas páginas 24 e 25.

SUTSKEVER, I. *et al.* On the importance of initialization and momentum in deep learning. In: **30th International Conference on Machine Learning**. Atlanta: JMLR, 2013. Citado 3 vezes nas páginas 87, 153 e 154.

TAN *et al.* Three-level market optimization model of virtual power plant with carbon capture equipment considering copula-CVaR theory. **Energy**, v. 237, 2021. Citado na página 64.

TAN, H. H.; LIM, K. H. Review of second-order optimization techniques in artificial neural networks backpropagation. **IOP Conference Series: Materials Science and Engineering**, v. 495, n. 1, 2019. Citado na página 150.

TENFEN, D.; FINARDI, E. C. A mixed integer linear programming model for the energy management problem of microgrids. **Electric Power Systems Research**, Elsevier Ltd, v. 122, p. 19–28, 2015. Citado 2 vezes nas páginas 32 e 35.

THEOCHARIDES, S. *et al.* Day-ahead photovoltaic power production forecasting methodology based on machine learning and statistical post-processing. **Applied Energy**, v. 268, 2020. Citado na página 121.

TUOHY, A. *et al.* Solar Forecasting: Methods, Challenges, and Performance. **IEEE Power and Energy Magazine**, Institute of Electrical and Electronics Engineers Inc., v. 13, n. 6, p. 50–59, nov 2015. Citado na página 46.

UPADHYAYULA, S. M.; VENKATARAMANAN, K. Mathematical Methods in Interdisciplinary Sciences. In: CHAKRAVERTY, S. (Ed.). **Mathematical Methods in Interdisciplinary Sciences**. [S.l.]: Wiley, 2020. cap. Mathematic. Citado 3 vezes nas páginas 150, 156 e 157.

VEEN, F. van; LEIJNEN, S. **The Neural Network Zoo**. 2019. Citado 2 vezes nas páginas 30 e 59.

VILLALVA, M. G.; GAZOLI, J. R. **Energia Solar Fotovoltaica**. 2. ed. São Paulo: Érica, 2012. Citado 4 vezes nas páginas 40, 41, 43 e 45.

VIRIDIAN SOLAR. **Theoretical Limits to PV Efficiency**. 2022. Disponível em: <<https://www.viridiansolar.co.uk/resources-4-1-photovoltaic-effect.html>>. Acesso em: 2 jun 2022. Citado na página 42.

WAN, C. *et al.* Photovoltaic and solar power forecasting for smart grid energy management. **CSEE Journal of Power and Energy Systems**, Power System Technology Press, v. 1, n. 4, p. 38–46, jan 2016. Citado na página 46.

WANG, H. *et al.* Taxonomy research of artificial intelligence for deterministic solar power forecasting. **Energy Conversion and Management**, Elsevier Ltd, v. 214, jun 2020. Citado na página 45.

WANG, Y. *et al.* Data-Driven Probabilistic Net Load Forecasting With High Penetration of Behind-the-Meter PV. **IEEE Transactions on Power Systems**, Institute of Electrical and Electronics Engineers Inc., v. 33, n. 3, p. 3255–3264, maio 2018. Citado na página 45.

WEATHER SPARK. Compare the Climate and Weather Between Two+ Cities Worldwide. Disponível em: <<https://weatherspark.com/compare>>. Acesso em: 15 de dez. de 2022. Citado na página 121.

WILSON, A. C. *et al.* The Marginal Value of Adaptive Gradient Methods in Machine Learning. maio 2017. Citado na página 86.

World Meteorological Organization. **Technical Regulations: Volume II – Meteorological Service for International Air Navigation**. World Meteorological Organization, 2018. Disponível em: <https://library.wmo.int/doc_num.php?explnum_id=10733>. Acesso em: 10 jun 2022. Citado na página 65.

XU, K. *et al.* How Neural Networks Extrapolate: From Feedforward to Graph Neural Networks. 2020. Citado 2 vezes nas páginas 79 e 151.

XU, Y.; AHOKANGAS, P.; REUTER, E. **EaaS: Electricity as a Service?** [S.l.], 2018. v. 6, n. 3, 1 p. Citado na página 82.

YALÇIN, O. G. **The Brief History of Convolutional Neural Networks**. 2021. Disponível em: <<https://towardsdatascience.com/the-brief-history-of-convolutional-neural-networks-45afa1046f7f>>. Acesso em: 16 jun 2022. Citado na página 59.

YANG, H. T. *et al.* A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. **IEEE Transactions on Sustainable Energy**, Institute of Electrical and Electronics Engineers Inc., v. 5, n. 3, p. 917–926, 2014. Citado na página 24.

BENGIO, Y.; GRANDVALET, Y. **No Unbiased Estimator of the Variance of K-Fold Cross-Validation** Yoshua Bengio Yves Grandvalet. [S.l.], 2004. v. 5, 1089–1105 p. Citado na página 53.

ZANG, H. *et al.* Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning. **International Journal of Electrical Power and Energy Systems**, Elsevier Ltd, v. 118, jun 2020. Citado na página 151.

ZEILER, M. D. ADADELTA: An Adaptive Learning Rate Method. dez 2012. Citado na página 87.

ZHANG, Y. *et al.* Data-Driven Day-Ahead PV Estimation Using Autoencoder-LSTM and Persistence Model. **IEEE Transactions on Industry Applications**, Institute of Electrical and Electronics Engineers Inc., v. 56, n. 6, p. 7185–7192, nov 2020. Citado na página 24.

ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 67, n. 2, p. 301–320, abr 2005. Citado na página 161.

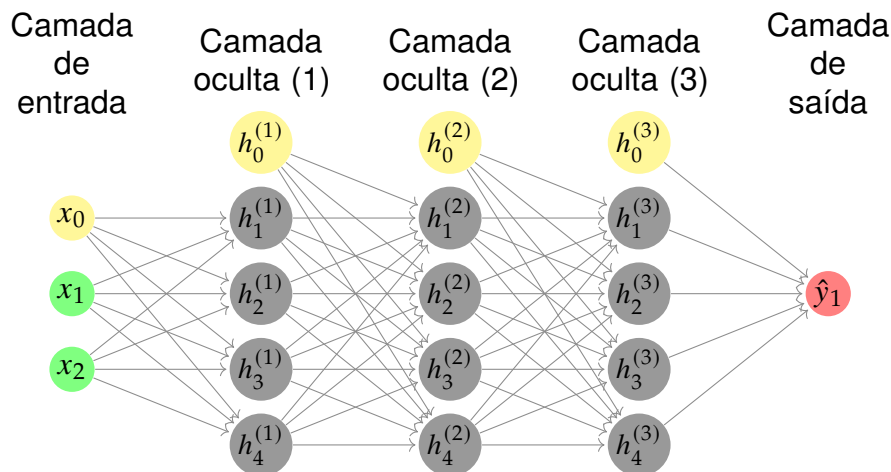
Apêndices

APÊNDICE A – PERCEPTRON MULTICAMADAS

O *perceptron* multicamadas (MLP) é um tipo de RNA de AP por excelência. Ao aumentar a comunicabilidade dos neurônios da rede, as camadas ocultas revelam computacionalmente as relações desconhecidas entre os dados em nível profundo de abstração (MASSAOUDI *et al.*, 2021). Do início ao final da rede neural, a cada camada oculta, há uma transformação gradual dos dados – complicados no espaço das entradas – até que sejam linearmente separáveis no espaço das saídas (NAITZAT; ZHITNIKOV; LIM, 2020).

A Figura A.1 apresenta a estrutura de um *perceptron* multicamadas (MLP) contendo dois atributos de entrada (x_1 e x_2), três camadas ocultas ($h^{(1)}$, $h^{(2)}$ e $h^{(3)}$) de quatro neurônios (h_1, h_2, h_3 e h_4), vieses de entrada para todas as camadas (x_0 ou h_0), e uma única saída real (\hat{y}_1), tudo densamente conectado. Cada neurônio das camadas ocultas e de saída opera segundo a lógica da Figura 13, porém com a substituição da função de ativação de forma apropriada para o tipo de problema ou sua função na arquitetura. Os neurônios da camada anterior compõem as entradas da nova camada.

Figura A.1 – Exemplo de um *Perceptron* Multicamadas (MLP)



Fonte: elaboração própria (2022).

O objetivo nesse tipo de aprendizado supervisionado é encontrar os pesos e vieses que garantam a menor diferença entre os valores reais e os previstos para a saída da rede neural. Os vieses, como na Figura A.1, costumam ser tratados como constantes de valor unitário por camada (exceto a de saída), mas aplicados a um vetor de pesos, como os demais neurônios. Do ponto de vista da camada seguinte, são como um neurônio comum da camada anterior.

O processo completo de ajuste dos pesos e vieses até a obtenção dos

valores para atingir as saídas desejadas – ou seja, a otimização – é chamado de treinamento da rede neural. A otimização é considerada a espinha dorsal do aprendizado das redes neurais artificiais. Ela se utiliza de algoritmos de aprendizado derivativos, que podem ser de primeira ou de segunda ordem. Os métodos de derivação de primeira ordem – gradiente descendente – utilizam informações do gradiente para construir a nova iteração, enquanto os métodos de segunda ordem (NESTEROV, 2018) – Newton, gradiente conjugado, quasi-Newton, Gauss-Newton, Levenberg-Marquardt etc. – usam a Hessiana para computar a próxima iteração com base na trajetória da otimização. Devido ao uso computacional excessivo para o cálculo da hessiana – principalmente de memória, pois é preciso computar a amostra de treinamento completa –, os métodos de segunda ordem são menos utilizados na prática, embora venham ganhando destaque com o avanço do poder computacional e de alocação de memória. Os algoritmos de otimização de primeira ordem não performam tão bem quanto os de segunda ordem e são dependentes de um número maior de ajustes finos dos hiperparâmetros para se ajustar a uma trajetória ótima na superfície dos erros. No entanto, várias melhorias foram desenvolvidas para reduzir o número de hiperparâmetros dos algoritmos de primeira ordem através de soluções adaptativas (TAN; LIM, 2019).

Há múltiplos algoritmos de treinamento de primeira ordem das redes neurais, mas a otimização é mais comumente solucionada através de um dos métodos do gradiente descendente com *back propagation* (SALOMON, 1998; UPADHYAYULA; VENKATARAMANAN, 2020). O gradiente descendente por *back propagation* (RUMELHART; HINTON; WILLIAMS, 1986 apud KUHN; JOHNSON, 2013) é uma técnica de treinamento onde, inicialmente, calculam-se os gradientes da função de custo em relação a cada parâmetro do modelo (todos os pesos e vieses) e, em seguida, executa-se uma etapa de redução dos pesos (e novos gradientes) em função desses gradientes. Essa etapa de retropropagação normalmente é executada milhares ou milhões de vezes, usando muitos exemplos de treinamento, até que os parâmetros do modelo convirjam para valores que – espera-se – minimizem a função de custo (GÉRON, 2019). O algoritmo de *back propagation* é uma metodologia muito eficiente na busca dos parâmetros ótimos, mas não garante que a hipótese resultante seja uniformemente melhor que qualquer outra (KUHN; JOHNSON, 2013).

Os pesos iniciais devem ser definidos de forma aleatória para quebrar a simetria e permitir a diversidade dos neurônios, e nunca devem ser iguais ou estar zerados, pois neste caso o processo de *back propagation* levaria à atualização idêntica de todos eles, pois apesar do tamanho que tiver a rede neural, cada nível seria tratado como se houvesse um único neurônio, com todos pesos de um mesmo nível atualizados da mesma forma a cada iteração do algoritmo do gradiente descendente (GÉRON, 2019).

Até o presente, os estudos que tratam da eficiência na utilização de um certo número de camadas ou neurônios para a solução dos problemas reais são limitados, sendo majoritariamente estudos empíricos envolvendo casos e bases de dados abertos específicas – como Golubeva, Neyshabur e Gur-Ari (2020) –, ou opiniões em função da experiência adquirida – em fóruns de discussão como StackExchange (ATWOOD; SPOLSKY, 2022) – mas que não provam a generalização de aplicações para as características propostas.

Hornik, Stinchcombe e White (1989) demonstram que as redes neurais de aprendizado profundo são capazes de aproximar qualquer função mensurável e com qualquer grau de precisão, desde que um número suficientemente grande de unidades ocultas (camadas ou neurônios) seja fornecido. O eventual insucesso em um caso específico está relacionado com:

- a) o aprendizado inadequado;
- b) um número insuficiente de unidades ocultas (neurônios);
- c) a relação estocástica e não determinística entre as variáveis de entrada e as saídas.

Para o sucesso do treinamento pode ser necessária uma quantidade de dados extraordinariamente grande, com tempo de execução impraticável (AGGARWAL, 2018). Além disso, as RNAs estão sujeitas às instabilidades do processo numérico envolvido e ao sobreajuste. Embora o sobreajuste seja menor no caso das redes neurais do que em outros métodos de AM, várias técnicas podem ser necessárias, como: regularização, parada antecipada – ou *early stopping* –, balanço entre número de neurônios (ou parâmetros da rede) e número de camadas ocultas (profundidade), e combinação de métodos – ou *ensemble methods* (ZANG *et al.*, 2020).

Um dos problemas das RNAs, e mais especificamente do MLP associado ao método do gradiente descendente com *back propagation*, é sua dificuldade em extrapolar a maioria das funções não lineares, embora esteja provada sua capacidade no caso de funções lineares, se os dados de treinamento forem diversos o suficiente (XU *et al.*, 2020).

Alguns problemas de regressão não se beneficiam da aplicação das redes neurais (solução genérica), notadamente quando já existe um algoritmo tradicional especializado para a solução, pois estas podem depender da busca de ajustes ótimos para atingir seu melhor desempenho, o que nem sempre compensa o esforço a dispendar, sem a garantia de chegar ao ótimo. Outro ponto é a eventual insuficiência de dados de entrada, o que pode limitar a capacidade de aprendizado de um modelo para o problema.

A.1 Método de otimização do gradiente descendente

Originalmente, o gradiente descendente utiliza todos os exemplos da amostra de treinamento no processo de otimização ao mesmo tempo, o que resulta em lentidão em sua execução, como no caso dos métodos de segunda ordem. Nessa configuração é conhecido por *batch gradient descent*. Desta forma, sua variação, conhecida por gradiente descendente estocástico (SGD) (ROBBINS; MONRO, 1951), e as variações dela são mais utilizadas.

O SGD executa, a cada iteração completa dos exemplos de treinamento – ou época – várias iterações parciais envolvendo lotes de exemplos escolhidos aleatoriamente de cada vez, em geral sem repetição. Assim, vários gradientes parciais gerados a partir de subamostras menores são tratados em série com muito mais facilidade computacional. Como efeito-colateral tem-se um processo menos estável, com variações da função de custo a favor e contra o objetivo, até atingir um valor próximo do mínimo, mas nunca o próprio (GÉRON, 2019).

O treinamento de uma RNA com o SGD é o equivalente a resolver o problema de otimização não-convexa, matematicamente representado por:

$$\min_{w \in \mathbb{R}^n} f(w) = \frac{1}{M} \sum_{i=1}^M f_i(w), \quad (\text{A.1})$$

sendo que: f_i é a função de custo para o ponto de dados $i \in \{1, 2, \dots, M\}$; e w é o vetor dos pesos sendo otimizados. Cada passo do SGD pode ser descrito por:

$$w_{k+1} = w_k - \alpha \left(\frac{1}{|B_k|} \sum_{i \in B_k} \hat{\nabla} f_i(w_k) \right), \quad (\text{A.2})$$

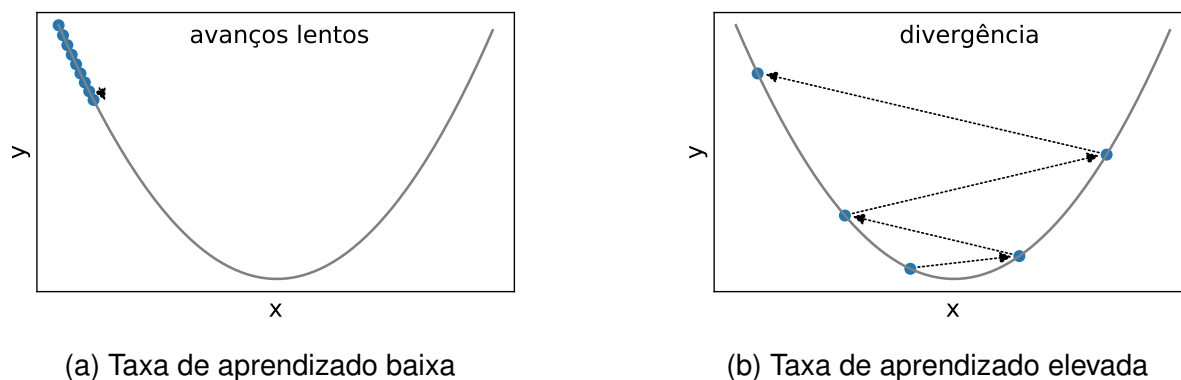
sendo que: w_k denota a k -ésima iteração; α é um tamanho de avanço (ajustado), conhecido por taxa de aprendizado; $|B_k|$ é o tamanho do lote de treinamento; e $\hat{\nabla} f_i(w_k)$ é o gradiente estocástico de w_k para o ponto de dados i (KESKAR *et al.*, 2017).

Os parâmetros de ajuste das redes neurais – diferentes dos pesos e vieses – são chamados de hiperparâmetros, como é o caso da taxa de aprendizado ou do tamanho do lote do SGD.

Se a taxa de aprendizado é muito baixa, o algoritmo leva muitas iterações até a convergência, ou seja, muito tempo computacional. Porém, se essa taxa é elevada, o algoritmo pode até mesmo divergir, ao ultrapassar o valor mínimo na superfície das soluções. A Figura A.2 ilustra essas situações.

Além disso as funções de custo podem apresentar todo tipo de superfície irregular, e dependendo da inicialização dos pesos e vieses, pode atingir um mínimo

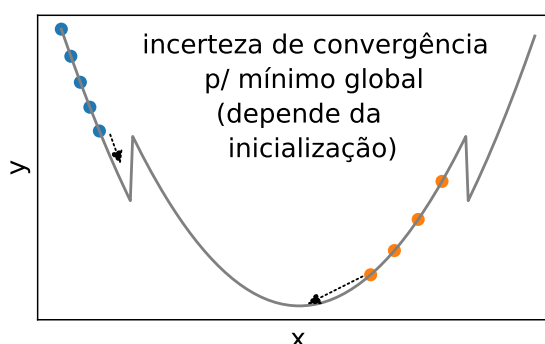
Figura A.2 – Convergência da otimização para diferentes taxas de aprendizado



Fonte: adaptado de Géron (2019).

local, como ilustra a Figura A.3.

Figura A.3 – Convergência da otimização em superfície irregular

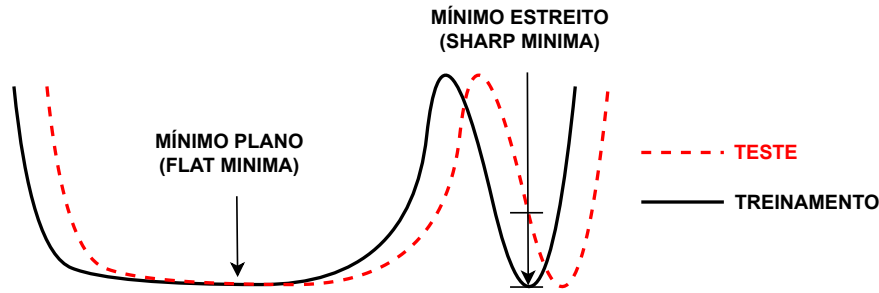


Fonte: adaptado de Géron (2019).

Já o tamanho do lote de treinamento é tipicamente $|B_k| = \{32, 64, \dots, 512\}$. Esses valores têm sido usados com sucesso em várias aplicações mencionadas por Keskar *et al.* (2017). Um efeito-colateral dos métodos em regime de pequenos-lotes – ou *small-batches* – é que são difíceis de paralelizar já que são calculados iterativamente em sequência. De outro lado, experiências com o aumento do tamanho do lote, visando a paralelização, demonstraram que o desempenho em teste é reduzido quando aplicados lotes grandes (LECUN *et al.*, 2012 apud KESKAR *et al.*, 2017). Keskar *et al.* (2017) demonstra empiricamente que com o crescimento do tamanho do lote de treinamento tem-se uma tendência de obtenção de um mínimo estreito – ou *sharp minima* – com baixo desempenho em teste, diante de dados nunca vistos, independentemente dos bons resultados de validação, pois qualquer variação da amostra de teste conduz a erros muito elevados. Já os lotes pequenos relacionam-se com os mínimos planos, menos sensíveis às variações da amostra de treinamento. A Figura A.4 ilustra esse fenômeno.

Sutskever *et al.* (2013) testaram e sugerem o uso do chamado SGD com momento (SGDM) (POLYAK, 1964), uma versão do SGD que utiliza a inércia das

Figura A.4 – Exemplo do mínimo estreito resultante de lotes grandes de treinamento



Fonte: adaptado de Keskar *et al.* (2017).

iterrações para acelerar o processo de treinamento na direçāo relevante e amortecer oscilações (RUDER, 2016). Os passos do SGDM podem ser descritos conforme as Equações A.3 e A.4, onde $\mu \in [0, 1)$ é um parâmetro do momento e v_0 é inicializado zerado (KESKAR *et al.*, 2017; SUTSKEVER *et al.*, 2013).

$$v_{k+1} = \mu v_k - \alpha \left(\frac{1}{|B_k|} \sum_{i \in B_k} \hat{\nabla} f_i(w_k) \right) \quad (\text{A.3})$$

$$w_{k+1} = w_k + v_{k+1} \quad (\text{A.4})$$

Uma variaçāo do SGDM procura impedir um salto muito rapido na direçāo relevante, em prol de um desempenho maior da RNA (RUDER, 2016), o chamado momento de Nesterov (NESTEROV, 1983). Essa tecnica acumula um vetor de velocidade na direçāo das reduçōes persistentes da funçāo de custo ao longo das iterrações (SUTSKEVER *et al.*, 2013). As Equações A.5 e A.6 apresentam a representaçāo de um passo do SGDM c/ Nesterov.

$$v_{k+1} = \mu v_k - \alpha \left(\frac{1}{|B_k|} \sum_{i \in B_k} \hat{\nabla} f_i(w_k + \mu v_k) \right) \quad (\text{A.5})$$

$$w_{k+1} = w_k + v_{k+1} \quad (\text{A.6})$$

Mas o SGDM tambem tem a desvantagem de escalar o gradiente uniformemente em todas as direçōes (relativas às suas componentes), o que pode ser prejudicial ao treinamento em alguns problemas e ao processo de seleçāo da taxa de aprendizado. Assim, varios metodos adaptativos foram sendo propostos, os quais escalam o gradiente diagonalmente atraves de estimativas da curvatura da funçāo de custo, como Adadelta (DUCHI; SINGER, 2011), Adam (KINGMA; BA, 2014) e RMSprop (HINTON; SRIVASTAVA; SWERSKY, 2012), por exemplo.

A.2 Funções de custo

As funções de custo – ou *loss functions* – são empregadas para expressar o valor de uma determinada utilidade durante o processo de otimização. Ela pode ser definida como a (quantidade de) utilidade perdida em função da previsão realizada, em comparação com o valor real. O algoritmo de AM pode teoricamente maximizar a utilidade selecionando a hipótese que minimiza o erro de generalização esperado com base nos pares entrada-saída conhecidos (RUSSELL; NORVIG, 2016).

As razões pela qual o resultado da otimização não coincide com o resultado real podem ser: irrealidade, variância, ruído, e complexidade computacional. A hipótese ótima pode não ser realizável, não estar no conjunto das possíveis soluções do algoritmo, ou estar presente numa forma preterida pelo método de otimização. O algoritmo retornará diferentes hipóteses para diferentes conjuntos de exemplos de entrada, mesmo que essas amostras sejam originadas de uma mesma função real, e essas hipóteses farão diferentes previsões para exemplos nunca vistos. O problema pode não ser determinístico ou ter muito ruído, e ruídos não podem ser previstos, por definição. Pode ser computacionalmente intratável a busca em todo espaço das hipóteses de solução (RUSSELL; NORVIG, 2016).

Alguns exemplos de funções de custo utilizadas na solução dos problemas de regressão através das RNAs são:

- a) MSE: é conhecido por função de custo ℓ^2 – ou ℓ^2 *loss function* – e é definido por:

$$\text{MSE}(X, h) = \frac{1}{m} \sum_{i=1}^m (h(w_i) - y_i)^2, \quad (\text{A.7})$$

sendo que: m é o número total de exemplos da amostra verificada; w_i é o valor dos pesos do i -ésimo exemplo da amostra X ; h é a transformação avaliada; e y_i é o correspondente valor real de saída – conhecido – do i -ésimo exemplo. Na prática, é o quadrado da RMSE, embora este último é que seja derivado dele. O MSE valoriza os erros maiores mais intensamente. É muito utilizado e possui derivadas parciais de segunda ordem, além de farta teoria estatística associada. É formado de duas componentes implícitas: o viés dos erros e sua variância. Portanto, atua nas duas frentes e tende ao valor médio dos erros (MURPHY, 2012);

- b) MAE: é conhecido por função de custo ℓ^1 – ou ℓ^1 *loss function*. O MAE como função de custo conduz a soluções menos sensíveis aos pontos fora da curva (*outliers*) e tende à mediana dos erros (MURPHY, 2012);
- c) Huber: funciona como uma composição das funções ℓ^1 e ℓ^2 de forma

controlada. Os erros menores são tratados por ℓ^2 e um parâmetro absoluto é utilizado para definir o ponto a partir do qual o erro é tratado por ℓ^1 . Sua formulação é dada por:

$$\text{Huber}(X, h) = \begin{cases} \frac{1}{2} (y_i - h(w_i))^2, & |y_i - h(w_i)| \leq \delta \\ \delta |y_i - h(w_i)| - \frac{1}{2} \delta^2, & \text{demais casos} \end{cases}, \quad (\text{A.8})$$

sendo que, δ é o valor absoluto do erro a partir do qual há a transição do MSE para o MAE;

- d) LogCosh: também funciona como uma composição das normas ℓ^1 e ℓ^2 , porém, com transição suave entre elas, algo necessário para os algoritmos de otimização de segunda ordem, e custo computacional maior. Sua formulação é dada por:

$$\text{LogCosh}(X, h) = \sum_{i=1}^m \log (\cosh(h(x_i) - y_i)), \quad (\text{A.9})$$

sendo que a função \cosh é o cosseno hiperbólico¹.

A.3 Funções de ativação

As funções de ativação são parte fundamental da estrutura de uma RNA pois são as responsáveis pela capacidade de representação das características não-lineares do problema, mapeando os dados de forma a torná-los linearmente separáveis (AGGARWAL, 2018). O aumento do número de neurônios (funções de ativação em paralelo) ou, principalmente, de camadas ocultas (funções de ativação em série) da RNA amplificam esta habilidade. Além disso, algumas vezes as funções de ativação têm outros propósitos, como o de restringir a faixa de valores de saída da RNA. Em problemas de classificação podem ser usadas para tornar as saídas booleanas, em regressões podem ser usadas para limitar as saídas aos valores reais positivos, dentre outros usos.

As funções de ativação mais triviais são: identidade, sigmoideal, tangente hiperbólica e unidade linear retificada (UPADHYAYULA; VENKATARAMANAN, 2020). Outras funções foram sendo criadas para endereçar problemas específicos, dificuldades de treinamento, generalização etc. Duas das funções de ativação de especial interesse nesse trabalho são: unidade linear retificada (ReLU) e unidade linear exponencial (ELU), ilustradas pela Figura A.5.

¹O cosseno hiperbólico é uma função da geometria hiperbólica, definido como: $\cosh(z) = \frac{1}{2} (e^z + e^{-z})$. Na área da engenharia elétrica é utilizado, por exemplo, no dimensionamento de linhas de transmissão longas de alta tensão e corrente alternada.

A ReLU (NAIR; HINTON, 2010) – também conhecida por função rampa – é a função de ativação mais usada em todas as formas de redes neurais. Sua formulação é dada por:

$$f(x) = \max(0, x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}. \quad (\text{A.10})$$

Para o AP, ela oferece fácil otimização e treinamento em relação a outras funções de ativação não lineares, pela simplicidade de obtenção de seu gradiente diante de múltiplas camadas ocultas de grande número de neurônios. A função tem como saída o próprio valor de entrada, se positivo, caso contrário zero. Assim, seu gradiente, dado por:

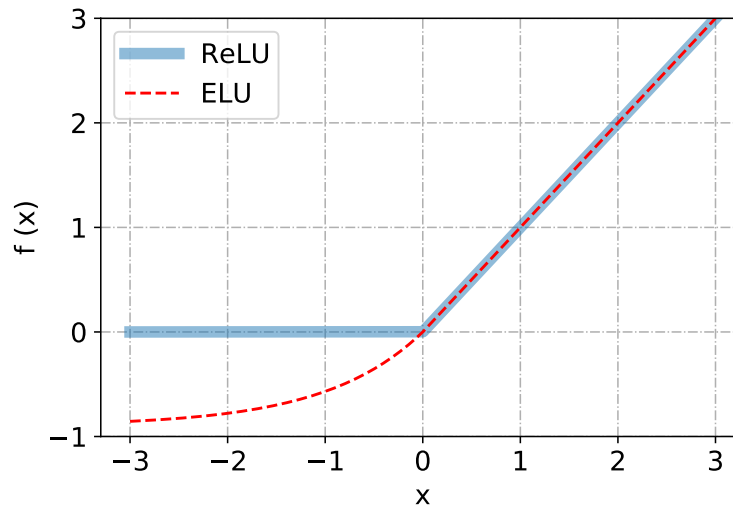
$$f'(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}, \quad (\text{A.11})$$

assume a unidade para qualquer entrada no intervalo $[0, \infty)$.

Em algumas situações, seja pelo valor das entradas em função da inicialização aleatória dos pesos e vieses, ou pela velocidade do aprendizado, o gradiente da função ReLU pode se tornar zero, e a partir de então o neurônio é inativado (ou "morto"), situação referida por ReLU morta – ou *dying* ReLU. Esse neurônio não participa mais da solução. No limite, a rede neural pode atingir um ponto em que vários neurônios são inativados e não há mais variação das saídas, independentemente das entradas (UPADHYAYULA; VENKATARAMANAN, 2020).

Mas isso pode ser positivo, em certas circunstâncias, pois ter alguns neurônios rapidamente fora do problema funciona para tornar a RNA mais esparsa, o que aumenta a velocidade de convergência, embora com tendência observada para o sobreajuste. Naitzat, Zhitnikov e Lim (2020) propõem que a descontinuidade da ReLU tem papel importante na capacidade de transformação dos dados de entrada de um espaço não separável em uma saída do espaço linearmente separável.

Várias alternativas foram criadas para reduzir a possibilidade de se atingir esse ponto de "morte" da função de ativação. A unidade linear exponencial (ELU) – comparada graficamente com a ReLU na Figura A.5 – foi proposta com o objetivo de tornar a média das ativações próximas de zero, prática que oferece alguns benefícios para a estabilidade do treinamento, como detalhado por Clevert, Unterthiner e Hochreiter (2016).

Figura A.5 – Sobreposição das funções de ativação ReLU e ELU

Fonte: elaboração própria (2022).

Com sua formulação, dada por:

$$f(x) = \begin{cases} x, & x > 0 \\ \alpha (e^x - 1), & x \leq 0 \end{cases}, \quad (\text{A.12})$$

os valores negativos de entrada ainda geram uma saída residual, e os neurônios conseguem persistir durante o treinamento, podendo ser eventualmente reativados, o que evita a *dying* ReLU. Porém, o gradiente da função ELU,

$$f'(x) = \begin{cases} 1, & x > 0 \\ f(x) + \alpha, & x \leq 0 \end{cases}, \quad (\text{A.13})$$

é mais difícil de computar, o que torna o processo de treinamento mais lento.

A percepção deste autor, a partir da literatura, é que devido à natureza "caixa preta" das RNAs, as proposições para a estrutura ou os hiperparâmetros das redes são empíricas e dependentes da metodologia de testes utilizada; há estudos e proposições conflitantes, ou que procuram resolver problemas indetectáveis sem muita experimentação, e que podem adicionar algum efeito colateral (complexidade, tempo etc.).

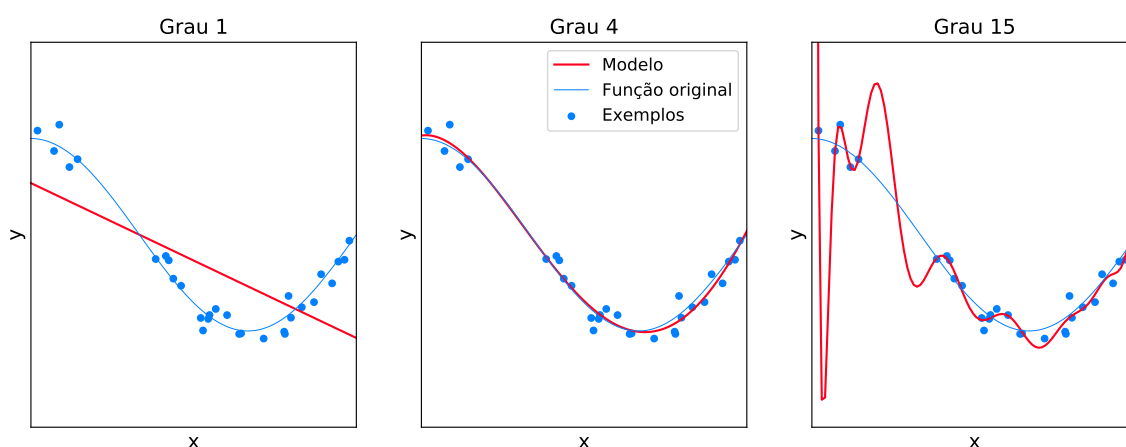
A.4 Critérios de parada do treinamento

Ao ajustar ou treinar modelos flexíveis de AM, é preciso ter cuidado para não causar o sobreajuste – ou *overfitting* –, ou seja, evitar que o modelo reproduza fielmente as menores variações dos dados de entrada, porque em geral, essas representam

apenas ruídos (MURPHY, 2012). Na outra ponta, o modelo pode não conseguir prever minimamente o caso real, o que se chama de sub-ajuste – ou *underfitting*.

A Figura A.6 exemplifica três regressões lineares polinomiais de ordem 1, 4 e 15 para os pontos gerados a partir da função cosseno com ruído aleatório uniforme de $\pm 10\%$. O primeiro gráfico representa o sub-ajuste, onde o modelo foi incapaz de representar o caso real. No segundo gráfico percebe-se uma aderência da solução ao problema real. No terceiro gráfico, o sobreajuste é caracterizado, o modelo aprendeu os ruídos presentes nos dados de treinamento e os incorporou à solução. Embora o modelo consiga representar melhor cada um dos pontos conhecidos, ele certamente falhará diante do objetivo da regressão de prever novas entradas, nunca vistas.

Figura A.6 – Regressões polinomiais de pontos da função cosseno com ruído



Fonte: adaptado de Pedregosa *et al.* (2011).

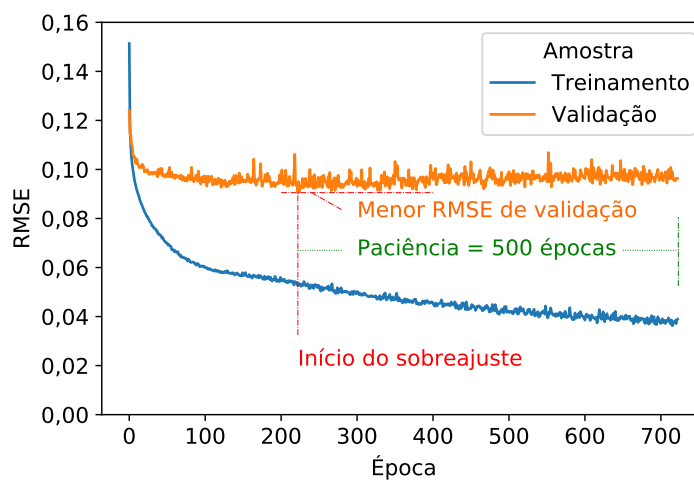
Para a construção dos modelos de AM é comum a segregação de uma parte dos dados para teste final dos resultados, além da aplicação de validação cruzada durante as fases de decisão da estrutura e dos hiperparâmetros, como visto na subseção 2.3.3. Mas com métodos iterativos, é importante ter alguma métrica – em tempo de execução – para decisão do momento de parada.

Durante a validação cruzada, as métricas relacionadas à parcela de validação correspondente a cada parcela de treinamento são usadas como referência para o início do sobre-ajuste. Senão, há a segmentação da amostra de treinamento para criar duas subamostras, de treinamento e validação, geralmente na proporção 75/25 ou 80/20, dependendo da disponibilidade de dados e do risco ao reduzir a representatividade dos dados de treinamento, não sendo incomuns as concentrações maiores na amostra de treinamento (AGGARWAL, 2018; GÉRON, 2019).

A Figura A.7 apresenta um exemplo real de evolução da função de custo durante o treinamento. Pode-se notar que a RMSE das parcelas de treinamento e validação caem fortemente durante as épocas iniciais, até que a parcela de validação

atinge seu mínimo e começa a piorar, enquanto a parcela de treinamento continua a melhorar. Como a evolução do método iterativo de otimização nem sempre é contínua, mas composta de avanços e retrocessos, não é possível saber de antemão se um determinado resultado parcial ainda poderá ser batido com o prosseguimento da execução do algoritmo. Então, costuma-se estabelecer um número máximo de épocas de treinamento da RNA a partir do qual considera-se que não haverá mais chance significativa de melhora da generalização da solução. Esse processo é conhecido por parada antecipada – ou *early stopping* (MURPHY, 2012).

Figura A.7 – Caracterização do sobreajuste durante o treinamento



Fonte: elaboração própria (2022).

Uma vez encerrado o treinamento, as matrizes dos pesos e vieses da época de menor validação são recuperadas para o modelo.

A.5 Métodos de regularização em caso de sobreajuste

A regularização é uma técnica para a penalização da complexidade da solução obtida pelo algoritmo de aprendizado de máquina, durante o treinamento, o que previne ou retarda o sobreajuste (RUSSELL; NORVIG, 2016). Trata-se de uma parcela adicionada à função de custo para a otimização. Nesse trabalho, são considerados três tipos de funções de regularização das RNAs: lasso, ridge e a combinação de ambas, a chamada rede elástica.

Lasso – do inglês *least absolute shrinkage and selection operator* –, ou regularização ℓ^1 , adiciona à função de custo a norma do vetor de pesos (W) da rede neural,

$$\|W\|_1 = \sum_{i=1}^n |w_i|, \quad (\text{A.14})$$

com o objetivo de tornar a solução mais esparsa (MURPHY, 2012), ou seja, descartar os pesos da rede neural que tenham pouca, a mesma de outro atributo correlacionado, ou nenhuma contribuição para a solução.

Já a ridge, ou regularização ℓ^2 , ou ainda decaimento de peso (do inglês *weight decay*), atua pela norma euclidiana,

$$\|W\|_2 = \sum_{i=1}^n (w_i)^2, \quad (\text{A.15})$$

para minimizar os pesos mais extremos da rede neural.

Assim, enquanto a norma ℓ^1 tende a reduzir o número de variáveis explicadoras (atributos) da solução, a norma ℓ^2 tende a reduzir o protagonismo de algumas delas. Zou e Hastie (2005) introduziram a chamada rede elástica, do inglês *elastic net*, que é uma combinação convexa de lasso e ridge. Essa estratégia pode apresentar performance superior a lasso ou ridge isoladamente, compartilhando o efeito de representação mais esparsa de lasso à minimização de extremos de ridge, e com tendência a agrupar e descartar – conjuntamente – os atributos muito correlacionados. A penalidade da rede elástica,

$$(1 - \alpha)\|W\|_1 + \alpha\|W\|_2, \quad (\text{A.16})$$

é estritamente convexa para qualquer valor α no intervalo $(0, 1)$.

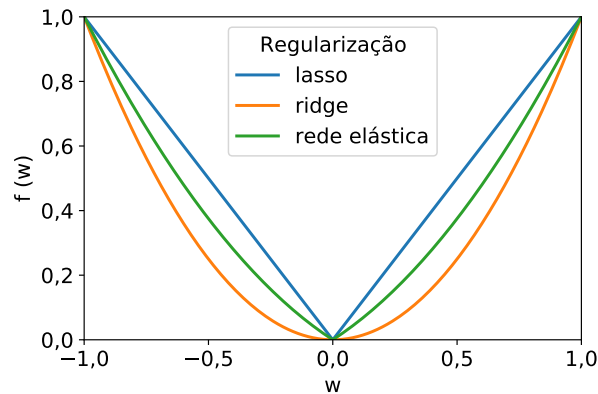
A Figura A.8 ilustra o efeito de cada uma dessas estratégias de regularização: linear no caso de lasso, quadrática no caso de ridge e uma composição dos dois efeitos na rede elástica. A adição dessa penalização (r) à função de custo original (f) gera uma nova função de custo (g),

$$g(X, h) = f(X, h) + \lambda r(W), \quad (\text{A.17})$$

sendo que um novo hiperparâmetro absoluto (λ) de ajuste é utilizado para ponderar o peso da regularização no caso concreto.

A.6 Normalização de entradas para estabilidade do treinamento

Os algoritmos de AP não oferecem boa performance quando os valores dos atributos de entrada possuem diferentes escalas (grandezas ou proporções), com raras exceções, porque as componentes de maior (desproporcional) gradiente tendem a direcionar o caminho da otimização, causando instabilidade e demandando um número maior de iterações até a solução. O processo de redimensionamento das variáveis de

Figura A.8 – Componentes de regularização da função de custo

Fonte: elaboração própria (2022).

entrada para a uniformização da escala é chamado de escalonamento de atributos – ou *feature scaling* (GÉRON, 2019).

Nesse trabalho, lida-se com dois tipos de escalonamento das entradas: normalização (ou escalonamento min-max) e estandardização. Na normalização os dados de cada atributo, analisados individualmente, são convertidos para uma nova escala, usualmente entre -1 e 1 (ou 0 e 1), seguindo a formulação:

$$f(x) = a + \frac{(x - x_{min})(b - a)}{x_{max} - x_{min}}, \quad (\text{A.18})$$

sendo que:

a e b são, respectivamente, os limites inferior e superior desejados após a transformação do atributo escalonado;

x_{min} é o menor valor do atributo na amostra de treinamento;

x_{max} é o maior valor do atributo na amostra de treinamento.

No caso da estandardização, também atributo a atributo, calcula-se a média (\bar{x}) e o desvio-padrão (σ) da amostra de treinamento, e aplica-se a transformação:

$$f(x) = \frac{x - \bar{x}}{\sigma}, \quad (\text{A.19})$$

a todos os elementos da amostra. Nesse caso da estandardização, os valores resultantes não têm limites, mas a média é zero e o desvio-padrão é a unidade.

Uma vez aplicado o escalonamento antes do treinamento, suas premissas (a e b , ou \bar{x} e σ) devem ser mantidas para serem aplicadas – intactas – às entradas referentes a novos exemplos de teste ou para as previsões com o modelo gerado.

Os termos utilizados para referir-se aos diferentes tipos de escalonamento de atributos de entrada são diversos na literatura. Nesse trabalho, o escalonamento e a

estandardização aparecem, doravante, referidos por:

- a) normalização Minmax: para o escalonamento min-max entre -1 e 1 ;
- b) normalização Z-score: para a estandardização.


APÊNDICE B – DADOS DO NREL DE USINAS DOS EUA

O *National Renewable Energy Laboratory* (NREL) disponibiliza em seu sítio de internet alguns dados energéticos dos Estados Unidos, incluindo energias renováveis e combustíveis alternativos, para uso pela comunidade científica e pelos desenvolvedores através de APIs, como mostra a Figura B.1 (NREL, 2021a).

Os dados são segmentados em:

- a) *Buildings*: dados relacionados às medidas de conservação de energia em unidades comerciais e residenciais;
- b) *Electricity*: dados associados à estrutura tarifária e preços da geração, distribuição, transmissão e monitoramento de energia;
- c) *Energy Optimization*: ferramentas e modelos para otimização de energia renovável, geração convencional e armazenamento de energia para edificações, campus, comunidades e microrredes;
- d) *Partnering*: uma coleção de APIs provendo informações relacionadas às tecnologias energéticas, especialistas e patentes associadas do *Lab Partnering Service* (LPS) do *Department of Energy* (DOE) dos Estados

Figura B.1 – Sítio de internet *Developer Network* da NREL

Developer Network 

HOME DOCUMENTATION COMMUNITY

The National Renewable Energy Laboratory's developer network helps developers access and use energy data via Web services, including renewable energy and alternative fuel data.

Web Service Documentation ▶

Explore our available Web services.

- [Buildings](#)
- [Electricity](#)
- [Energy Optimization](#)
- [Partnering](#)
- [Solar](#)
- [Transportation](#)
- [Wave](#)
- [Wind](#)

API Key ▶

Sign up to use our APIs.

- [Get an API Key](#)

Community ▶

Join the discussion and stay updated.

- [Follow Us on Twitter](#)

[Help Improve this Content](#)

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

[Need Help?](#) | [Security & Privacy](#) | [Disclaimer](#) | [NREL Home](#)

Fonte: developer.nrel.gov (2021).

Unidos;

- e) **Solar**: serviços de dados e análises que provêm acesso aos dados de recursos solares e modelos de geração do NREL;
- f) *Transportation*: documentação sobre as tecnologias alternativas de distribuição e transmissão de energia;
- g) *Wave*: dados de recursos ondomotrizados e modelos de geração do NREL;
- h) *Wind*: dados de recursos eólicos e modelos de geração do NREL.

No segmento *Solar*, sempre no contexto dos Estados Unidos, são disponibilizados cinco grupos de APIs:

- a) *NSRDB Data Downloads*: diversos dados abertos do *National Solar Radiation Database* (NSRDB);
- b) **PVDAQ (PV Data Acquisition)**: acesso aos dados de performance de sistemas fotovoltaicos, coletados pela NREL;
- c) *PVWatts*: modelo de previsão mensal de geração de sistemas fotovoltaicos conectados à rede elétrica a partir de poucos dados de entrada;
- d) *Solar Dataset Query*: informação sobre a estação meteorológica mais próxima de uma localidade (ou coordenada geográfica);
- e) *Solar Resource Data*: dados anuais e mensais médios de insolação por localidade (ou coordenada geográfica).

O PVDAQ do NREL foi utilizado para a prospecção de usinas e seus dados de medição (dos Estados Unidos) para este trabalho. O serviço é subdividido em quatro APIs:

- a) **Aggregated Site Data**: dados de geração agregados (horários, diários, semanais ou mensais) do PVDAQ para o sistema e período selecionados, no formato IEC-61724;
- b) *Annual Data CSV for a System*: arquivo CSV de dados brutos anuais do PVDAQ para o sistema e ano selecionados (com a discretização variável de cada sistema);
- c) *Raw Data*: dados brutos do PVDAQ para o sistema e período selecionados (com a discretização variável de cada sistema);
- d) **Sites Metadata**: informações sobre as usinas do PVDAQ, podendo incluir proprietário, localização, dados da instalação, tecnologia envolvida nos painéis solares e sistema de potência.

A base de dados de usinas obtida através do API *Sites Metadata* continha 71

usinas com dados abertos. A Tabela B.1 apresenta um exemplo do tipo de informação disponível, podendo-se notar a presença de lacunas nos dados: nem todas usinas possuem informações completas na base de dados.

Tabela B.1 – Exemplos de registros de usinas do PVDAQ da NREL (como disponibilizados)

Campo	Exemplo 1	Exemplo 2
system id	1201	1332
available years	[2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020]	[2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020]
comments		Array tilt is a weighted average of the two tilts (main array at 10 deg, south facade at 60 deg)
confidential		
inverter mfg		Satcon
inverter model		PVS-500(x2), PVS-135(x1)
module mfg		SunPower
module model		318W(x3136), 315W(496)
module tech	1	1
site power	140140	1153488
name private	Distributed Sun - 5 Executive Campus	NREL, Parking Garage
name public	[1201] Distributed Sun - 5 Executive Campus	[1332] NREL Parking Garage
site area	959.058	5922.7534
site azimuth	230	180
site tilt	10	16.77
site elevation	18	1770
site latitude	39.9283	39.7388
site longitude	75.0481	105.1732

Fonte: elaboração própria (2021).

Assim, para a utilização dos dados da NREL foi necessário selecionar usinas que continham dados mínimos – suficientes e necessários – para a operacionalização da metodologia proposta. Também foi preciso inferir o significado de alguns dos campos de dados. As informações básicas das usinas e as inferências adotadas foram:

- a) *system id*: número de referência do sistema PV entre as bases de dados;
- b) *name private*: nome do usuário (e/ou proprietário);
- c) *site power*: potência nominal em W;
- d) *site elevation*: altitude em metros;

e) *site latitude*: latitude em graus da instalação;

f) *site longitude*: longitude em graus da instalação no sentido leste¹.

Para cada uma dessas usinas pode-se obter através do API *Aggregated Site Data* as medições da geração agregada, em base horária (por opção), para uma parcela ou a totalidade dos anos mencionados no campo *available years* (Tabela B.1), como mostra o exemplo da Tabela B.2.

Tabela B.2 – Exemplos de registros de medições de geração das usinas do PVDAQ da NREL (como disponibilizados)

Campo	Exemplo 1	Exemplo 2
measdatetime	2020-07-10 12:00:00	2020-08-20 12:00:00
array capture losses		-2.5033
array energy fraction	0.993679	1
array performance ratio		
array yield	0.046739	2.5033
availability	0.6	0.179167
bos efficiency	0.840967	3.1071
bos losses	0.007433	-5.2747
energy from array	6.55	2887.53
energy from grid	0.041667	0
energy to grid	5.55	8971.83
final yield	0.039353	7.778
load efficiency	0.841972	3.1071
mean array efficiency		
poa irradiation		0
reference yield		0
system id	1201	1332
system performance ratio		
total energy input	6.59167	2887.53
total energy output	5.55	8971.83
total system efficiency		

Fonte: elaboração própria (2021).

Da mesma forma que o API anterior, não há explicação suficiente dos campos de informação no sítio de internet da NREL. Assim, da comparação entre os valores, foram inferidos:

a) *system id*: número de referência do sistema PV entre as bases de dados;

¹ Parece comum a adoção nos EUA da referência à longitude leste, ou seja, sem o sinal negativo, em sistemas locais.

- b) *measdatetime*: ano, mês, dia, hora, minuto e segundo de início da agregação horária;
- c) *energy from array*: energia proveniente dos painéis fotovoltaicos em kWh;
- d) *energy from grid*: energia consumida da rede de distribuição em kWh (pelo inversor nos momentos de baixa insolação);
- e) *energy to grid*: energia líquida entregue à rede de distribuição no período em kWh;
- f) *total energy input*: energia total consumida pelo(s) inversor(es) solar(es) em kWh (soma de *energy from array* e *energy from grid*);
- g) *total energy output*: potência de saída do(s) inversor(es) solar(es) em kWh (mesmo que *energy to grid*).

Nem todas as usinas possuíam dados consistentes, como mostra o Exemplo 2 da Tabela B.2, onde os volumes de entrada (*total energy input*) são inferiores aos volumes de saída (*total energy output*).

APÊNDICE C – DADOS DO API DA OPENWEATHER

A OpenWeather (2020)¹ fornece dados históricos e previsões das condições atmosféricas de qualquer ponto do globo terrestre. Esta opção é interessante em função de que a empresa trabalha com a coleta de dados e a interpolação das condições do tempo para todas as coordenadas geográficas do globo terrestre, fornecendo dados históricos dos últimos 40 anos, além de previsões de tempo regulares e gratuitas para o horizonte desejado de previsão de 48 horas, o que facilita a integração por ferramenta de previsão.

Neste trabalho foram acessados os serviços meteorológicos da Tabela C.1.

Tabela C.1 – Fonte de Dados Meteorológicos

Fonte	Meio	Informação
OpenWeather	Arquivo CSV	Histórico das condições de tempo para as coordenadas geográficas de interesse. Esses dados instantâneos são fornecidos em intervalos horário e podem abranger até 40 anos. Há um custo de US\$ 10 por coordenada geográfica, independente do período desejado (OPENWEATHER, 2020).
	API	A empresa mantém um serviço de fornecimento de informações de previsão e histórico das condições de tempo por API. Tanto as previsões quanto o histórico recente são gratuitos.

Fonte: elaboração própria (2021).

A estrutura dos dados obtidos é apresentada pela Tabela C.2. A etiqueta temporal desses dados é baseada no UTC, o que lhes garante melhor confiabilidade ao excluir o impacto das variações sazonais de fusos horários das diversas regiões do planeta.

¹A OpenWeather é uma empresa de Londres (Reino Unido), composta por um time de especialistas e cientistas de dados meteorológicos atuantes desde 2014.

Tabela C.2 – Estrutura dos arquivos CSV/Json do serviço *History Bulk*

Campo	Descrição
city_name	Nome da cidade
lat	Coordenadas geográficas do local (latitude)
lon	Coordenadas geográficas do local (longitude)
main	
main.temp	Temperatura (°C)
main.temp_min	Temperatura mínima (°C)
main.temp_max	Temperatura máxima (°C)
main.feels_like	Sensação térmica (°C)
main.pressure	Pressão atmosférica ao nível do mar (hPa)
main.humidity	Umidade relativa do ar (%)
wind	
wind.speed	Velocidade do vento (m/s)
wind.deg	Direção do vento (°), meteorológica
clouds	
clouds.all	Cobertura de nuvens (%)
rain	
rain.1h	Volume de chuva da última hora (mm)
rain.3h	Volume de chuva das últimas 3 horas (mm)
snow	
snow.1h	Volume de neve da última hora (mm, em estado líquido)
snow.3h	Volume de neve das últimas 3 horas (mm, em estado líquido)
weather	
weather.id	Identificador da condição do tempo
weather.main	Grupo de condição do tempo (chuva, neve, extremo etc.)
weather.description	Condição do tempo
weather.icon	Ícone da condição do tempo
dt	Momento do cálculo (Unix, UTC)
dt_iso	Data e hora em formato UTC
timezone	Diferença em segundos a partir do UTC

Fonte: elaboração própria (2021).

APÊNDICE D – USINAS DE REFERÊNCIA

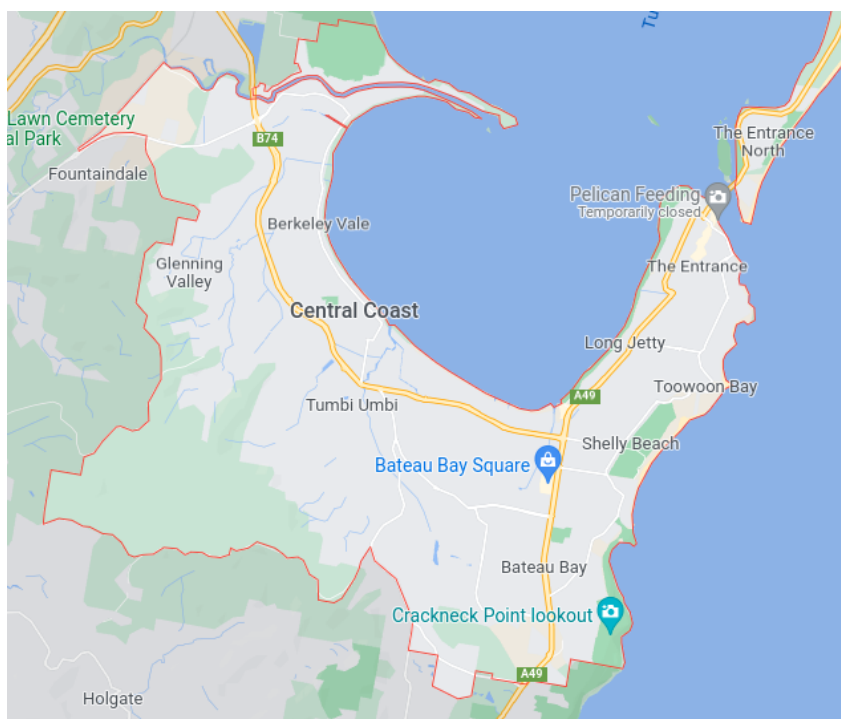
Este apêndice apresenta os fatos conhecidos ou assumidos verdadeiros a respeito das usinas cujos dados foram utilizados no desenvolvimento deste trabalho.

D.1 AUS: Usina residencial

A primeira usina considerada neste estudo é residencial e pertence à área de cobertura da concessionária Ausgrid¹ em New South Wales na Austrália, conforme material explicativo do Anexo B. Ela é referida por AUS neste trabalho.

Os dados públicos permitem saber apenas a região em que está localizada esta residência – código postal 2261, The Entrance North/NSW (Figura D.1) – e diferenciar os volumes de geração e de consumo² com discretização semi-horária.

Figura D.1 – Mapa da região de The Entrance North/NSW na Austrália



Fonte: Google Maps (2021c).

Esta usina possui dados íntegros ao longo do período de três anos de julho/2010 a junho/2013. Suas características são apresentadas pela Tabela D.1.

¹ A Ausgrid é a maior distribuidora de energia da costa leste da Austrália, fornecendo energia a cerca de 1,8 milhões de consumidores das áreas de Sydney, Central Coast e Hunter Valley.

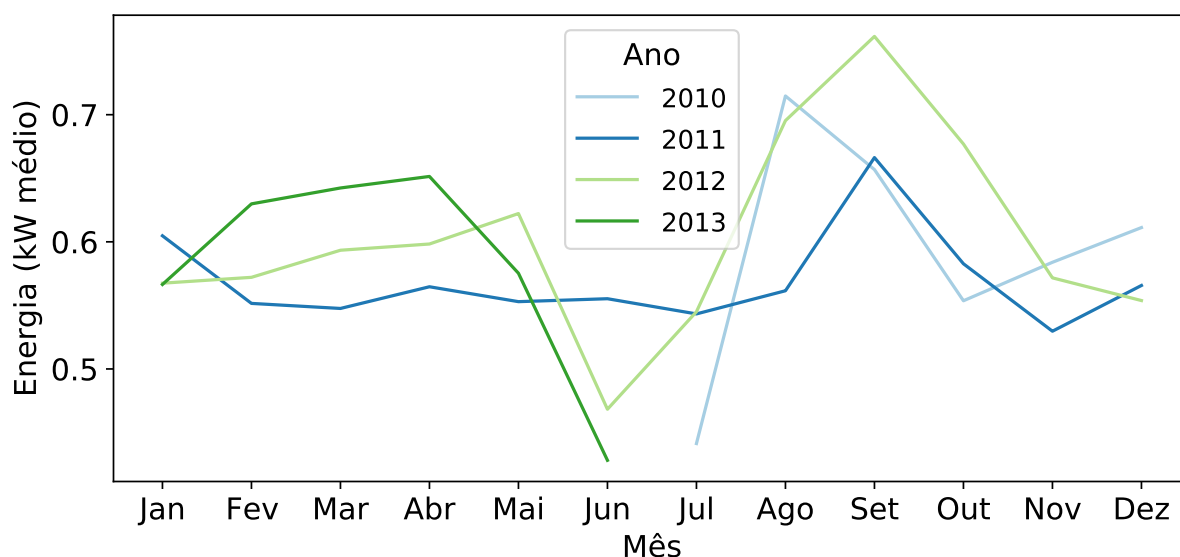
² O consumo não é utilizado neste trabalho. Vale destacar a atratividade desses dados para outras pesquisas envolvendo consumo e geração distribuída de energia (sistemas de armazenamento, microrredes etc.).

Tabela D.1 – Características da Usina AUS

Identificação Ausgrid	69
Local	
País	Austrália
Cidade	The Entrance North/NSW
Fuso Horário	UTC+10
Coordenadas geográficas	
Latitude	-33,36°
Longitude	151,43°
Altitude (aproximada)	25 m
Sistema FV	
Potência Nominal	2,04 kWp
Fator de Capacidade ³	28,8 %

Fonte: elaboração própria (2021).

A Figura D.2 apresenta a distribuição mensal da geração de energia da usina, ano a ano. Pode-se observar uma variação significativa ano a ano da geração média de cada mês.

Figura D.2 – Distribuição mensal da geração de energia da Usina AUS

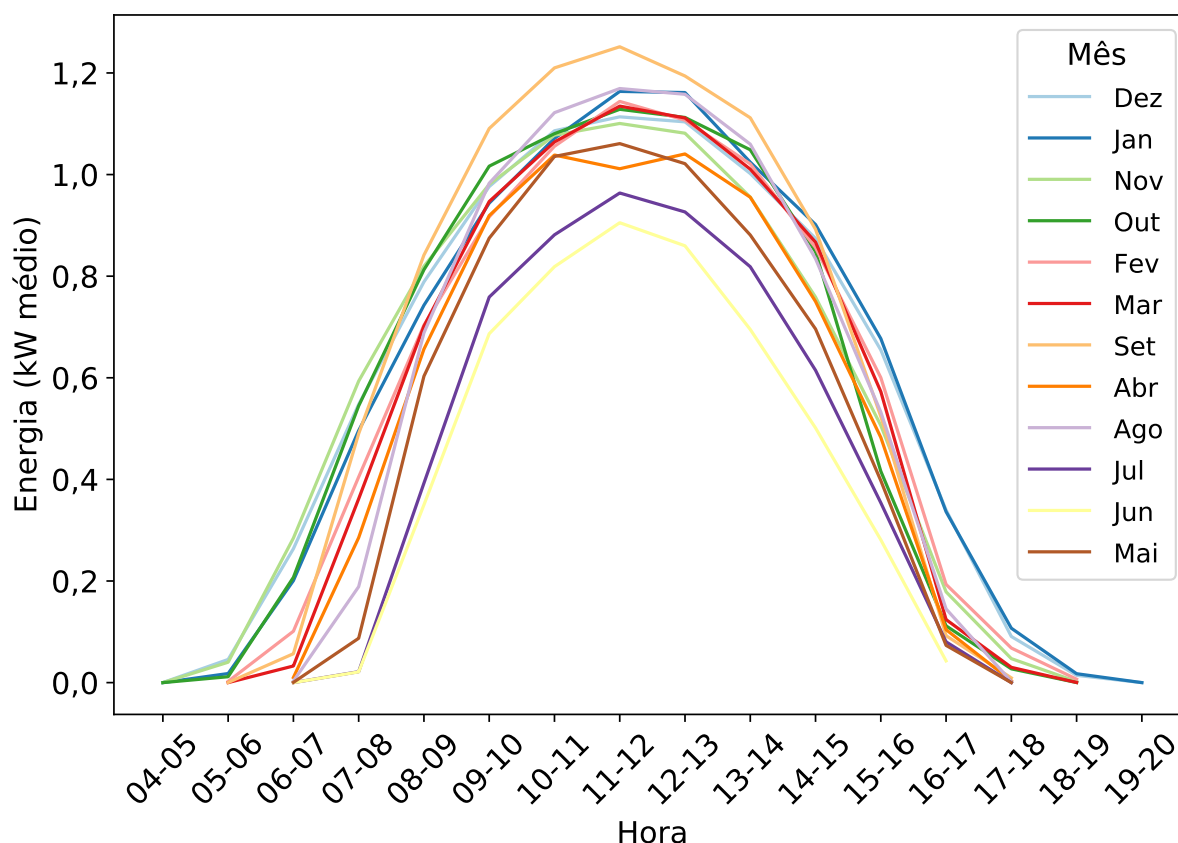
Fonte: elaboração própria (2022).

Já a Figura D.3 apresenta a distribuição horária média da geração de energia da usina por mês do ano (considerando todos os anos). Percebe-se a variação do

³Calculado a partir dos registros de medição para as horas claras do período de Jul/2010 a Jun/2013.

número de horas claras, com os dias começando mais cedo (no verão do hemisfério Sul) ou mais tarde (no inverno).

Figura D.3 – Distribuição horária da geração de energia da usina AUS



Fonte: elaboração própria (2022).

D.2 BRA: COGE AES Bauru

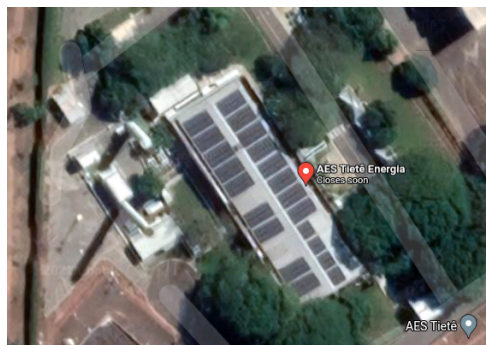
A segunda usina considerada neste estudo, referida por BRA, está instalada no telhado do Centro de Operações da Geração de Energia (COGE) da AES Brasil⁴ em Bauru, estado de São Paulo, Brasil (Figura D.4).

As características da usina são apresentadas pela Tabela D.2. O controlador da microrrede local mantém o acompanhamento da potência de saída do sistema de forma assíncrona. O dado horário de geração de energia obtido da AES Brasil, no contexto do projeto P&D ANEEL Microrredes, é calculado a partir da média das medições de potência em cada intervalo de uma hora.

A Figura D.5 apresenta a distribuição mensal da geração de energia da usina ano a ano. Pode-se visualizar a variação da média entre os anos para um mesmo mês.

⁴A AES Brasil é um dos maiores produtores privados de energia elétrica do Brasil. Com foco exclusivo em energia renovável, possui concessões hidrelétricas e fazendas fotovoltaicas no Estado de São Paulo e um complexo eólico na Bahia.

⁵Calculado a partir dos registros de medição para as horas claras do período 2019 e 2020.

Figura D.4 – Imagem de satélite do telhado do COGE AES Brasil

Fonte: Google Maps (2021b).

Tabela D.2 – Características da Usina BRA

Nome	COGE AES Brasil
Local	
País	Brasil
Cidade	Bauru/SP
Fuso Horário	UTC-3
Coordenadas geográficas	
Latitude	-22,2886°
Longitude	49,0986°
Altitude	600 m
Sistema FV	
Potência Nominal	95 kWp
Fator de Capacidade ⁵	31,0 %

Fonte: elaboração própria (2021).

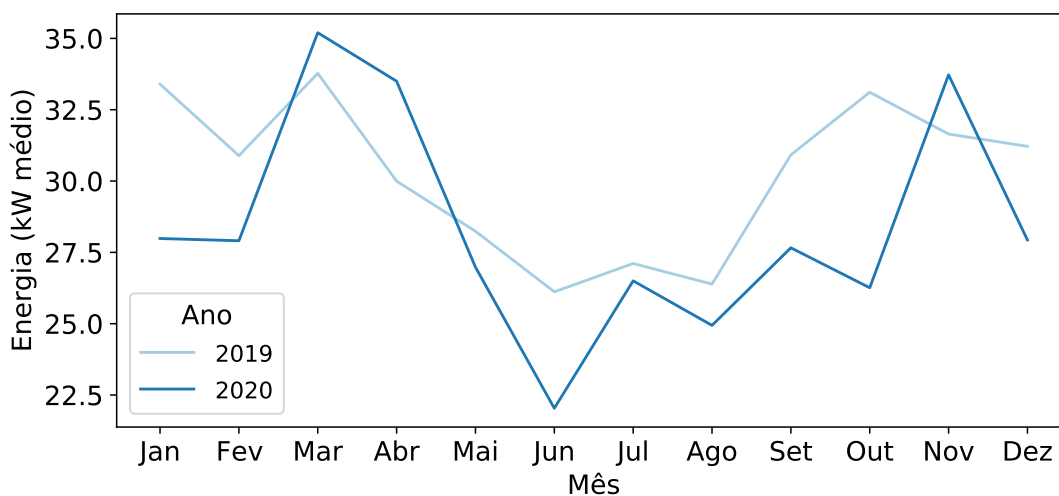
De 2019 para 2020, no mês de outubro, por exemplo, houve uma variação negativa de 20,7%, a maior do período observado.

A distribuição horária média da geração de energia elétrica da usina por mês do ano (considerando todos os anos) é apresentada pela Figura D.6.

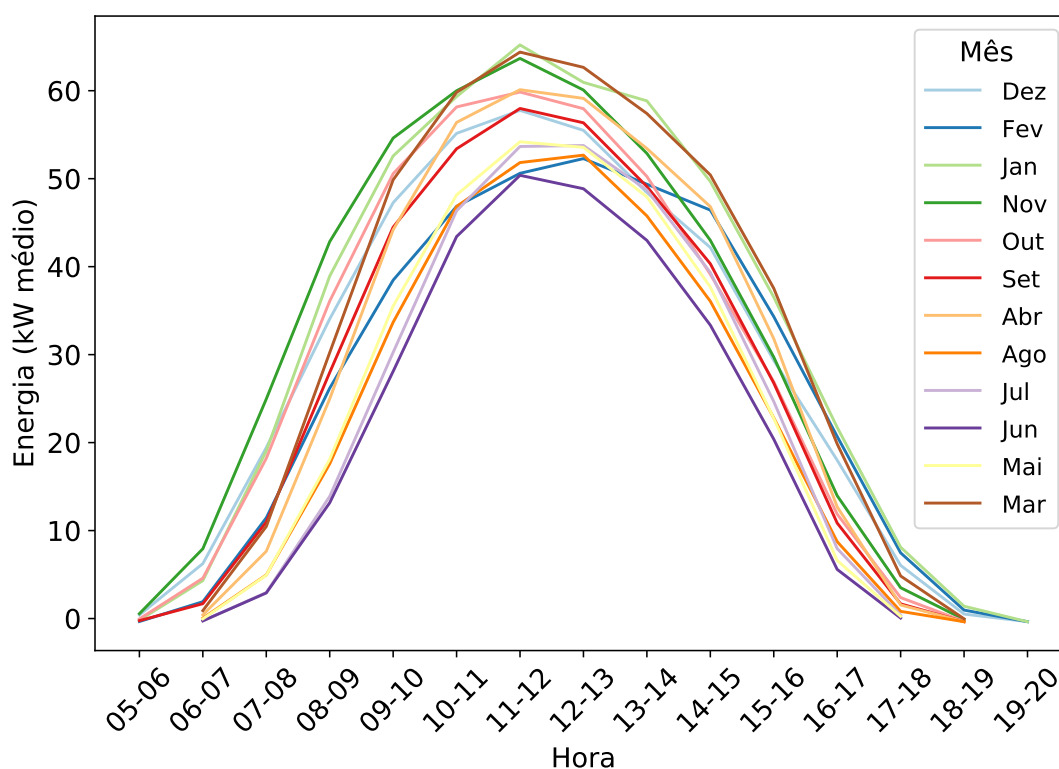
D.3 EUA: 5 Executive Campus

A terceira e última usina considerada neste estudo, referida EUA, está instalada no topo de um edifício comercial de escritórios chamado 5 Executive Campus, localizado na cidade de Cherry Hill, estado de New Jersey, nos Estados Unidos (Figura D.7).

As características da usina, de acordo com o banco de dados do NREL

Figura D.5 – Distribuição mensal da geração de energia da usina BRA

Fonte: elaboração própria (2022).

Figura D.6 – Distribuição horária da geração de energia da usina BRA

Fonte: elaboração própria (2022).

(NREL, 2021a), são apresentadas pela Tabela D.3.

Os dados de geração de energia da usina, obtidos conforme o Apêndice B, continuam registros horários para os anos de 2013 a 2019. Os dados até 25/01/2016 eram agregados de dados brutos com frequência de 3 minutos. Após essa data os

⁶Calculado a partir dos registros de medição para as horas claras do período de 2013 a 2019.

Figura D.7 – Imagem de satélite do telhado do Edifício 5 Executive Campus

Fonte: Google Maps (2021a).

Tabela D.3 – Características da Usina EUA

Identificador NREL	1201
Nome	Distributed Sun – 5 Executive Campus
Local	
País	Estados Unidos
Cidade	Cherry Hill/NJ
Fuso Horário	UTC-5
Coordenadas geográficas	
Latitude	39,9283°
Longitude	-75,0481°
Altitude	18 m
Sistema FV	
Potência Nominal	140,14 kWp
Fator de Capacidade ⁶	23,5%
Painéis solares	
Área	959,058 m ²
Azimute	230°
Inclinação	10°

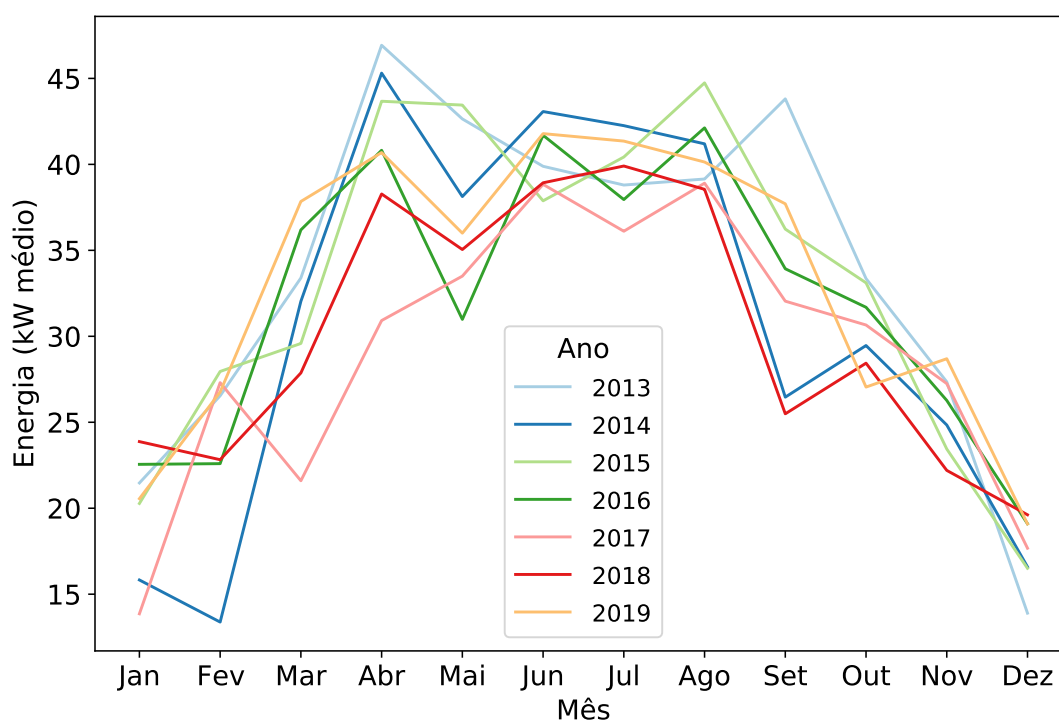
Fonte: elaboração própria (2021).

dados passaram a ser registrados com frequência de 5 minutos. Assim, a base de dados apresenta um indicador de disponibilidade incorreto que inicialmente assume o valor máximo de 1 e posteriormente cai a 0,6. Uniformizada a disponibilidade, as horas com menos de 80% dos registros esperados foram excluídas (cerca de 80 horas claras no período).

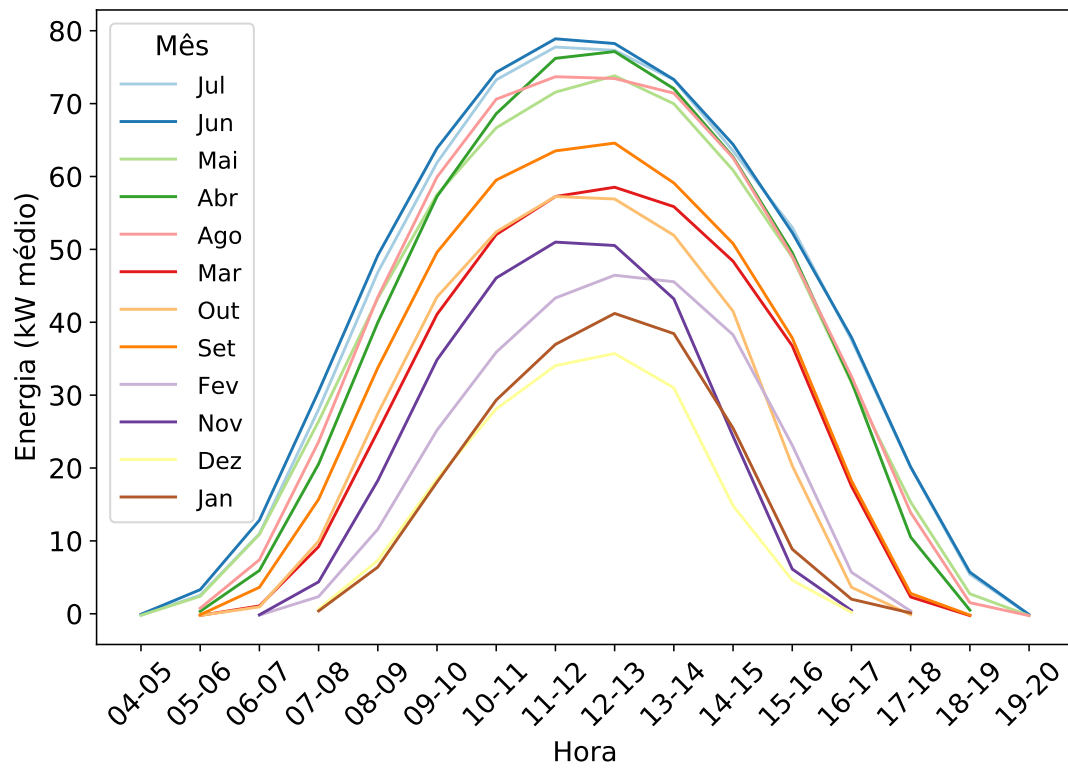
A Figura D.8 apresenta a distribuição mensal da geração de energia da

usina ano a ano. A variação pode ser significativa, como no caso do crescimento entre os anos de 2014 e 2015, no mês de fevereiro, da ordem de 109%. A distribuição horária da geração de energia elétrica da usina por mês do ano (considerando todos os anos) é apresentada pela Figura D.9. Além da variação das horas claras, pode-se notar como a área sob a curva – energia – é maior no mês de junho em relação a fevereiro, por exemplo. E como essa variação é maior que nas demais usinas (Figura D.3 e Figura D.6).

Figura D.8 – Distribuição mensal da geração de energia da usina EUA



Fonte: elaboração própria (2022).

Figura D.9 – Distribuição horária da geração de energia da usina EUA

Fonte: elaboração própria (2022).

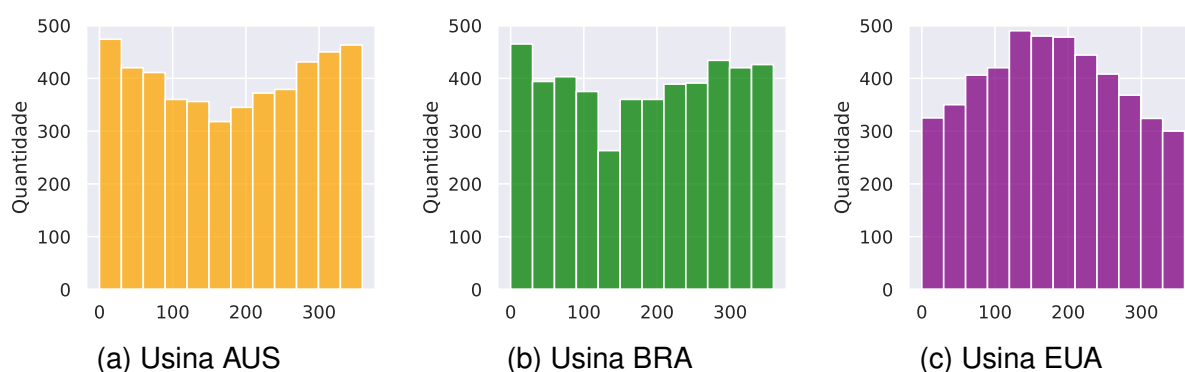
APÊNDICE E – HISTOGRAMAS DOS DADOS DE DESENVOLVIMENTO

Neste apêndice são apresentados os histogramas dos dados utilizados no desenvolvimento. Eles são segmentados por usina e compreendem todos os atributos de entrada e a geração de energia. Estão compreendidas as horas claras, ou seja, aquelas em que há alguma possibilidade de irradiação solar (do ponto de vista do geoposicionamento em relação ao Sol) ao longo de 12 meses, sendo:

- usina AUS: 2º semestre de 2010 ao 1º semestre de 2011;
- usina BRA: o ano cheio de 2019. Os dados de parcela significativa das horas do mês de maio estão faltando mas, acredita-se, não impedem sua utilização;
- usina EUA: o ano cheio de 2013.

A Figura E.1 apresenta o histograma do atributo Ângulo Diário Solar das três usinas. O início de sua escala ocorre às zero horas do primeiro dia do ano e compreende 0° a 360°. Como no verão os dias são mais longos que no inverno, principalmente em altas latitudes, a frequência não é uniforme ao longo dos doze meses (30° cada) plotados.

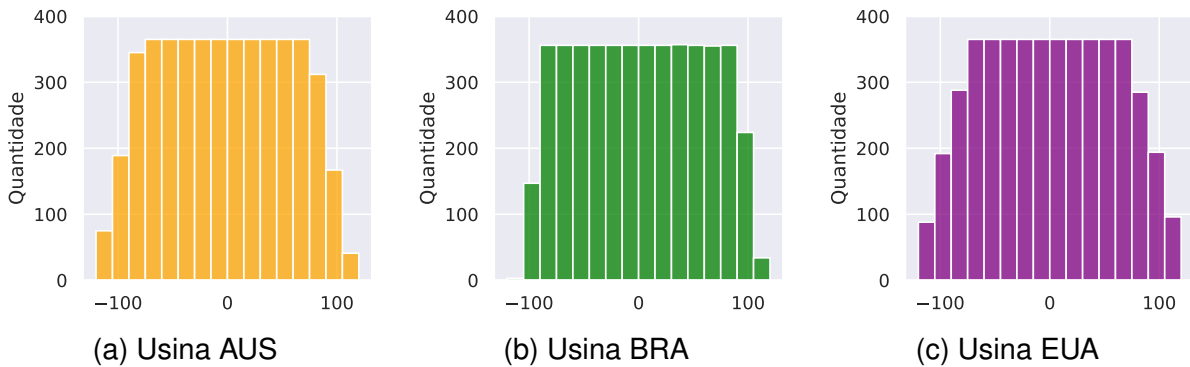
Figura E.1 – Histograma do atributo Ângulo Diário Solar (graus, 0° a 360°)



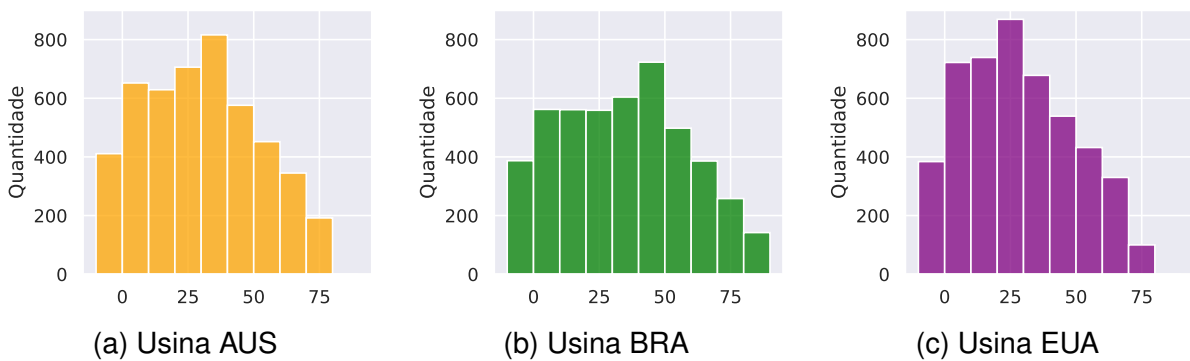
Fonte: elaboração própria (2022).

O Ângulo Horário Solar é apresentado pela Figura E.2. Neste caso, zero graus refere-se ao meio dia solar. Cada intervalo plotado corresponde a uma hora.

A Figura E.3 apresenta os dados do atributo Ângulo de Elevação Solar. O zero é o horizonte e quando o Sol está a pino o valor é 90°. Durante a primeira e a última hora do dia a média da elevação solar pode gerar um valor abaixo de zero, ainda que numa hora clara, ou seja, com alguma incidência solar (resultante da integração de toda a hora).

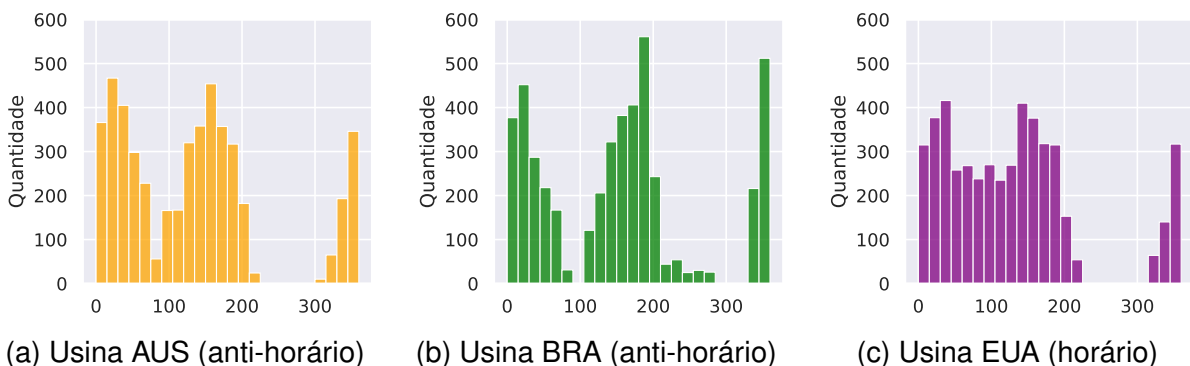
Figura E.2 – Histograma do atributo Ângulo Horário Solar (graus, -180° a 180°)

Fonte: elaboração própria (2022).

Figura E.3 – Histograma do atributo Ângulo de Elevação Solar (Horizonte= 0°)

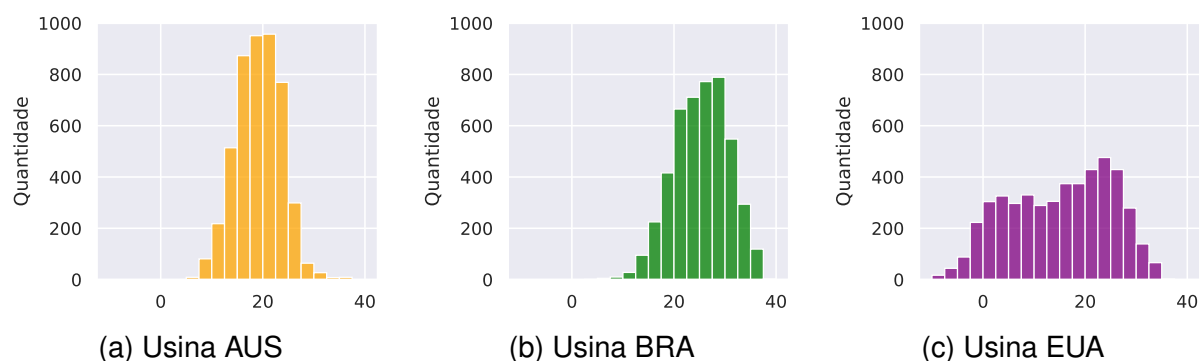
Fonte: elaboração própria (2022).

O Azimute (Figura E.4) está apresentado na forma ajustada, ou seja, o zero ao Leste e o crescimento (até 360°) ocorre no sentido anti-horário para as usinas AUS e BRA (hemisfério Sul) e no sentido horário para a usina EUA (hemisfério Norte).

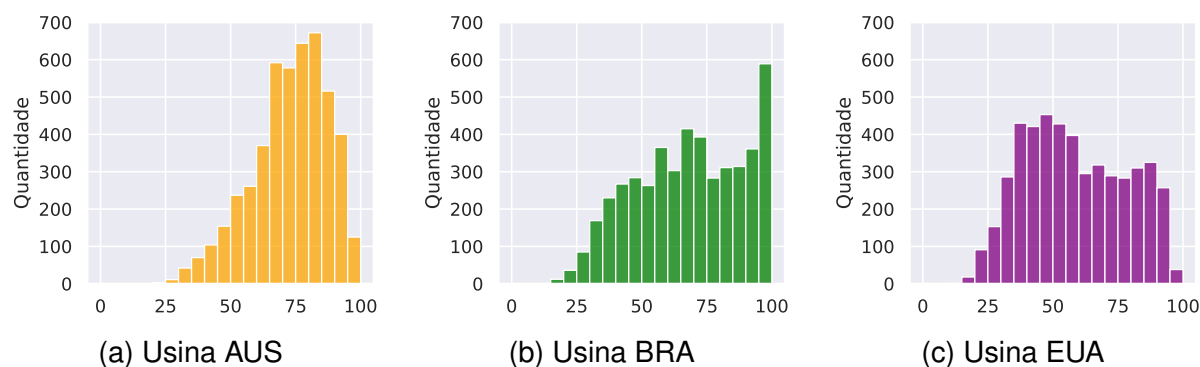
Figura E.4 – Histograma do atributo Azimute – ajustado (Leste= 0°)

Fonte: elaboração própria (2022).

As Figuras E.5, E.6 apresentam o histograma dos registros instantâneos de Temperatura Ambiente e Umidade Relativa do Ar, respectivamente. Apenas a usina EUA possui registros de temperatura abaixo de 0°C (zero graus Celsius) no período.

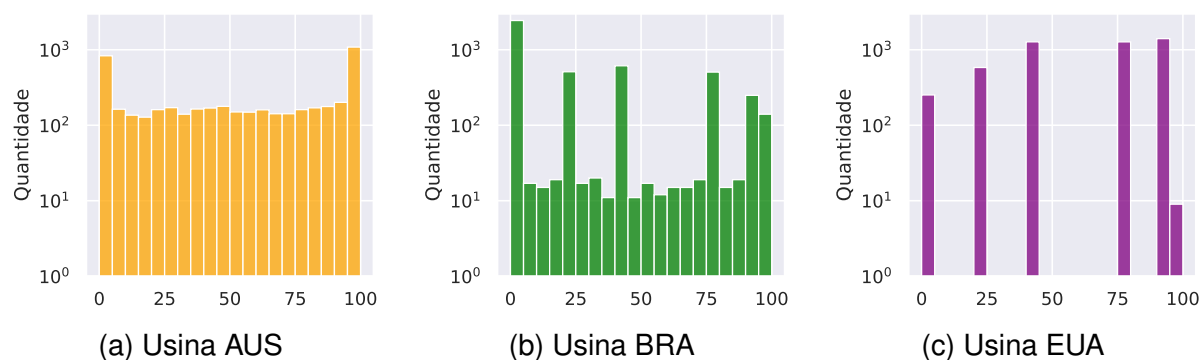
Figura E.5 – Histograma do atributo Temperatura Ambiente (°C)

Fonte: elaboração própria (2022).

Figura E.6 – Histograma do atributo Umidade Relativa do Ar (%)

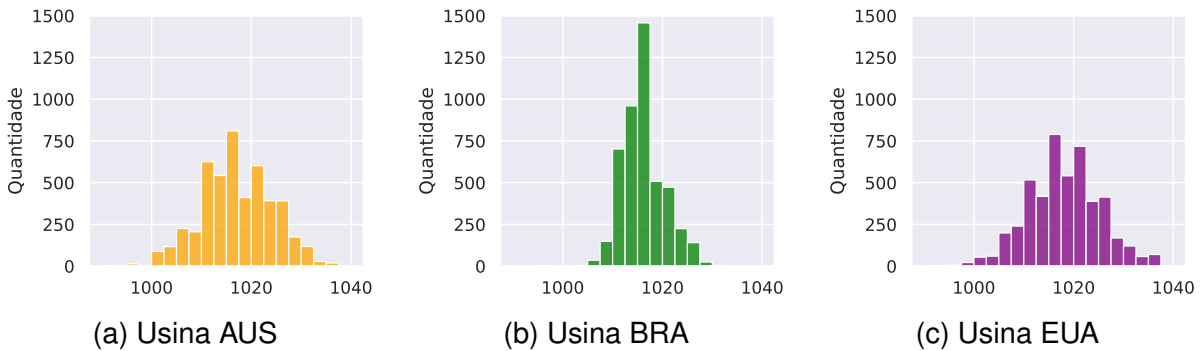
Fonte: elaboração própria (2022).

A Cobertura de Nuvens (Figura E.7) é um dado oriundo, geralmente, de verificações visuais (humanas), provável caso da usina EUA, mas podem ser geradas por sistemas automatizados, com maior discretização, provável caso da usina AUS. A usina BRA aparentemente contém ambos tipos de verificação.

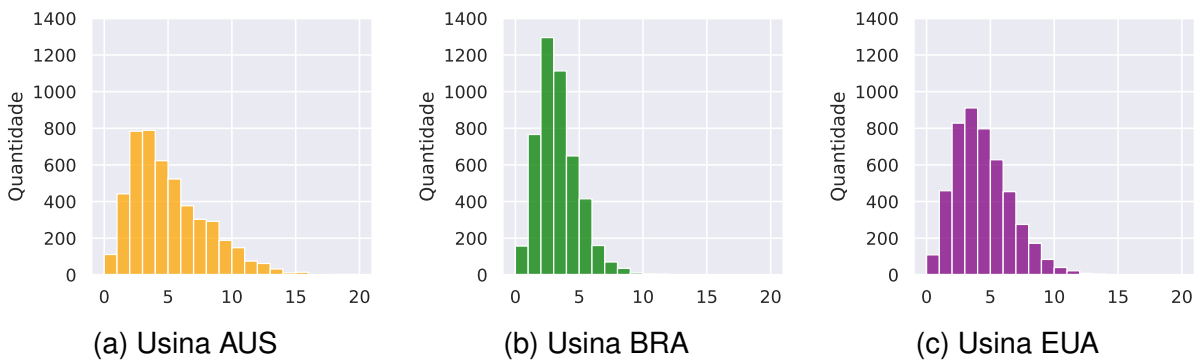
Figura E.7 – Histograma do atributo Cobertura de Nuvens (%)

Fonte: elaboração própria (2022).

A Figura E.8 apresenta o histograma da Pressão Atmosférica e a Figura E.9 a Velocidade dos Ventos.

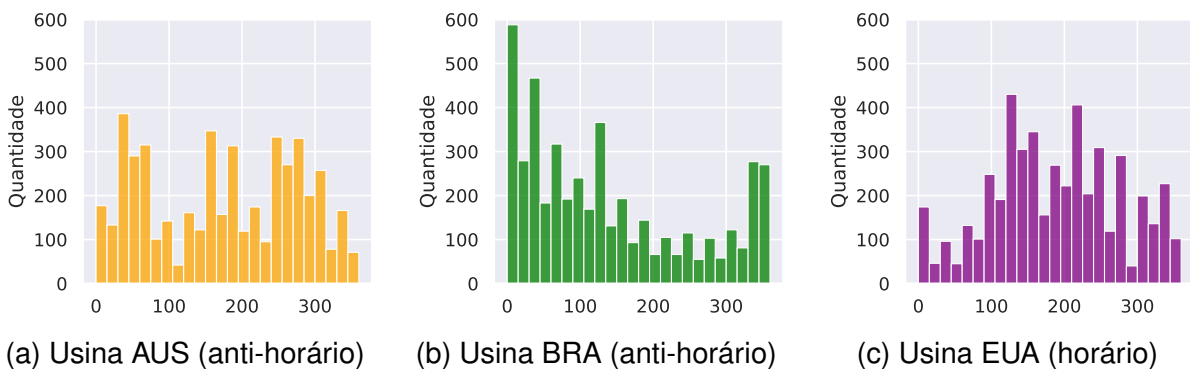
Figura E.8 – Histograma do atributo Pressão Atmosférica (hPa)

Fonte: elaboração própria (2022).

Figura E.9 – Histograma do atributo Velocidade dos Ventos (m/s)

Fonte: elaboração própria (2022).

Da mesma forma que o Azimute, a Direção dos Ventos (Figura E.10) é apresentada em sua forma ajustada, conforme a proposta deste estudo. Além das influências individuais, facilita-se o estabelecimento de relacionamentos entre as duas variáveis.

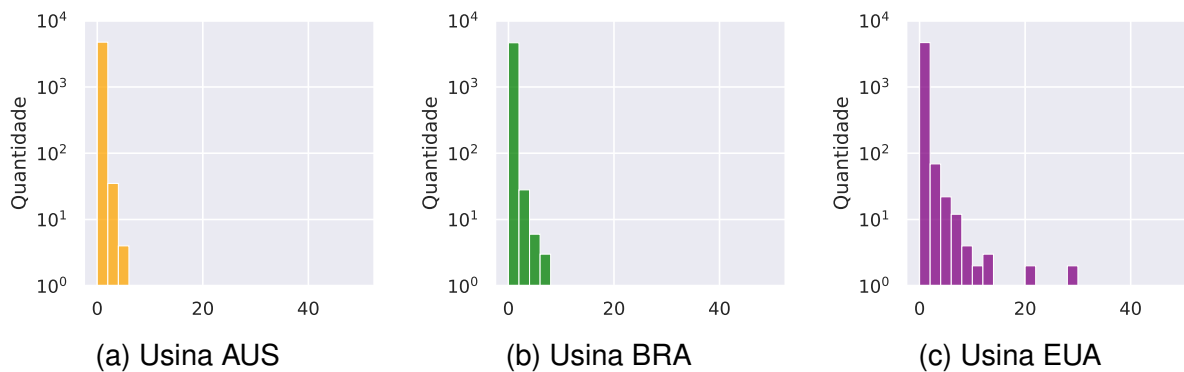
Figura E.10 – Histograma do atributo Direção dos Ventos – ajustado (Leste=0°)

Fonte: elaboração própria (2022).

O histograma do Índice de Chuva (e Neve), apresentado pela Figura E.11, corresponde à proposta deste estudo de junção desses atributos (chuva e neve) –

considerando-os mutuamente exclusivos – e com escalas similares em milímetros de água.

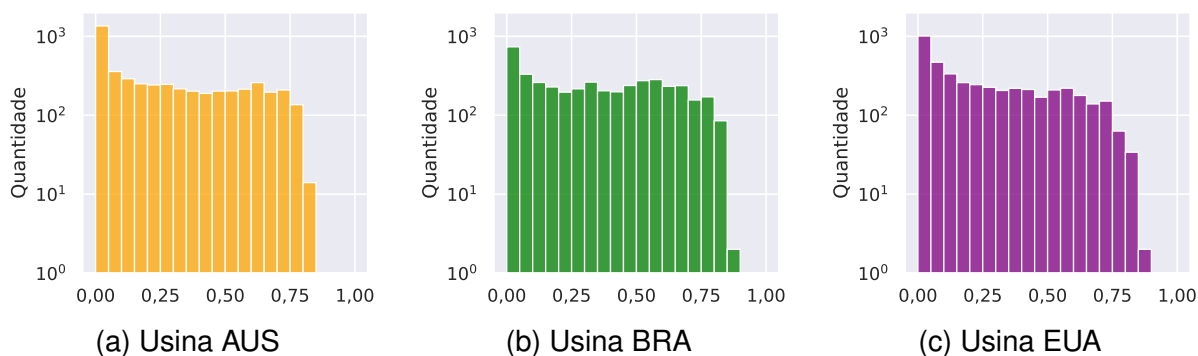
Figura E.11 – Histograma do atributo Índice de Chuva (e Neve) (mm, estado líquido)



Fonte: elaboração própria (2022).

Os dados de geração horária de energia apresentados nos histogramas da Figura E.12 estão no sistema por unidade (p.u.). A base é a capacidade máxima de cada usina: 2,04 kWp (AUS), 95 kWp (BRA) e 140,14 kWp (EUA).

Figura E.12 – Histograma da Geração de Energia (p.u.)



Fonte: elaboração própria (2022).

APÊNDICE F – FORMAÇÃO DAS PARCELAS DE VALIDAÇÃO CRUZADA

A validação cruzada foi a metodologia adotada para suportar o processo de escolha dos parâmetros e hiperparâmetros utilizados nos experimentos deste trabalho. Todas etapas de desenvolvimento utilizaram os mesmos critérios para todas usinas, como mostra a Tabela F.1.

Tabela F.1 – Condições gerais da validação cruzada

Aspecto	Opção
Método	
Tipo	k-fold, estratificada
Número de parcelas (k)	10 (dez)
Estratificação	Por ranqueamento e distribuição em zigue-zague
Ordenamento	1°) cobertura de nuvens do instante inicial (<i>clouds</i>): abaixo de 25%, intermediário e acima de 75% 2°) geração de energia crescente
Ferramenta	Algoritmo próprio

Fonte: elaboração própria (2021).

A seleção aleatória dos componentes de cada parcela foi preterida em favor do ranqueamento. O objetivo foi reduzir o risco de distribuição desproporcional dos exemplos relativos a horas de céu mais ou menos encoberto, bem como de horas do dia mais claras (próximas do meio-dia solar) ou mais escuras (amanhecer ou entardecer). A distribuição desses exemplos, antes da ordenação por geração de energia e ranqueamento, não é igual entre as parcelas e nem mantém a mesma proporcionalidade entre as usinas, como mostra a Tabela F.2.

Tabela F.2 – Distribuição da cobertura de nuvens das amostras de desenvolvimento

Cobertura de nuvens	AUS	BRA	EUA
< 25%	29,7%	64,1%	17,4%
entre 25% e 75%	33,5%	26,6%	53,1%
> 75%	36,8%	9,2%	29,5%

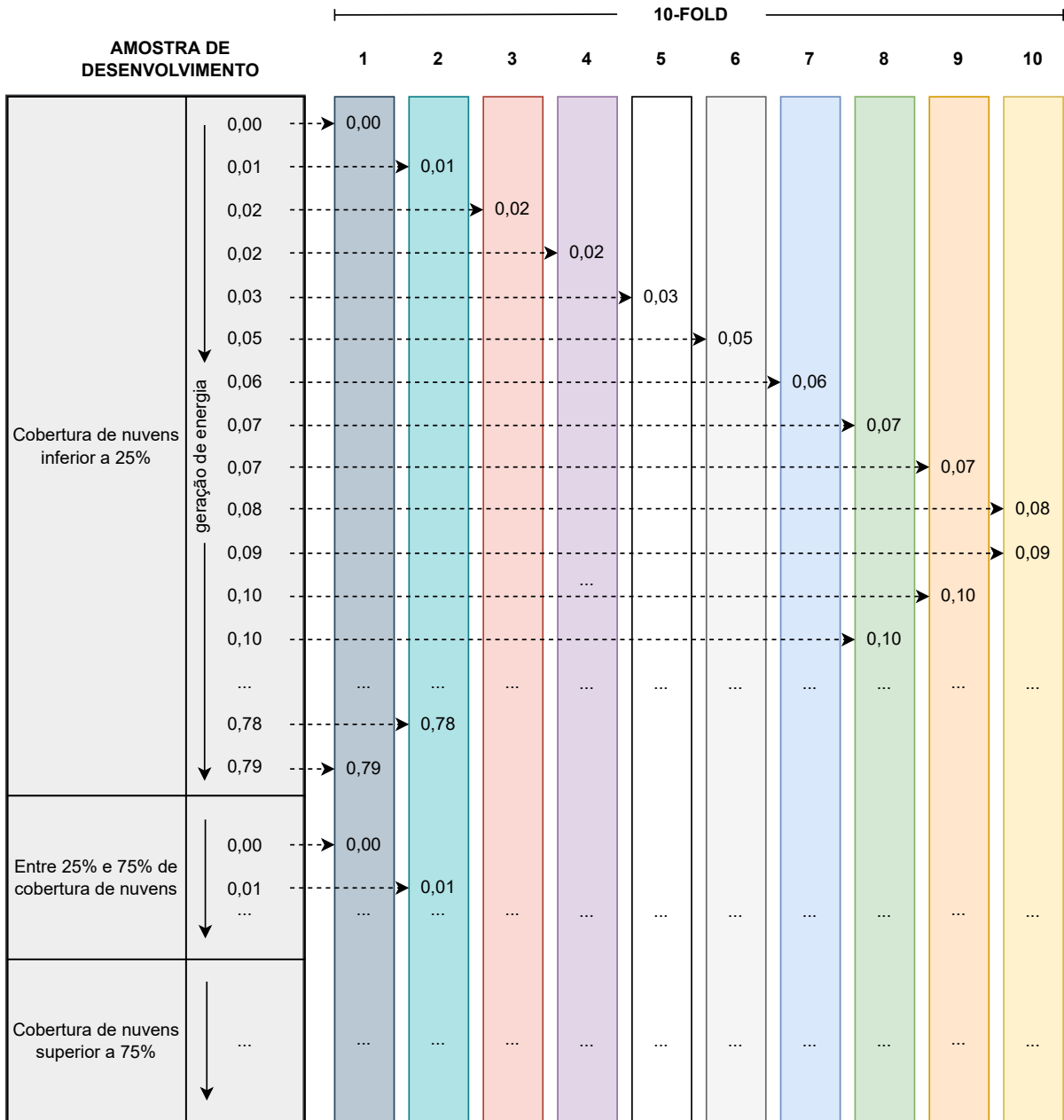
Fonte: elaboração própria (2022).

Uma outra vantagem é que a estratificação é sempre a mesma para uma mesma amostra, o que facilita a execução de testes comparativos envolvendo outros

aspectos do problema. Um último ponto é que as médias e desvios-padrão tendem a ser próximos entre todas as parcelas e a amostra completa.

A Figura F.1 demonstra o processo de seleção das k parcelas.

Figura F.1 – Ranqueamento para formação de 10 parcelas de validação cruzada



Fonte: elaboração própria (2022).

APÊNDICE G – EXPERIMENTO 1

Este experimento foi realizado de acordo com as condições gerais apresentadas pela Tabela G.1. O objetivo foi verificar o comportamento do treinamento das RNAs para as três usinas (AUS, BRA, EUA) diante de diferentes configurações e antes de um processo de regularização, através de validação cruzada, comparando as métricas: RMSE, MBE, MAE e R^2 , além do número de épocas necessárias até a solução.

Tabela G.1 – Condições gerais do Experimento 1

Aspecto	Valor(es) testado(s)
Ferramenta	Keras (TensorFlow) sem GPU
Estrutura MLP	
Camada de entrada	17 neurônios (11 atributos ^a)
Camada(s) oculta(s)	1 a 3 camadas de: 105, 36x36 e 27x27x27 neurônios (~ 2.000 parâmetros)
Camada de saída	1 neurônio
Ajustes	
Normalização de atributos	Minmax ou Z-score
Parâmetros iniciais	Glorot uniforme ou He normal, com vieses zerados
Função de ativação	ReLU, ELU
Função de custo	MSE
Otimizador	SGD c/ momento e nesterov, sem regularização
Lote	32 exemplos
N° de épocas	máximo de 10.000 c/ paciência ^b de 500
Atributos	<i>temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h, elevref, azim, hrang e dayang</i>
Validação	Cruzada de 10 parcelas c/ estratificação pela cobertura de nuvens do início do período (< 25%, intermediário e > 75%) e a geração de energia

^a Os atributos *temp, pressure, humidity, clouds, wind_speed* e *wind_deg* são pares de valores relativos ao instante inicial e final de cada hora.

^b Um algoritmo interrompe o treinamento na ausência de melhoria (*early stopping*), conforme os critérios detalhados no Apêndice F.

Fonte: elaboração própria (2022).

As configurações comparadas foram:

- a) dois tipos de normalização de entradas: Minmax e Z-score;

- b) duas funções de ativação: ReLU e ELU;
- c) duas estratégias de inicialização de parâmetros: He com distribuição normal e Glorot com distribuição uniforme. As distribuições testadas são aquelas dos artigos originais de He *et al.* (2015) e Glorot e Bengio (2010);
- d) três níveis de camadas ocultas: uma, duas ou três.

As amostras de desenvolvimento das três usinas (AUS, BRA e EUA) relativas a um ano de dados horários (horas claras) foram parceladas pela técnica de validação cruzada conforme o Apêndice F. Médias e desvios-padrão de geração de energia¹ (p.u.) das parcelas de treinamento e de validação ficaram muito próximos, como mostra a Tabela G.2, o que possibilita a comparação entre as métricas absolutas (MBE, MAE e RMSE) de amostras diferentes de uma mesma usina.

Tabela G.2 – Média e desvio-padrão da geração de energia por usina

	Média (p.u.)	DP (p.u.)
AUS	0,284 ± 0,001	0,254 ± 0,001
BRA	0,320 ± 0,002	0,260 ± 0,001
EUA	0,251 ± 0,003	0,240 ± 0,002

Fonte: elaboração própria (2022).

A Figura G.1 apresenta os melhores resultados de validação obtidos para as usinas, estruturas (uma, duas ou três camadas ocultas) e demais ajustes testados ao longo do treinamento. As estruturas não são identificadas porque não houve a pretensão de compará-las mas de elevar a representatividade dos resultados. Assim, para cada alternativa há três pontos plotados nos gráficos e cada ponto representa a média de 10 execuções (amostras de treinamento e validação distintas).

MAE e R^2 não apresentaram dispersões significativas (na forma do DP) em função da configuração da RNA, e as correlações² entre suas médias e a \overline{RMSE} entre as diferentes estruturas e configurações testadas foram boas, como mostra a Tabela G.3.

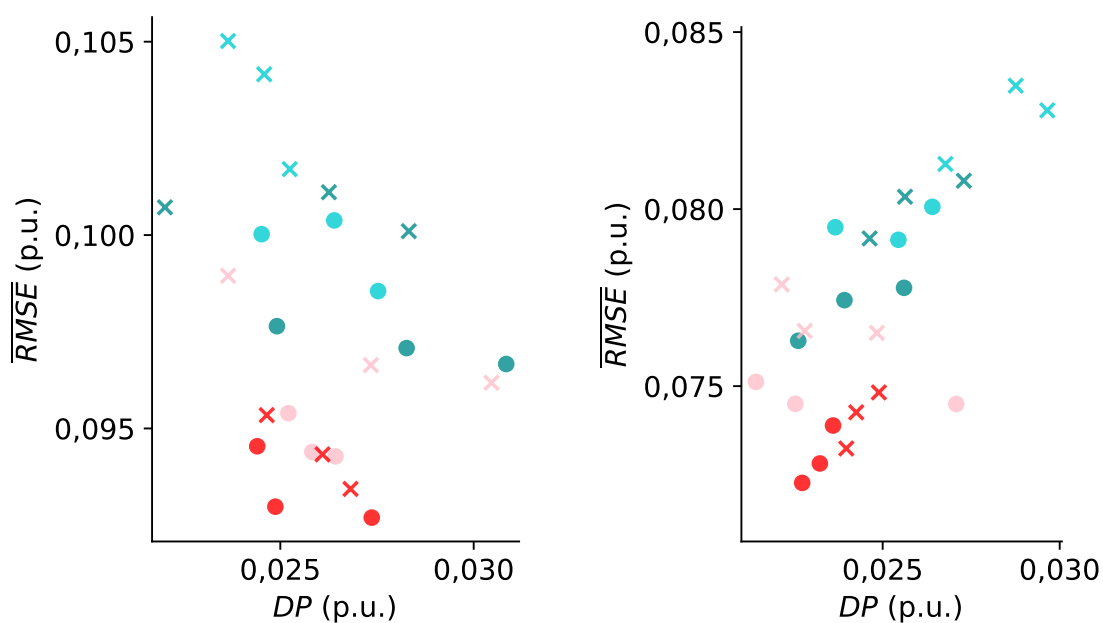
Essa correlação forte é importante porque sugere que os aspectos representados por essas métricas estão sendo otimizados concomitantemente pela rede neural, com o otimizador selecionado – o MSE. Já a \overline{MBE} ³ não está correlacionado com a \overline{RMSE} ou outra métrica, e nem é igual entre as usinas. Isso indica que o balanço

¹Os dados de geração horária de energia foram normalizados por unidade (p.u.) com base no pico de capacidade de cada usina: 2,04 kWp (AUS), 95 kWp (BRA) e 140,14 kWp (EUA).

²Correlação de Pearson (MURPHY, 2012).

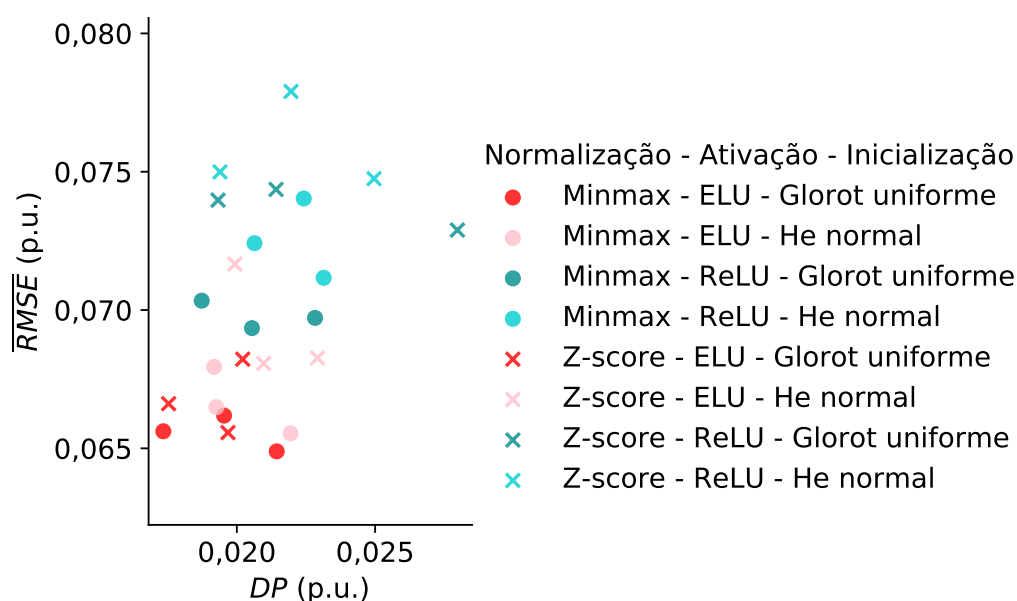
³Em última análise, o MBE é uma componente da RMSE, em função da decomposição viés-variância do MSE (JAMES *et al.*, 2013)

Figura G.1 – \overline{RMSE} e DP de validação cruzada por tipo de normalização, ativação e inicialização



(a) Usina AUS

(b) Usina BRA



(c) Usina EUA

Fonte: elaboração própria (2022).

entre viés e variância dos resultados é diferente entre as usinas e entre as estruturas ou configurações testadas.

O esforço em termos do número de épocas até o melhor resultado, ou seja, até o início da caracterização de sobreajuste (*overfitting*) ocorreu de forma similar entre

Tabela G.3 – Correlação com $\overline{\text{RMSE}}$ das demais métricas dos resultados

Métrica	AUS	BRA	EUA
$\overline{\text{MBE}}$	0.1672	0.3360	0.1140
$\overline{\text{MAE}}$	0.9494	0.9629	0.9417
$\overline{\text{R}^2}$	-0.9998	-0.9998	-0.9997

Fonte: elaboração própria (2022).

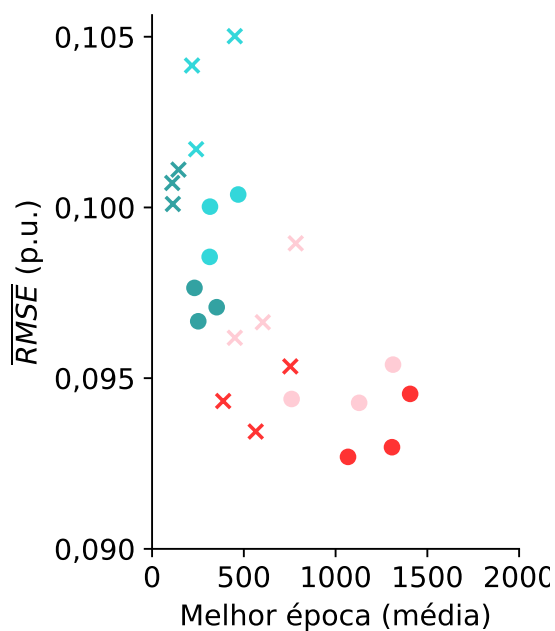
as usinas e diferente entre as configurações, como mostra o exemplo da Figura G.2.

As principais conclusões deste experimento são que:

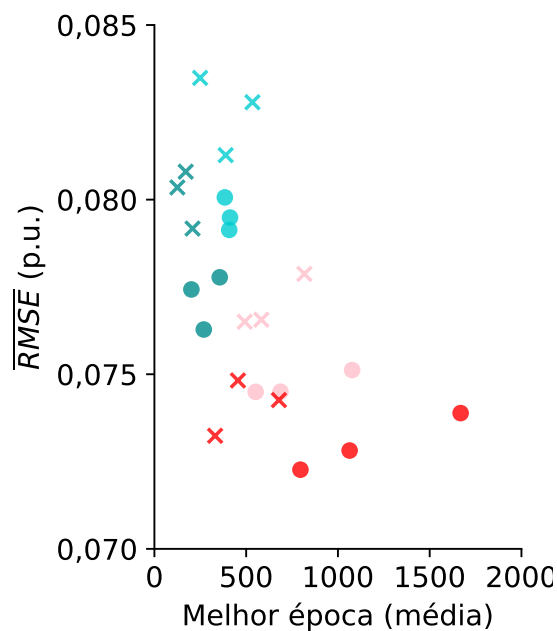
- a) as respostas ao treinamento são similares entre as usinas, para as diversas alternativas, ainda que os valores absolutos sejam diferentes;
- b) a normalização Minmax ("o") apresenta melhor generalização que a normalização Z-score ("x"), ou seja, $\overline{\text{RMSE}}$ mais reduzida, porém, toma um número de épocas maior;
- c) a função de ativação ELU (itens avermelhados) também apresenta médias mais baixas que a função de ativação ReLU (itens esverdeados), porém, toma um número de épocas maior;
- d) as execuções com inicialização do tipo Glorot e distribuição uniforme (cores fortes) apresentam resultados um pouco melhores que suas equivalentes com inicialização do tipo He e distribuição normal (cores claras).

Assim, as combinações que apresentaram a melhor generalização e a maior velocidade de treinamento são apresentadas pela Tabela G.4. Embora essa conclusão seja baseada em uma única execução da validação cruzada estratificada de 10 parcelas, caso a caso, a coincidência de valores relativos sob mesmas configurações entre as usinas, as quais representam distintas amostras (sistemas FVs, condições do tempo, geoposicionamento solar etc.) parece dotá-la de certa confiabilidade como representativa dos sistemas FVs – em geral – sob as premissas deste trabalho.

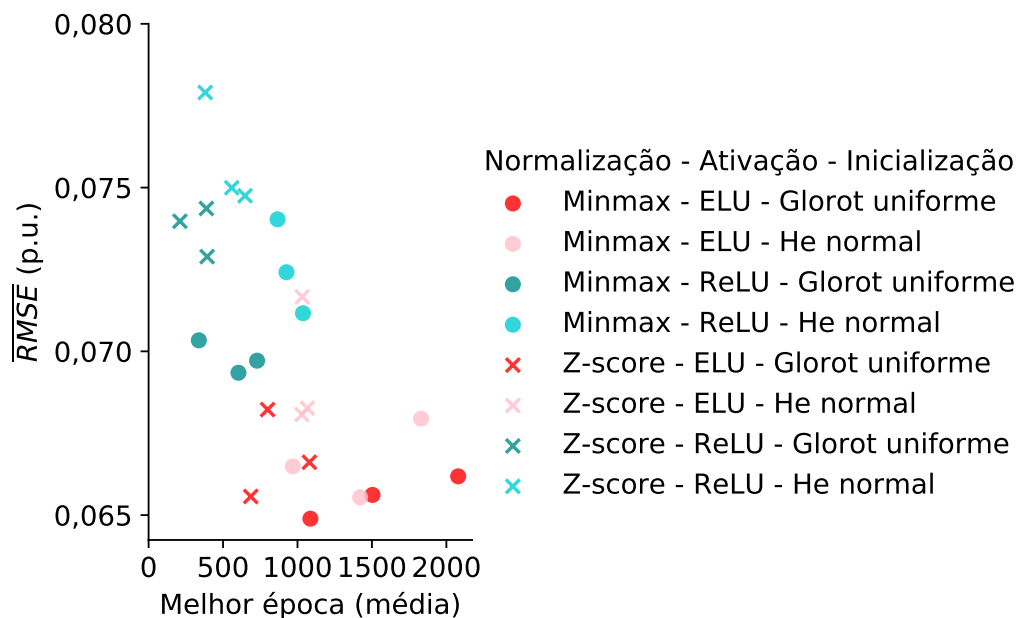
Figura G.2 – \overline{RMSE} e DP de validação cruzada por tipo de normalização, ativação e inicialização



(a) Usina AUS



(b) Usina BRA



(c) Usina EUA

Fonte: elaboração própria (2022).

Tabela G.4 – Combinações de interesse

Aspecto	↑ Generalização ^a	↑ Velocidade de treinamento ^b
Normalização	Minmax	Z-score
Ativação	ELU	ReLU
Inicialização	Glorot uniforme	Glorot uniforme

^a Em termos do menor \overline{RMSE} .

^b Em termos do menor número de épocas até o melhor resultado (antes do sobreajuste).

Fonte: elaboração própria (2022).

APÊNDICE H – EXPERIMENTO 2

Este experimento foi realizado de acordo com as condições gerais apresentadas pela Tabela H.1. O objetivo foi verificar os aspectos relacionados ao porte da rede neural para as três usinas (AUS, BRA, EUA), verificados em termos do número de camadas ocultas e de parâmetros, antes de um processo de regularização, através de validação cruzada, comparando a distribuição da RMSE a partir das diversas execuções.

Tabela H.1 – Condições gerais do Experimento 2

Aspecto	Valor(es) testado(s)
Ferramenta	Keras (TensorFlow) sem GPU (exceto a referência)
Estrutura MLP	
Camada de entrada	17 neurônios (11 atributos ^a)
Camada(s) oculta(s)	1, 2, 3 e 4 camadas 6 camadas (referência)
Camada de saída	1 neurônio
Ajustes	
Normalização de atributos	Minmax e Z-score
Parâmetros iniciais	Glorot uniforme, com vieses zerados
Função de ativação	ELU
Função de custo	erro quadrático médio (MSE)
Otimizador	SGD c/ momento e nesterov, sem regularização
Lote	32 exemplos
N° de épocas	máximo de 10.000 c/ paciência ^b de 200
Atributos	<i>temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h, elevref, azim, hrang e dayang</i>
Validação	Cruzada de 10 parcelas c/ estratificação pela cobertura de nuvens do início do período (< 25%, intermediário e > 75%) e a geração de energia

^a Os atributos *temp, pressure, humidity, clouds, wind_speed* e *wind_deg* são pares de valores relativos ao instante inicial e final de cada hora.

^b Um algoritmo interrompe o treinamento na ausência de melhoria (*early stopping*), conforme os critérios detalhados no Apêndice F.

Fonte: elaboração própria (2022).

Uma estrutura arbitrariamente selecionada para representar um porte grande o suficiente, de 6 camadas e 200.000 parâmetros (ou cerca de 40 vezes o número de

exemplos anual de treinamento das usinas) foi utilizada como referência para verificação dos resultados de:

- estruturas de diferentes números de parâmetros: 2 mil, 10 mil e 20 mil (valores aproximados);
- com mesmo número de neurônios em cada camada oculta; e
- para dois tipos de normalização das entradas: Minmax e Z-score.

As amostras de desenvolvimento das três usinas (AUS, BRA e EUA) relativas a um ano de dados horários (horas claras), foram parceladas pela técnica de validação cruzada de 10 parcelas, estratificadas segundo a cobertura de nuvens no início do período (*clouds*) e a geração de energia, conforme o Apêndice F. Médias e desvios-padrão de geração de energia¹ (p.u.) das parcelas de treinamento e de validação ficaram muito próximos, como mostra a Tabela H.2, o que possibilita a comparação entre as métricas absolutas (RMSE) de amostras diferentes de uma mesma usina. O R^2 é utilizado para comparação dos resultados relativos entre as usinas.

Tabela H.2 – Média e desvio-padrão da geração de energia por usina

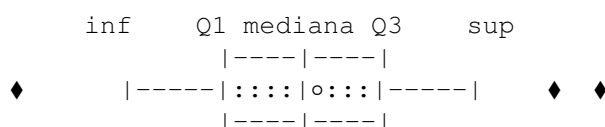
	Média (p.u.)	DP (p.u.)
AUS	0,284 ± 0,001	0,254 ± 0,001
BRA	0,320 ± 0,002	0,260 ± 0,001
EUA	0,251 ± 0,003	0,240 ± 0,002

Fonte: elaboração própria (2022).

As Figuras H.1 e H.2 apresentam os diagramas de caixa² (*boxplot*) dos resultados individuais (RMSE e R^2 , respectivamente) de validação cruzada obtidos para as três usinas considerando as estruturas de até quatro camadas ocultas (com igual número de neurônios em cada camada) e diferentes números totais de parâmetros: pesos e vieses. O processo de validação cruzada foi repetido três vezes, portanto,

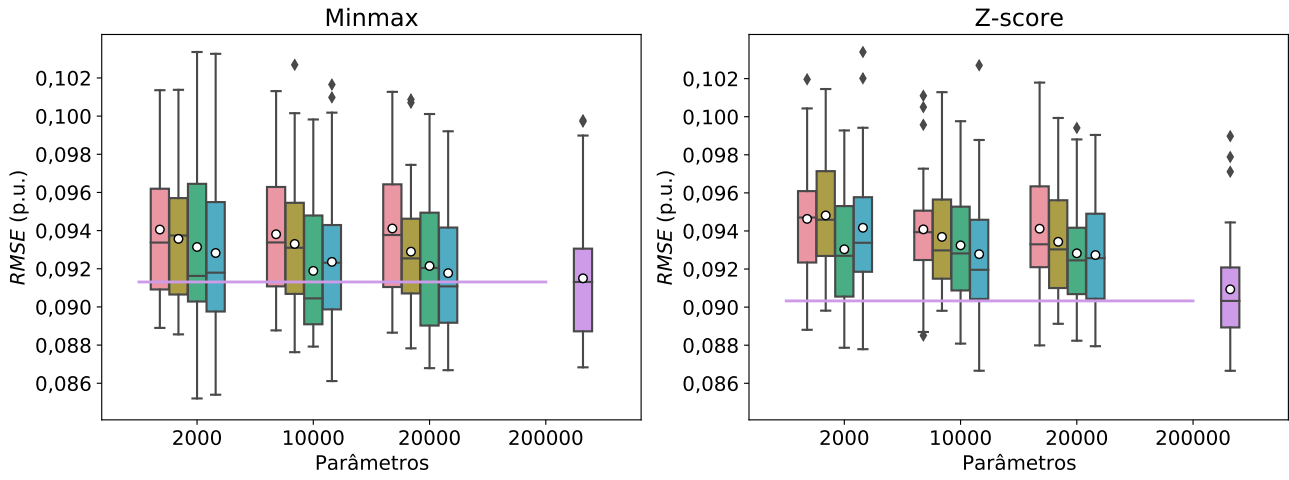
¹Os dados de geração horária de energia foram normalizados por unidade (p.u.) com base no pico de capacidade de cada usina: 2,04 kWp (AUS), 95 kWp (BRA) e 140,14 kWp (EUA).

²Diagrama de caixa é uma forma de representação de uma distribuição em termos da mediana (marca interna) e dos limites inferior (1º quartil) e superior (3º quartil) da área hachurada, limites extremos inferior (inf) e superior (sup) conectados por linhas, além de – eventualmente – pontos considerados fora da curva (♦). Na ausência desses pontos, mais distanciados do Q1 ou Q3 do que uma vez e meia a distância interquartil (Q3-Q1), os limites extremos coincidem com o mínimo (inf) e máximo (sup) (DEKKING *et al.*, 2005). Adicionalmente, pode ser exibida a média (○).

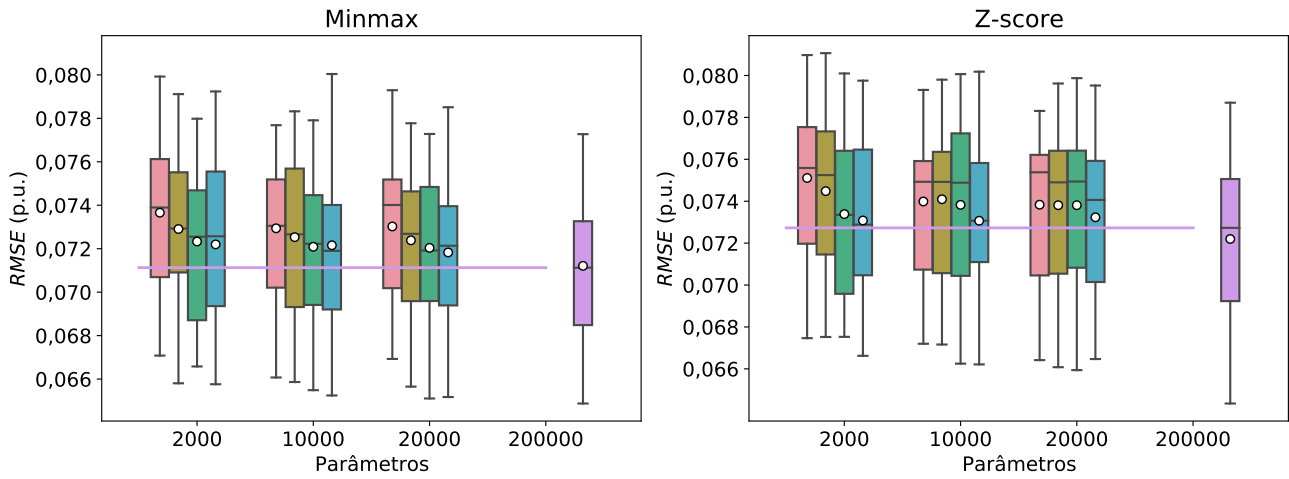


cada item no diagrama representa 30 execuções: três para cada parcela de validação cruzada.

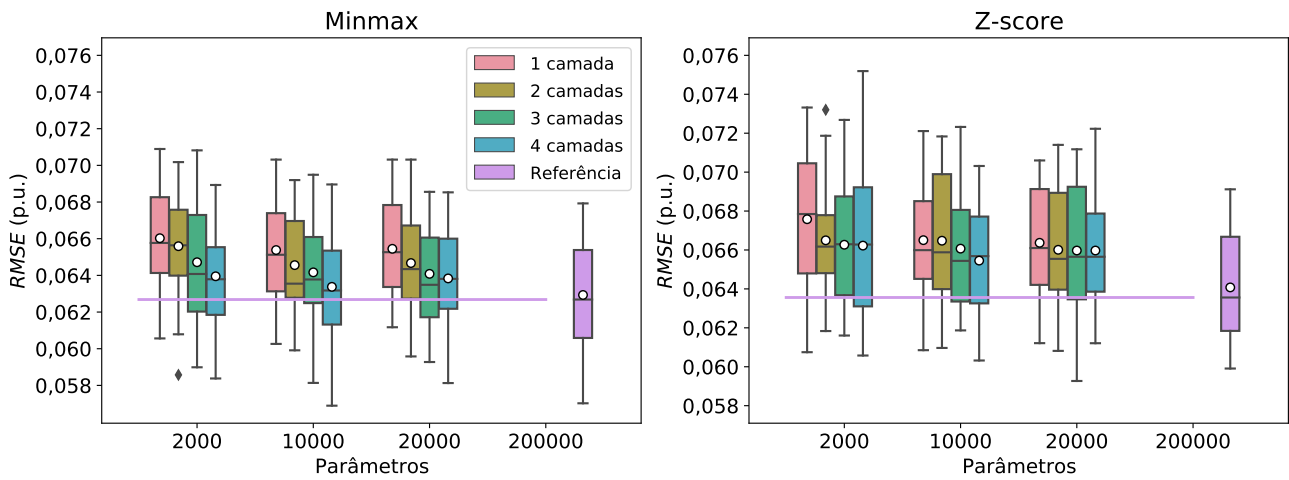
Figura H.1 – RMSE da validação cruzada por nº de camadas e nº de parâmetros



(a) Usina AUS



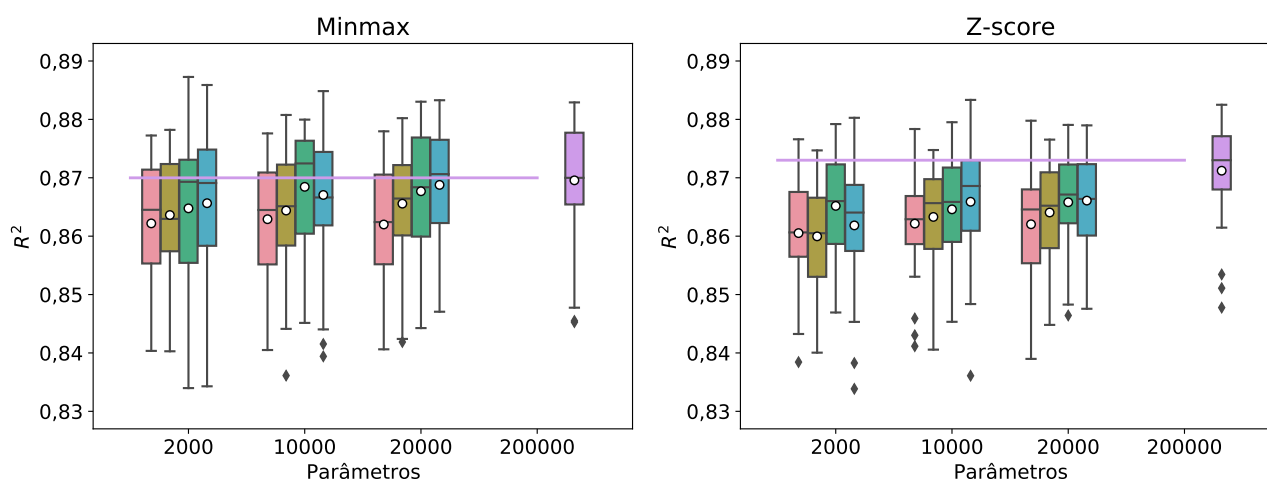
(b) Usina BRA



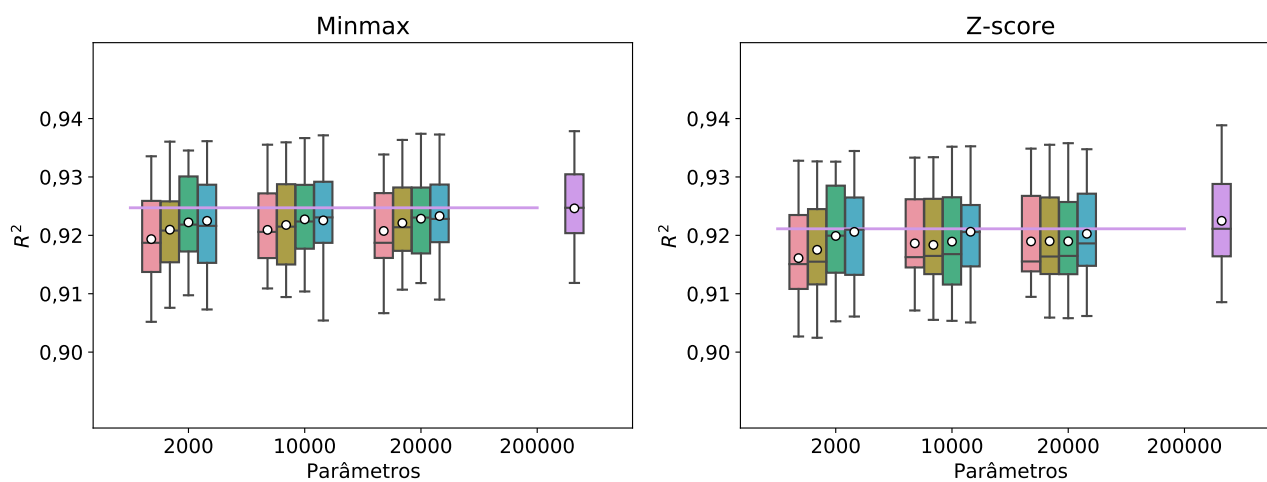
(c) Usina EUA

Fonte: elaboração própria (2022).

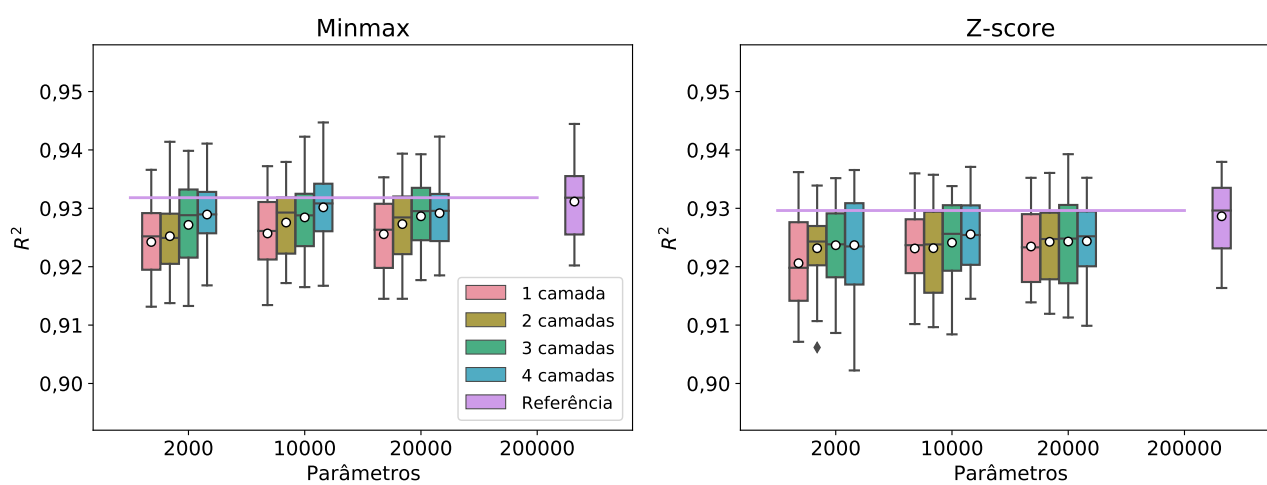
Figura H.2 – R^2 da validação cruzada por nº de camadas e nº de parâmetros



(a) Usina AUS



(b) Usina BRA



(c) Usina EUA

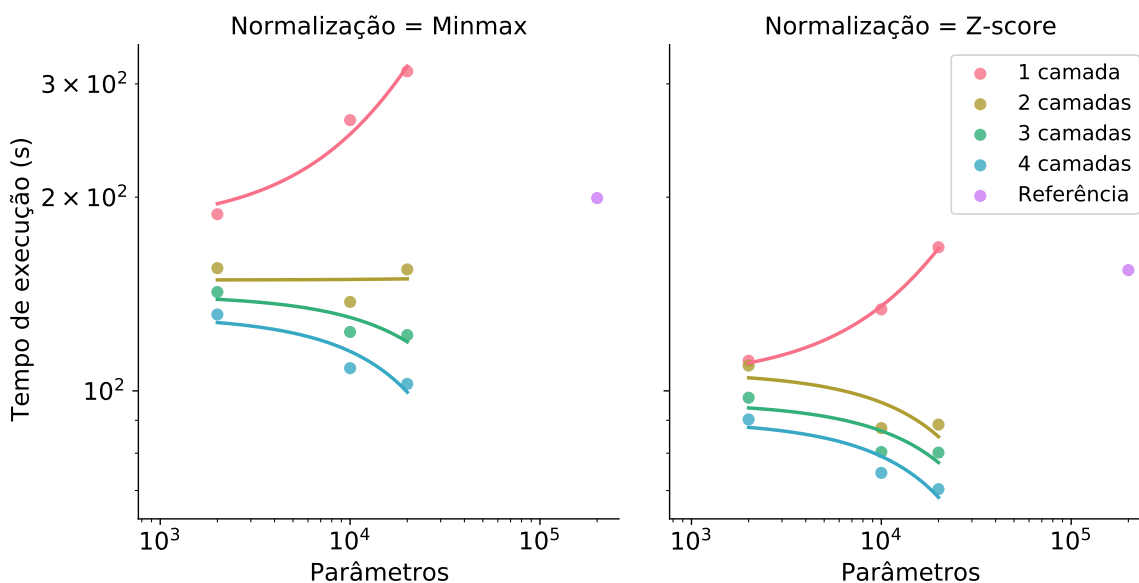
Fonte: elaboração própria (2022).

Os resultados – sem regularização – não mostraram um padrão determinante sobre a forma ideal de uma RNA para o propósito desejado. No entanto, é possível extrair algumas conclusões:

- um número grande de parâmetros (pesos e vieses) – algumas vezes maior que o número (quase 5000) de exemplos de treinamento – não resultou em aumento ou redução notável da dispersão dos resultados;
- as médias parecem indicar que o aumento do nº de camadas ocultas é mais interessante que o aumento do nº de parâmetros da rede neural;
- os melhores resultados não se distanciam significativamente daqueles obtidos pela rede neural de referência (de 6 camadas ocultas e 200.000 parâmetros), mas a distribuição dos resultados desta última parecem indicar que o aumento do porte da rede neural é benéfico, e não prejudicial, para a qualidade do resultado.

Finalmente, a Figura H.3 apresenta, de forma acessória, o tempo médio de execução do treinamento (todas usinas) com curvas de interpolação polinomial do 2º grau para legibilidade. Ressalta-se que as redes até quatro camadas ocultas foram resolvidas sem o uso da GPU (a CPU foi mais eficiente) ao passo que a RNA de referência foi resolvida com o uso da GPU. É possível perceber que as RNAs de uma camada são mais lentas, assim como a rede de referência, mas os motivos são diferentes: a primeira levou um número maior de épocas até atingir a melhor solução (início do sobreajuste), enquanto a segunda levou mais tempo por época.

Figura H.3 – Tempo de execução do treinamento em função do número de camadas ocultas e o número de parâmetros da rede neural (todas usinas)



APÊNDICE I – EXPERIMENTO 3

Este experimento foi realizado de acordo com as condições gerais apresentadas pela Tabela I.1. O objetivo foi verificar se o processo de regularização pode interferir na capacidade de generalização das RNAs estudadas. Para isto, foram testadas duas opções de regularização: Lasso e Ridge.

Tabela I.1 – Condições gerais do Experimento 3

Aspecto	Valor(es) testado(s)
Ferramenta	Keras (TensorFlow) sem GPU
Estrutura MLP	
Camada de entrada	17 neurônios (11 atributos ^a)
Camada(s) oculta(s)	4 camadas (20.000 parâmetros)
Camada de saída	1 neurônio
Ajustes	
Normalização de atributos	Minmax e Z-score
Parâmetros iniciais	Glorot uniforme, com vieses zerados
Função de ativação	ELU e ReLU
Função de custo	erro quadrático médio (MSE)
Função de regularização	Lasso (norma ℓ_1) e Ridge (norma ℓ_2)
Otimizador	SGD c/ momento e nesterov, sem regularização
Lote	32 exemplos
Nº de épocas	máximo de 10.000 c/ paciência ^b de 500
Atributos	<i>temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h, elevref, azim, hrang e dayang</i>
Validação	Cruzada de 10 parcelas c/ estratificação pela cobertura de nuvens do início do período (< 25%, intermediário e > 75%) e a geração de energia

^a Os atributos *temp, pressure, humidity, clouds, wind_speed* e *wind_deg* são pares de valores relativos ao instante inicial e final de cada hora.

^b Um algoritmo interrompe o treinamento na ausência de melhoria (*early stopping*), conforme os critérios detalhados no Apêndice F.

Fonte: elaboração própria (2022).

Uma estrutura de quatro camadas ocultas, 78 neurônios por camada e cerca de 20 mil parâmetros (ou cinco vezes o número de exemplos de treinamento) foi selecionada. Foram verificados, em termos do R^2 , as combinações de:

- a) dois tipos de normalização de entradas: Minmax e Z-score;

- b) duas funções de ativação: ELU e ReLU, ambas com distribuição do tipo Glorot uniforme;
- c) diferentes ajustes dos coeficientes ℓ_1 e ℓ_2 , de regularização.

As amostras de desenvolvimento das três usinas (AUS, BRA e EUA), relativas a um ano de dados horários (horas claras), foram parceladas pela técnica de validação cruzada de 10 parcelas, estratificadas segundo a cobertura de nuvens no início do período (*clouds*) e a geração de energia, conforme o Apêndice F. Médias e desvios-padrão de geração de energia¹ (p.u.) das parcelas de treinamento e de validação ficaram muito próximos, como mostra a Tabela I.2.

Tabela I.2 – Média e desvio-padrão da geração de energia por usina

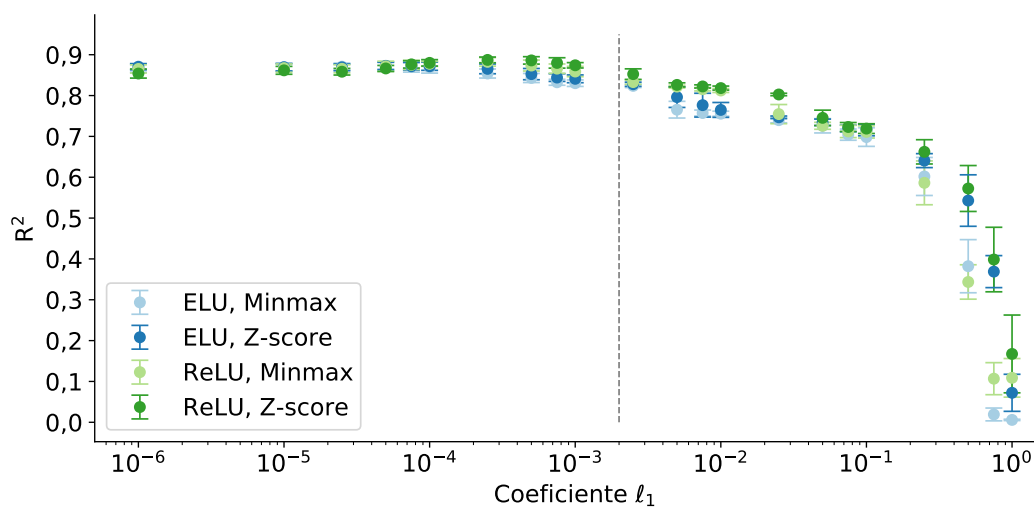
	Média (p.u.)	DP (p.u.)
AUS	0,284 ± 0,001	0,254 ± 0,001
BRA	0,320 ± 0,002	0,260 ± 0,001
EUA	0,251 ± 0,003	0,240 ± 0,002

Fonte: elaboração própria (2022).

A Figura I.1 ilustra a distribuição dos resultados da Usina AUS com aplicação de regularização Lasso, ao longo de uma faixa de coeficientes ℓ_1 que vai de 10^{-6} a 10^0 (um), e para as quatro alternativas testadas. São apresentadas médias ± 1 (um) desvio padrão. Os resultados de todas as usinas foram semelhantes no que concerne ao propósito de mostrar que a faixa de interesse está abaixo de 2×10^{-3} , linha vertical tracejada na figura.

¹Os dados de geração horária de energia foram normalizados por unidade (p.u.) com base no pico de capacidade de cada usina: 2,04 kWp (AUS), 95 kWp (BRA) e 140,14 kWp (EUA).

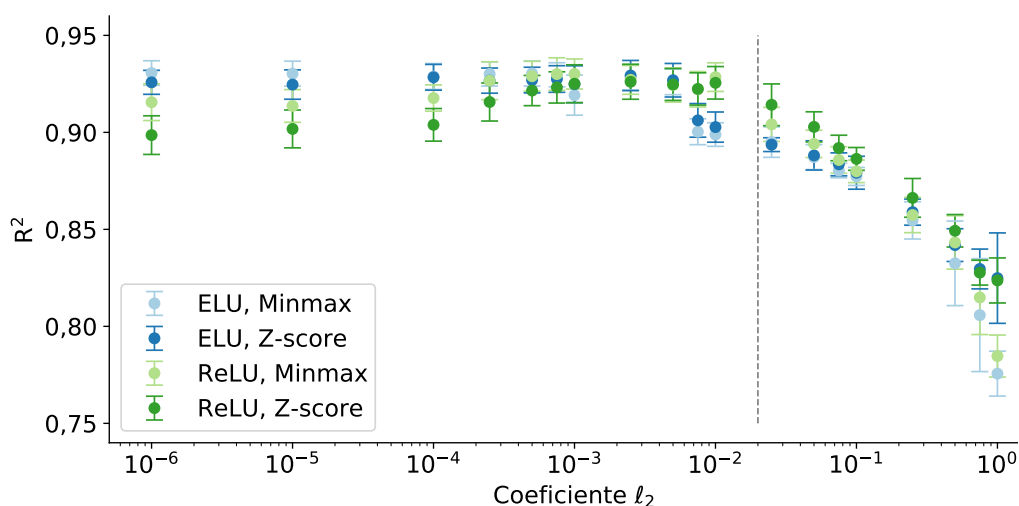
Figura I.1 – Impacto da regularização do tipo Lasso (ℓ_1) – Usina AUS



Fonte: elaboração própria (2022).

Analogamente, a Figura I.2 ilustra a forma de distribuição dos resultados da usina EUA com regularização Ridge, ao longo de uma faixa de coeficientes ℓ_2 que vai de 10^{-6} a 10^0 (um), e para as quatro alternativas testadas. Os resultados de todas as usinas também foram semelhantes, e mostram que a faixa de interesse está abaixo de 2×10^{-2} , linha vertical tracejada na figura.

Figura I.2 – Impacto da regularização do tipo Ridge (ℓ_2) – Usina EUA



Fonte: elaboração própria (2022).

As Figuras I.3 e I.4 apresentam diagramas de caixa² (*boxplot*) para os tipos de regularização ℓ_1 e ℓ_2 , respectivamente, restritas às faixas de interesse mencionadas, e agora com a adição dos valores antes da regularização (coeficientes zerados). Nota-se que os resultados relativos das Usinas BRA e EUA são similares, e diferem dos da Usina AUS. A linha horizontal tracejada é a mediana dos resultados para a normalização Minmax e função de ativação ELU, ou seja, a configuração que apresentou o melhor resultado antes da regularização para todas as Usinas.

O benefício foi mais significativo para a usina AUS do que para as demais, como mostra a Tabela I.3, que compara a melhor combinação sem regularização (normalização Minmax e ativação ELU) com a mais rápida (normalização Z-score e ativação ReLU) que teve bom resultado geral após a regularização.

²Diagrama de caixa é uma forma de representação de uma distribuição em termos da mediana (marca interna) e dos limites inferior (1º quartil) e superior (3º quartil) da área hachurada, limites extremos inferior (inf) e superior (sup) conectados por linhas, além de – eventualmente – pontos considerados fora da curva (♦). Na ausência desses pontos, mais distanciados do Q1 ou Q3 do que uma vez e meia a distância interquartil (Q3-Q1), os limites extremos coincidem com o mínimo (inf) e máximo (sup) (DEKKING *et al.*, 2005). Adicionalmente, pode ser exibida a média (◊).

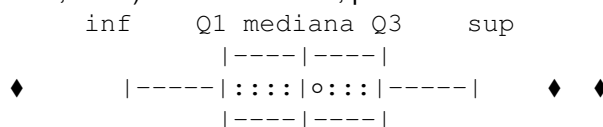
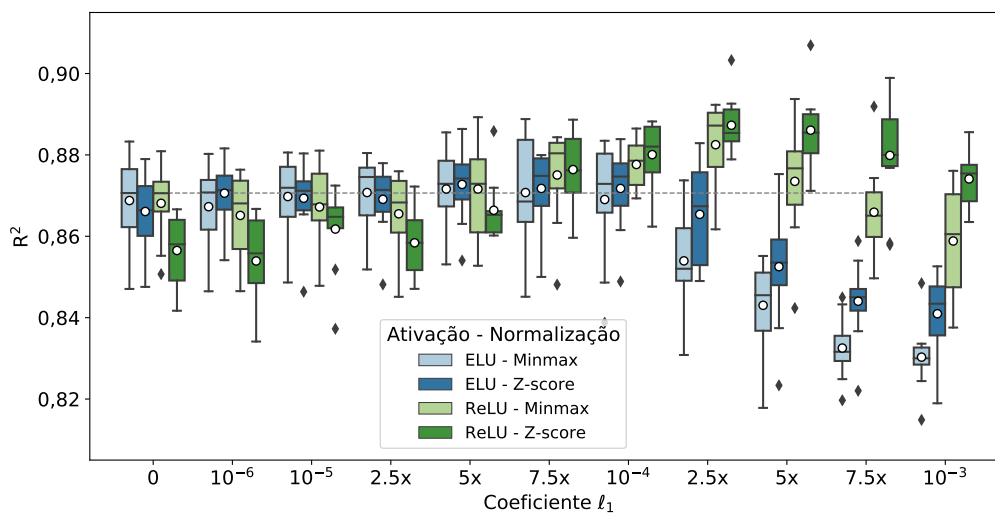
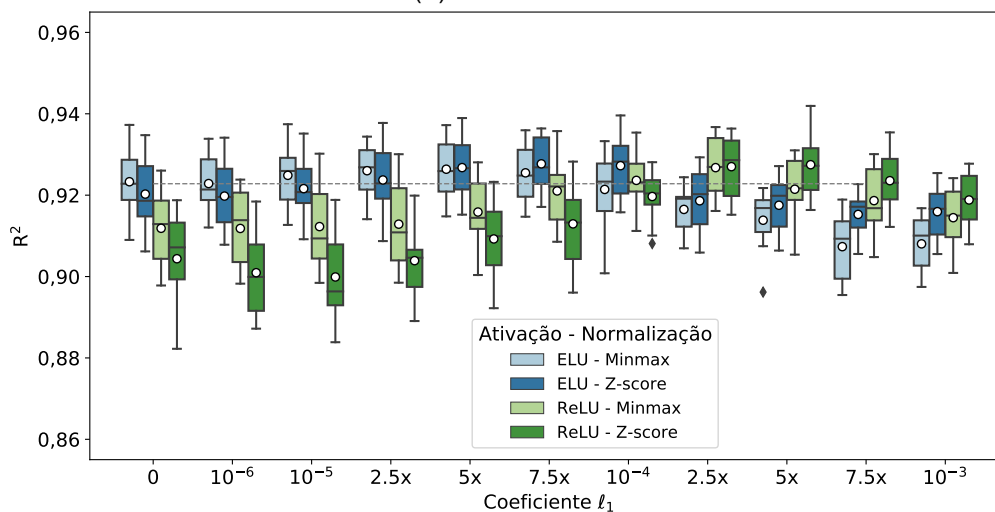


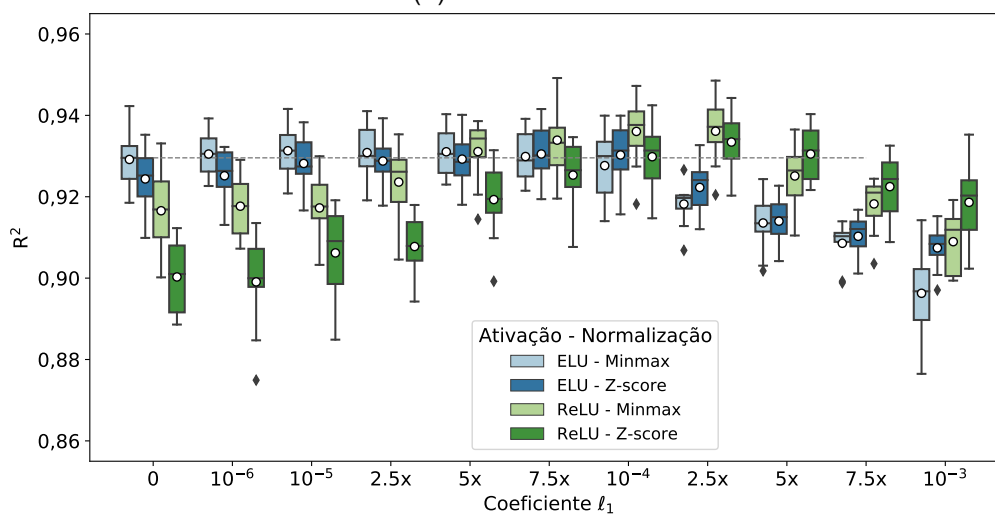
Figura I.3 – R^2 das usinas para o treinamento com regularização do tipo Lasso (ℓ_1)



(a) Usina AUS



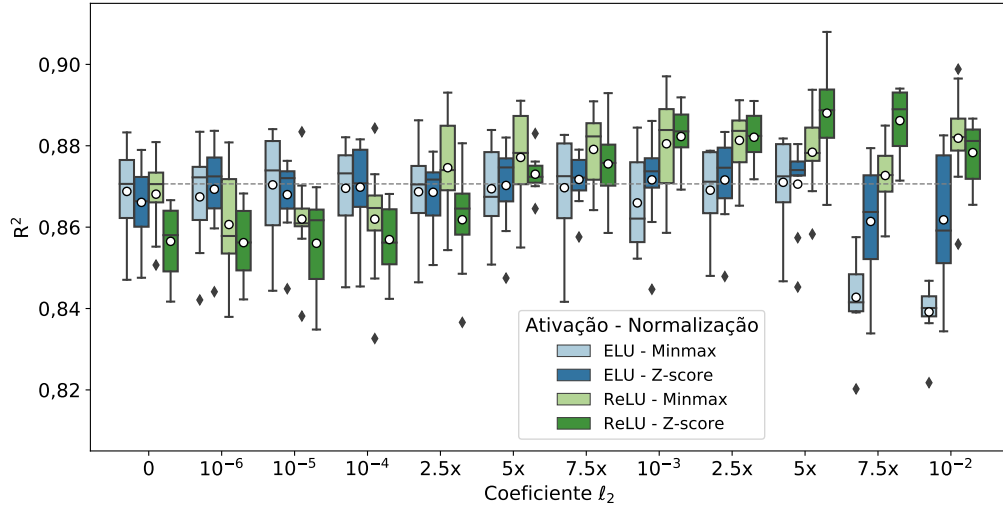
(b) Usina BRA



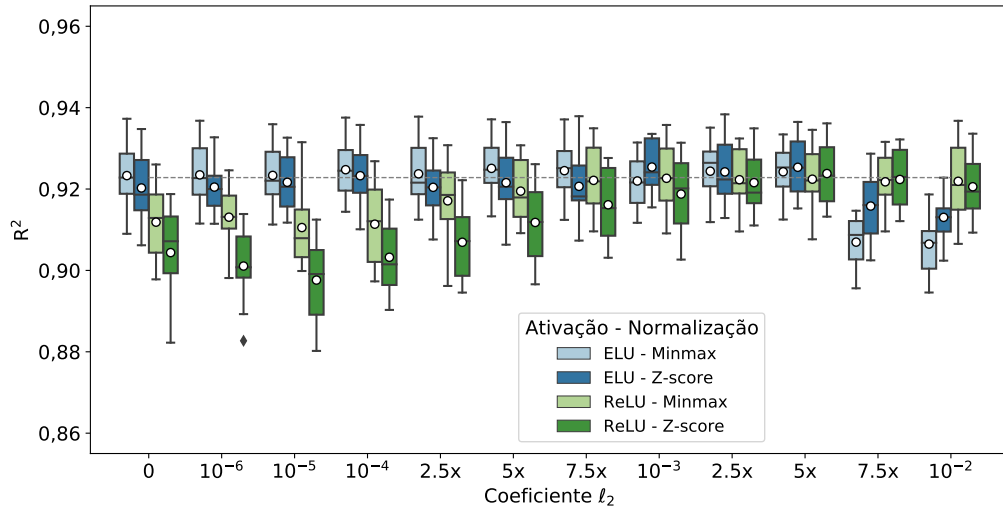
(c) Usina EUA

Fonte: elaboração própria (2022).

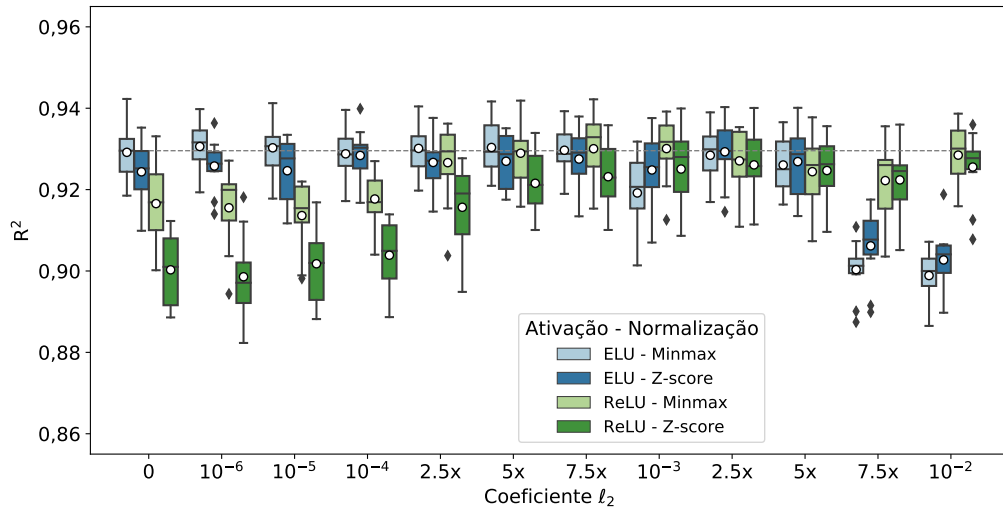
Figura I.4 – R^2 das usinas para o treinamento com regularização do tipo Ridge (ℓ_2)



(a) Usina AUS



(b) Usina BRA



(c) Usina EUA

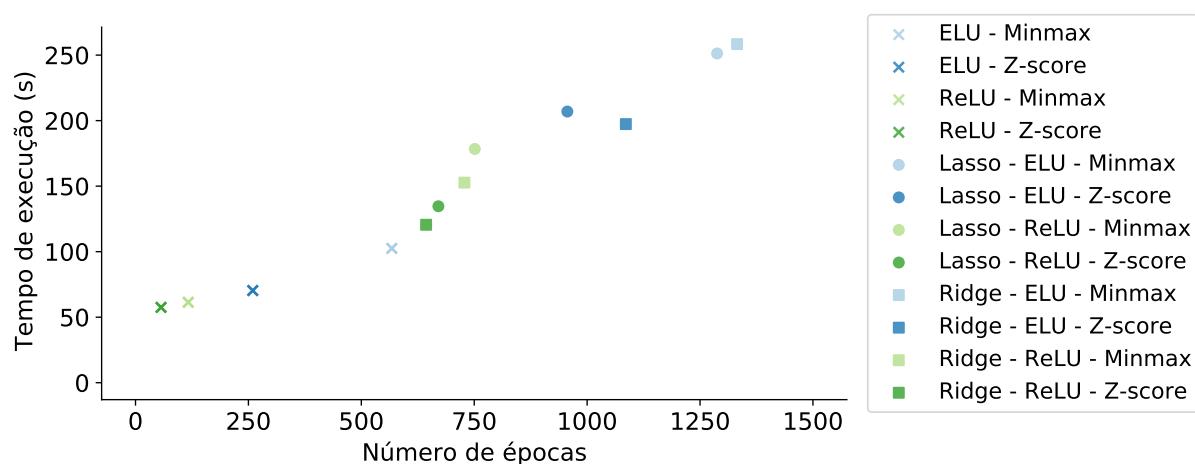
Fonte: elaboração própria (2022).

Tabela I.3 – Melhores resultados por usina – com e sem regularização

Usina	Antes	c/ Lasso	c/ Ridge
AUS			
configuração	Minmax/ELU	Z-score/ReLU	Z-score/ReLU
coeficiente de regularização		0,00025	0,005
resultado (R^2)	0,8688	0,8873	0,8880
BRA			
configuração	Minmax/ELU	Z-score/ReLU	Z-score/ReLU
coeficiente de regularização		0,0005	0,005
resultado (R^2)	0,9233	0,9275	0,9238
EUA			
configuração	Minmax/ELU	Z-score/ReLU	Z-score/ReLU
coeficiente de regularização		0,00025	0,0025
resultado (R^2)	0,9292	0,9334	0,9261

Fonte: elaboração própria (2022).

O tempo de execução mais que dobra com a regularização, como mostra a Figura I.5, sendo Lasso ligeiramente mais rápido que Ridge. ReLU e ELU, Z-score e Minmax mantêm proporções semelhantes antes e depois.

Figura I.5 – Tempo de execução versus nº de épocas até o sobreajuste – todas usinas

Fonte: elaboração própria (2022).

Os resultados do experimento parecem suficientes para sustentar que:

- a regularização pode contribuir para o aumento de generalização;
- principalmente no caso das configurações com maior sobreajuste prévio, que se tornam competitivas com a regularização; e
- o custo de treinamento é cerca de 2x maior com a regularização.

APÊNDICE J – EXPERIMENTO 4

Este experimento foi realizado de acordo com as condições gerais apresentadas pela Tabela J.1. O objetivo foi verificar se o processo de regularização do tipo Rede Elástica – ou *Elastic Net*, que combina Ridge e Lasso – é capaz de aumentar a capacidade de generalização das RNAs estudadas.

Tabela J.1 – Condições gerais do Experimento 4

Aspecto	Valor(es) testado(s)
Ferramenta	Keras (TensorFlow) sem GPU
Estrutura MLP	
Camada de entrada	17 neurônios (11 atributos ^a)
Camada(s) oculta(s)	4 camadas (20.000 parâmetros)
Camada de saída	1 neurônio
Ajustes	
Normalização de atributos	Z-score
Parâmetros iniciais	Glorot uniforme, com vieses zerados
Função de ativação	ReLU
Função de custo	erro quadrático médio (MSE)
Otimizador	SGD c/ momento e nesterov
Função de regularização	Rede elástica ($\ell_1 + \ell_2$)
Lote	32 exemplos
N° de épocas	máximo de 10.000 c/ paciência ^b de 500
Atributos	<i>temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h, elevref, azim, hrang e dayang</i>
Validação	Cruzada de 10 parcelas c/ estratificação pela cobertura de nuvens do início do período (< 25%, intermediário e > 75%) e a geração de energia

^a Os atributos *temp, pressure, humidity, clouds, wind_speed* e *wind_deg* são pares de valores relativos ao instante inicial e final de cada hora.

^b Um algoritmo interrompe o treinamento na ausência de melhoria (*early stopping*), conforme os critérios detalhados no Apêndice F.

Fonte: elaboração própria (2022).

Uma estrutura de quatro camadas ocultas, 78 neurônios por camada e cerca de 20 mil parâmetros (ou cinco vezes o número de exemplos de treinamento) foi selecionada. As amostras de desenvolvimento das três usinas (AUS, BRA e EUA) relativas a um ano de dados horários (horas claras) foram parceladas pela técnica de

validação cruzada de 10 parcelas, estratificadas segundo a cobertura de nuvens no início do período (*clouds*) e a geração de energia, conforme o Apêndice F. Médias e desvios-padrão de geração de energia¹ (p.u.) das parcelas de treinamento e de validação ficaram muito próximos, como mostra a Tabela J.2.

Tabela J.2 – Média e desvio-padrão da geração de energia por usina

	Média (p.u.)	DP (p.u.)
AUS	0,284 ± 0,001	0,254 ± 0,001
BRA	0,320 ± 0,002	0,260 ± 0,001
EUA	0,251 ± 0,003	0,240 ± 0,002

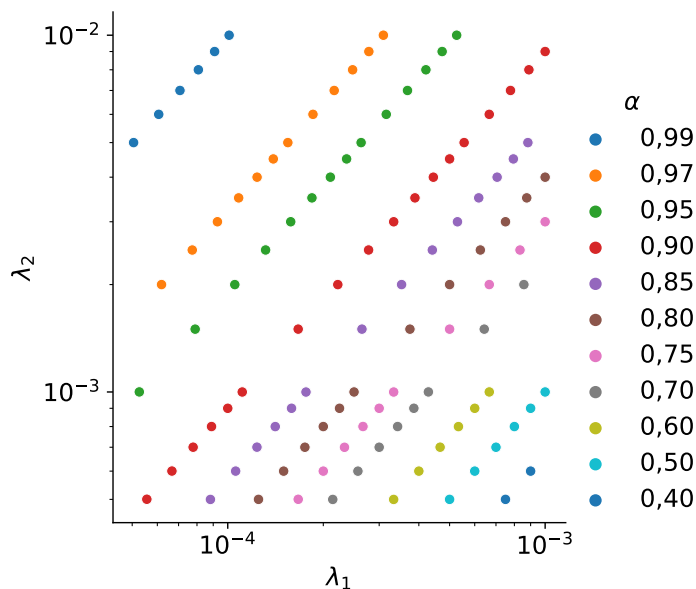
Fonte: elaboração própria (2022).

Foram comparados os resultados, em termos do R^2 , com a variação:

- dos coeficientes λ_2 de 5×10^{-4} a 10^{-2} ;
- da relação $\alpha = \lambda_2 / (\lambda_1 + \lambda_2)$ na faixa entre 0,4 e 1 (um);

resultando no mapeamento ilustrado pela Figura J.1.

Figura J.1 – Mapeamento dos coeficientes da regularização por rede elástica

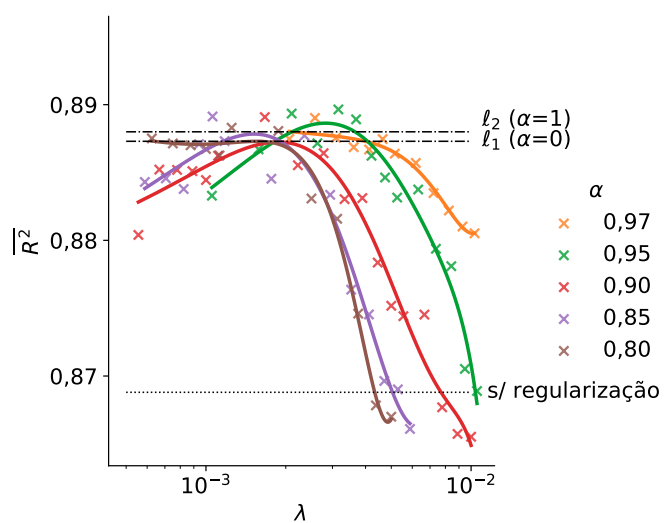


Fonte: elaboração própria (2022).

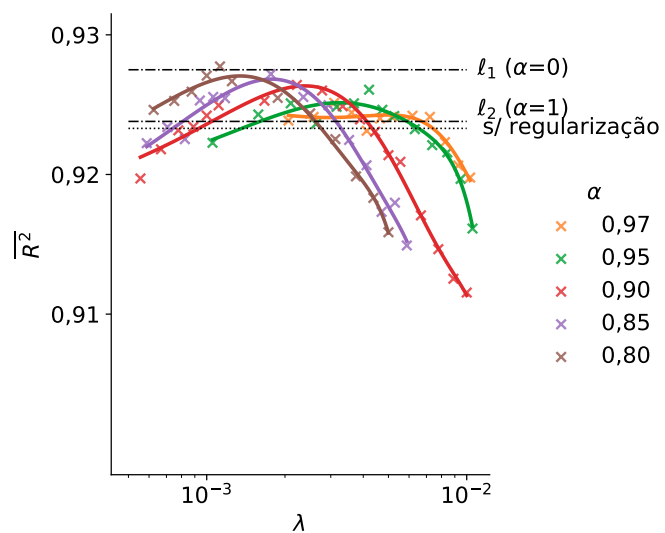
A Figura J.2 compara algumas das curvas α de melhor desempenho de cada usina com os resultados obtidos no Experimento 3 (Apêndice I), quando a regularização Lasso (ℓ_1) e Ridge (ℓ_2) foram testadas separadamente.

¹Os dados de geração horária de energia foram normalizados por unidade (p.u.) com base no pico de capacidade de cada usina: 2,04 kWp (AUS), 95 kWp (BRA) e 140,14 kWp (EUA).

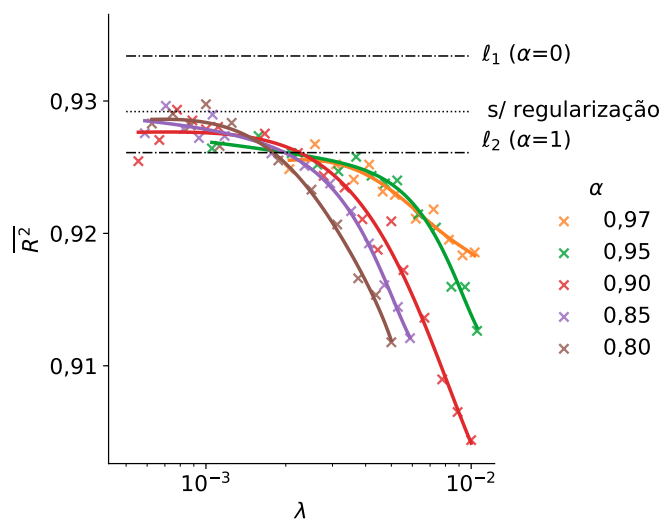
Figura J.2 – R^2 resultante do treinamento com regularização do tipo Rede Elástica



(a) Usina AUS



(b) Usina BRA



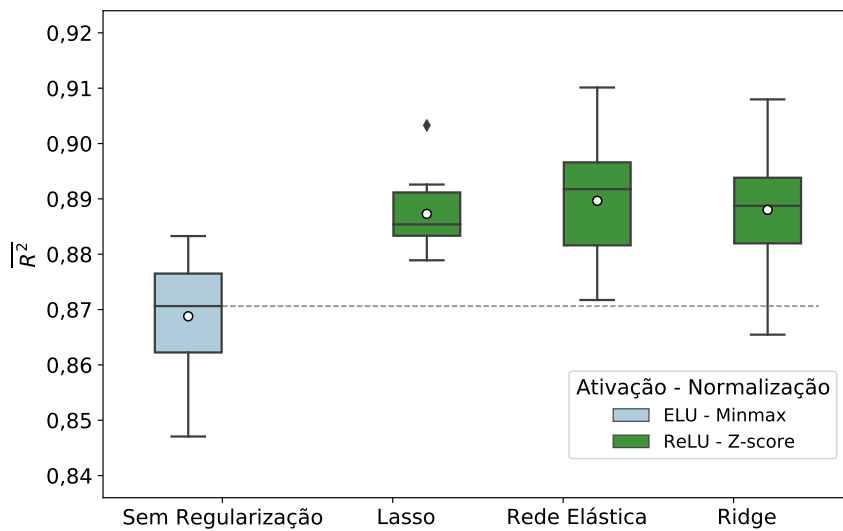
(c) Usina EUA

Fonte: elaboração própria (2022).

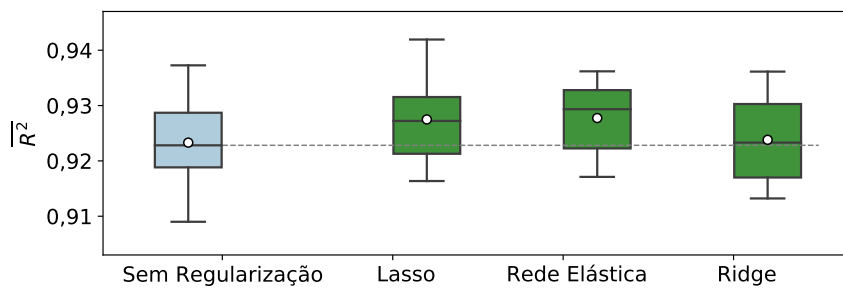
Os melhores resultados de validação por rede elástica foram superiores aos da regularização Ridge em todos os casos, porém, não tão significativos na comparação com Lasso. Se o processo de otimização de um coeficiente λ_1 é inconveniente, otimizar dois coeficientes (α e λ , ou λ_1 e λ_2) pode ser ainda mais.

A Figura J.3 apresenta a distribuição dos resultados de validação cruzada, incluindo também os resultados sem regularização. O custo-benefício da rede elástica parece pouco promissor para o tipo de problema em estudo.

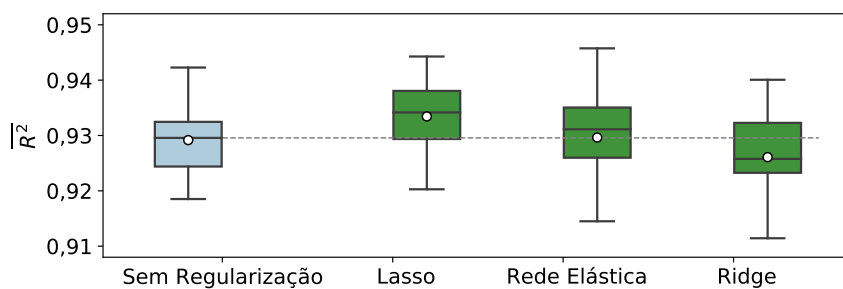
Figura J.3 – Comparativo dos resultados de validação com e sem regularização



(a) Usina AUS



(b) Usina BRA



(c) Usina EUA

Fonte: elaboração própria (2022).

APÊNDICE K – EXPERIMENTO 5

Este experimento foi realizado de acordo com as condições gerais apresentadas pela Tabela K.1. O objetivo foi verificar se o processo de regularização por Dados Aumentados através da aplicação de Ruído Gaussiano é capaz de melhorar a capacidade de generalização das RNAs estudadas.

Tabela K.1 – Condições gerais do Experimento 5

Aspecto	Valor(es) testado(s)
Ferramenta	Keras (TensorFlow) sem GPU
Estrutura MLP	
Camada de entrada	17 neurônios (11 atributos ^a)
Camada(s) oculta(s)	4 camadas (20.000 parâmetros)
Camada de saída	1 neurônio
Ajustes	
Normalização de atributos	Minmax e Z-score
Parâmetros iniciais	Glorot uniforme, com vieses zerados
Função de ativação	ELU e ReLU
Função de custo	erro quadrático médio (MSE)
Otimizador	SGD c/ momento e nesterov
Regularização	Dados aumentados por ruído gaussiano
Lote	32 exemplos
Nº de épocas	máximo de 10.000 c/ paciência ^b de 200
Atributos	<i>temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h, elevref, azim, hrang e dayang</i>
Validação	Cruzada de 10 parcelas c/ estratificação pela cobertura de nuvens do início do período (< 25%, intermediário e > 75%) e a geração de energia

^a Os atributos *temp, pressure, humidity, clouds, wind_speed* e *wind_deg* são pares de valores relativos ao instante inicial e final de cada hora.

^b Um algoritmo interrompe o treinamento na ausência de melhoria (*early stopping*), conforme os critérios detalhados no Apêndice F.

Fonte: elaboração própria (2022).

Uma estrutura de quatro camadas ocultas, 78 neurônios por camada e cerca de 20 mil parâmetros (ou cinco vezes o número de exemplos de treinamento) foi selecionada. As amostras de desenvolvimento das três usinas (AUS, BRA e EUA) relativas a um ano de dados horários (horas claras) foram divididas em 10 parcelas

estratificadas segundo a cobertura de nuvens no início do período (*clouds*) e a geração de energia, conforme o Apêndice F. Médias e desvios-padrão de geração de energia¹ (p.u.) das parcelas de treinamento e de validação ficaram muito próximos, como mostra a Tabela K.2.

Tabela K.2 – Média e desvio-padrão da geração de energia por usina

	Média (p.u.)	DP (p.u.)
AUS	0,284 ± 0,001	0,254 ± 0,001
BRA	0,320 ± 0,002	0,260 ± 0,001
EUA	0,251 ± 0,003	0,240 ± 0,002

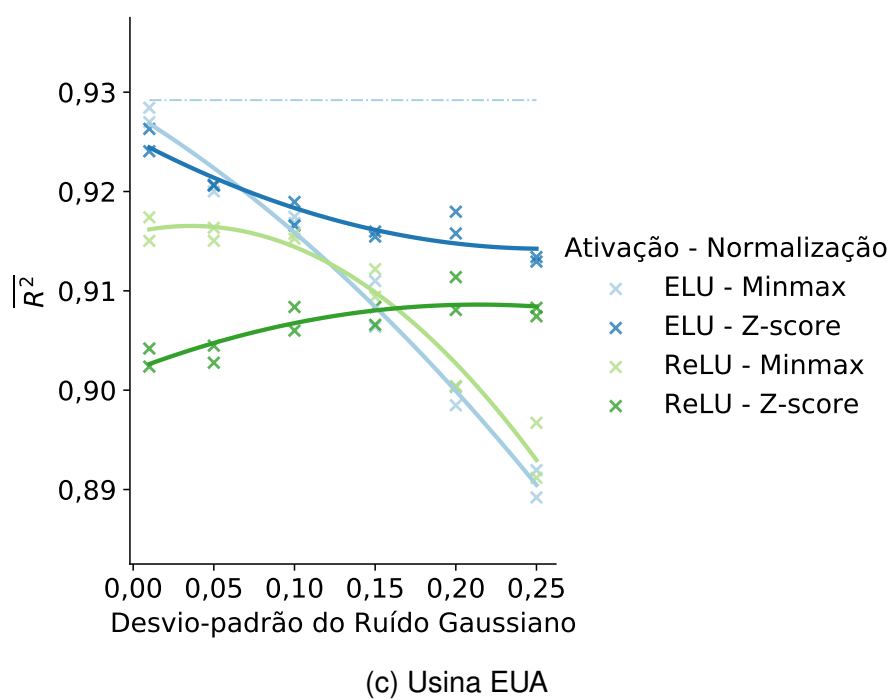
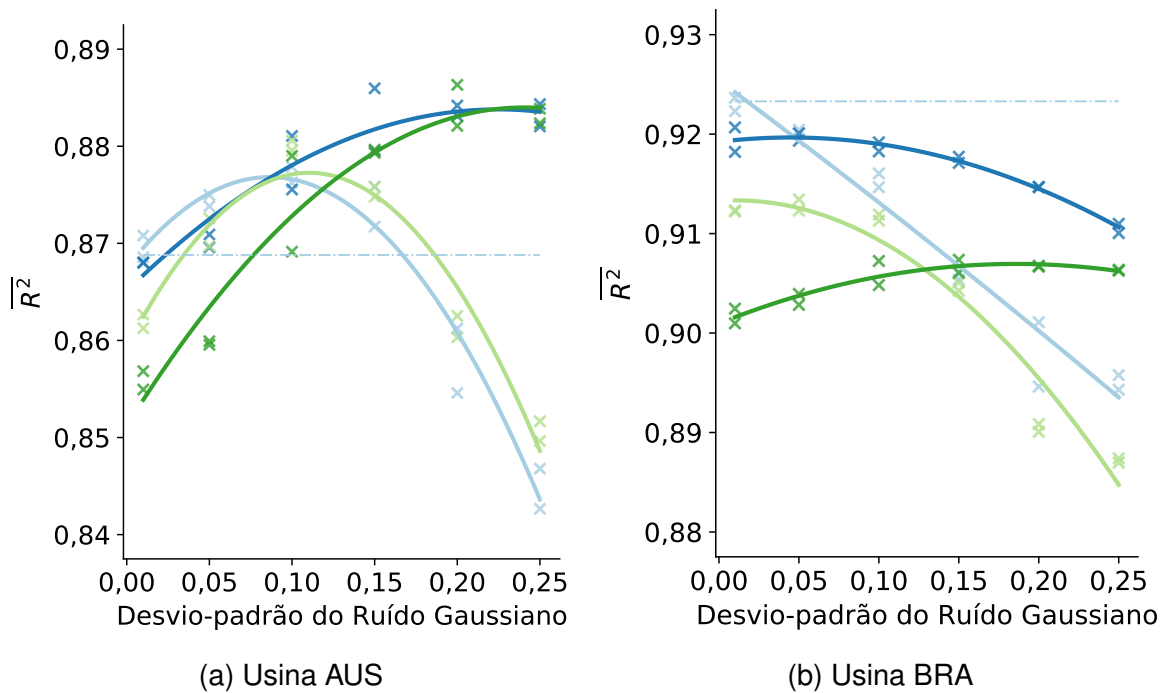
Fonte: elaboração própria (2022).

A Figura K.1 compara as curvas R^2 resultantes da aplicação de ruído gaussiano das três usinas e para os tipos de normalização Minmax e Z-score e as funções de ativação ReLU e ELU. São plotados os R^2 com a variação do desvio-padrão do ruído na faixa entre 0,01 (um centésimo) e 0,25 (vinte e cinco centésimos). Como os dados de entrada são previamente normalizados, o valor máximo testado representa 25% do desvio-padrão dos dados, no caso da normalização prévia do tipo Z-score, e até 62% para a normalização prévia do tipo Minmax. Portanto, o valor absoluto do ruído não deve ser comparado entre os tipos de normalização. O foco é a capacidade de regularização. A linha horizontal representa o resultado antes da regularização da configuração ELU/Minmax (de maior generalização). Foram duas realizações da validação cruzada.

A configuração ReLU/Z-score, naturalmente mais rápida e de menor generalização, teve o melhor ganho com a técnica, ao passo que a configuração naturalmente mais lenta e de menor generalização (ELU/Minmax) teve algum ganho apenas para a usina AUS. A primeira também foi a única que apresentou similaridade do melhor desvio-padrão do ruído entre as usinas, em torno de 0.2 (dois décimos).

¹Os dados de geração horária de energia foram normalizados por unidade (p.u.) com base no pico de capacidade de cada usina: 2,04 kWp (AUS), 95 kWp (BRA) e 140,14 kWp (EUA).

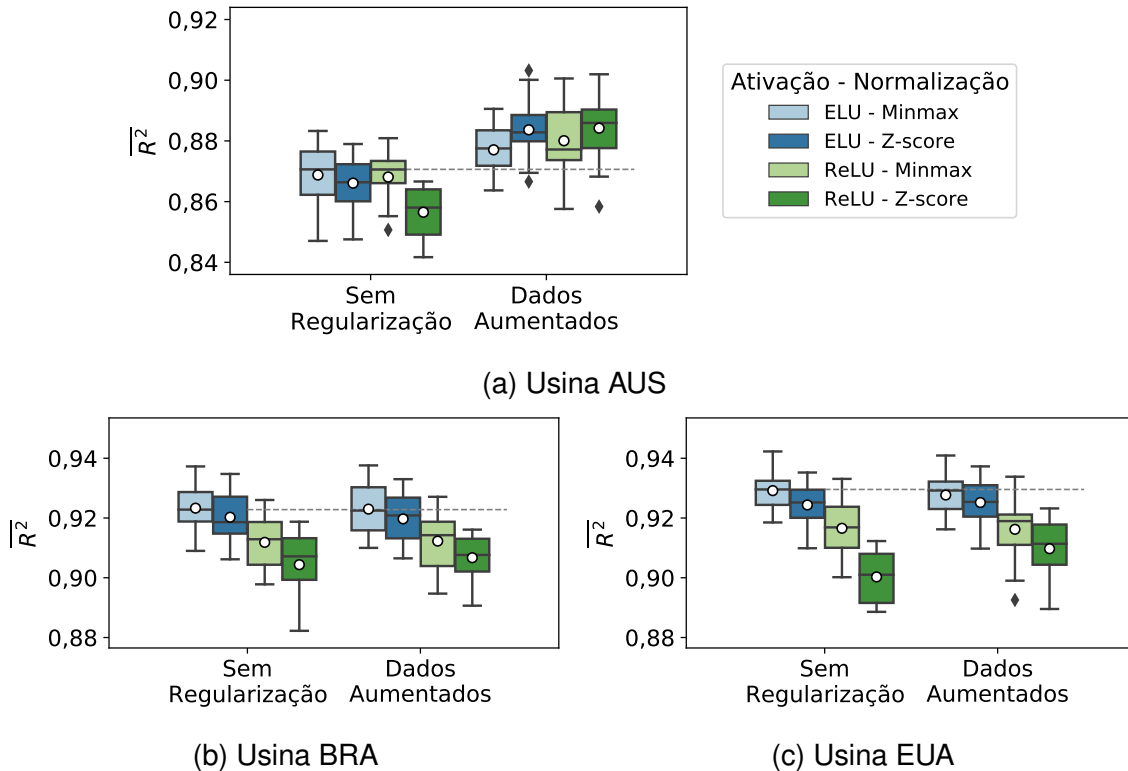
Figura K.1 – Curva R^2 do treinamento com regularização por aumento de dados



Fonte: elaboração própria (2022).

A Figura K.2 apresenta os diagramas de caixa² (*boxplot*) para comparação da distribuição dos resultados de validação cruzada antes da regularização e após aplicação do Aumento de Dados (melhores valores do desvio-padrão do ruído gaussiano).

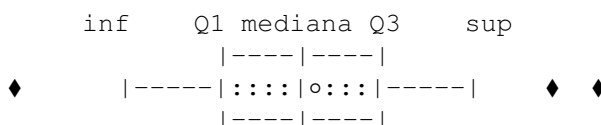
Figura K.2 – Resultados de validação com e sem regularização por Dados Aumentados



Fonte: elaboração própria (2022).

Os resultados mostram que, em alguns casos, é possível obter uma melhoria significativa com a aplicação de ruído gaussiano, como para a usina AUS, principalmente para a normalização de entradas do tipo Z-score (com ELU ou ReLU). No entanto, assim como as demais técnicas de regularização, esta depende de um ajuste ideal, e nem sempre é vantajosa, como visto para as usinas BRA e EUA, dependendo da normalização ou função de ativação utilizada.

²Diagrama de caixa é uma forma de representação de uma distribuição em termos da mediana (marca interna) e dos limites inferior (1º quartil) e superior (3º quartil) da área hachurada, limites extremos inferior (inf) e superior (sup) conectados por linhas, além de – eventualmente – pontos considerados fora da curva (♦). Na ausência desses pontos, mais distanciados do Q1 ou Q3 do que uma vez e meia a distância interquartil (Q3-Q1), os limites extremos coincidem com o mínimo (inf) e máximo (sup) (DEKKING *et al.*, 2005). Adicionalmente, pode ser exibida a média (o).



APÊNDICE L – EXPERIMENTO 6

Este experimento foi realizado de acordo com as condições gerais apresentadas pela Tabela L.1. O objetivo foi verificar se a variação do lote do treinamento com SGD é capaz de melhorar a capacidade de generalização das RNAs estudadas.

Tabela L.1 – Condições gerais do Experimento 6

Aspecto	Valor(es) testado(s)
Ferramenta	Keras (TensorFlow) sem GPU
Estrutura MLP	
Camada de entrada	17 neurônios (11 atributos ^a)
Camada(s) oculta(s)	4 camadas (20.000 parâmetros)
Camada de saída	1 neurônio
Ajustes	
Configurações	ELU/Minmax e ReLU/Z-score
Parâmetros iniciais	Glorot uniforme, com vieses zerados
Função de custo	erro quadrático médio (MSE)
Otimizador	SGD c/ momento e nesterov
Regularização	Nenhuma
Lote	8, 16, 24, 32, 40 e 48 exemplos
N° de épocas	máximo de 10.000 c/ paciência ^b de 500
Atributos	<i>temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h, elevref, azim, hrang e dayang</i>
Validação	Cruzada de 10 parcelas c/ estratificação pela cobertura de nuvens do início do período (< 25%, intermediário e > 75%) e a geração de energia

^a Os atributos *temp, pressure, humidity, clouds, wind_speed* e *wind_deg* são pares de valores relativos ao instante inicial e final de cada hora.

^b Um algoritmo interrompe o treinamento na ausência de melhoria (*early stopping*), conforme os critérios detalhados no Apêndice F.

Fonte: elaboração própria (2022).

Uma estrutura de quatro camadas ocultas, 78 neurônios por camada e cerca de 20 mil parâmetros (ou cinco vezes o número de exemplos de treinamento) foi selecionada. As amostras de desenvolvimento das três usinas (AUS, BRA e EUA) relativas a um ano de dados horários (horas claras) foram divididas em 10 parcelas estratificadas segundo a cobertura de nuvens no início do período (*clouds*) e a geração

de energia, conforme o Apêndice F. Médias e desvios-padrão de geração de energia¹ (p.u.) das parcelas de treinamento e de validação ficaram muito próximos, como mostra a Tabela L.2.

Tabela L.2 – Média e desvio-padrão da geração de energia por usina

	Média (p.u.)	DP (p.u.)
AUS	0,284 ± 0,001	0,254 ± 0,001
BRA	0,320 ± 0,002	0,260 ± 0,001
EUA	0,251 ± 0,003	0,240 ± 0,002

Fonte: elaboração própria (2022).

A Figura L.1 apresenta os diagramas de caixa² (*boxplot*) comparando a métrica R^2 resultante do uso de diferentes tamanhos de lotes do gradiente descendente estocástico. A linha horizontal tracejada identifica a mediana dos resultados com o lote de 32 exemplos, utilizado durante os experimentos anteriores. Não há ganho significativo com a redução ou aumento deste hiperparâmetro para a configuração mais generalista – ELU/Minmax. Para a mais rápida – ReLU/Z-score – há alguma vantagem na redução do lote para as usinas AUS e EUA. Chama a atenção, positivamente, a baixa suscetibilidade à variação do tamanho do lote observada para a configuração ELU/Minmax – de referência.

¹Os dados de geração horária de energia foram normalizados por unidade (p.u.) com base no pico de capacidade de cada usina: 2,04 kWp (AUS), 95 kWp (BRA) e 140,14 kWp (EUA).

²Diagrama de caixa é uma forma de representação de uma distribuição em termos da mediana (marca interna) e dos limites inferior (1º quartil) e superior (3º quartil) da área hachurada, limites extremos inferior (inf) e superior (sup) conectados por linhas, além de – eventualmente – pontos considerados fora da curva (♦). Na ausência desses pontos, mais distanciados do Q1 ou Q3 do que uma vez e meia a distância interquartil (Q3-Q1), os limites extremos coincidem com o mínimo (inf) e máximo (sup) (DEKKING *et al.*, 2005). Adicionalmente, pode ser exibida a média (o).

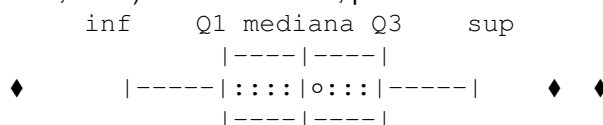
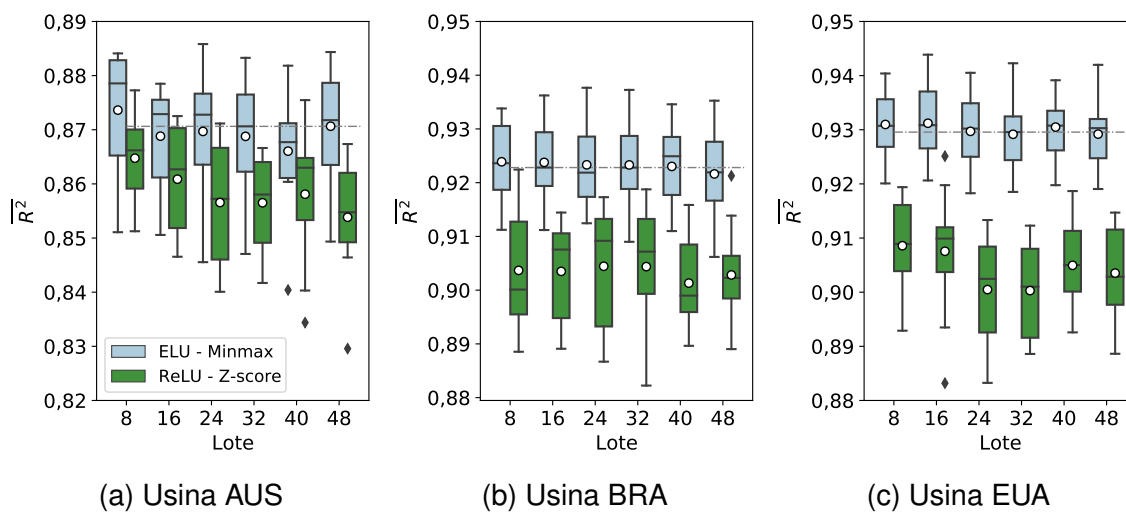


Figura L.1 – $\overline{R^2}$ do treinamento com diferentes tamanhos de lotes do SGD



Fonte: elaboração própria (2022).

APÊNDICE M – EXPERIMENTO 7

Com o objetivo de verificar se a taxa de aprendizado inicial padrão do pacote TensorFlow/Keras para a otimização por SGDM é satisfatória para a solução do problema proposto, este experimento verificou uma faixa de valores para as três usinas, utilizando-se das configurações apresentadas pela Tabela M.1.

Tabela M.1 – Especificação da rede neural artificial de teste

Característica	Valor
Entradas (atributos)	<i>azim, elevref, dayang, hrang, temp, pressure, humidity, clouds, wind_speed, wind_deg, rain_1h</i>
Normalização	Minmax {-1, 1}
Saída	Geração de energia (em p.u. ^b)
Tipo de rede	<i>perceptron</i> multicamadas (MLP)
Estrutura de camadas	
Entrada	18 neurônios (11 atributos ^a)
Ocultas	6 camadas de 61 neurônios
Saída	1 neurônio
Pesos e vieses	
Quantidade	20.131 parâmetros
Método de inicialização	Glorot uniforme c/ vieses zerados
Função de ativação	Unidade linear exponencial (ELU)
Otimizador	Gradiente descendente estocástico (SGD)
Variação	c/ Momento e Nesterov
Taxas de aprendizado	diferentes valores
Taxa de aceleração	0,9
Lote	diferentes valores
Nº de épocas	máximo de 10.000 c/ paciência ^c de 500
Função de custo	Erro quadrático médio (MSE)
Validação	Cruzada de 10 parcelas c/ estratificação pela cobertura de nuvens do início do período (< 25%, intermediário e > 75%) e a geração de energia

^a Os atributos de condições do tempo – *temp, pressure, humidity, clouds, wind_speed, wind_deg* – são pares de valores relativos ao instante inicial e final de cada hora. O atributo de geoposicionamento solar *dayang* é desdobrado em duas componentes: seno e cosseno. Somam-se a esses os demais atributos de geoposicionamento solar – *azim, elevref*, e *hrang*.

^b A base é a potência nominal de cada usina.

^c Um algoritmo interrompe o treinamento na ausência de melhoria (*early stopping*), conforme os critérios detalhados no Apêndice F.

As amostras de desenvolvimento das três usinas (AUS, BRA e EUA) relativas a um ano de dados horários (horas claras) foram divididas em 10 parcelas estratificadas segundo a cobertura de nuvens no início do período (*clouds*) e a geração de energia, conforme o Apêndice F. Médias e desvios-padrão de geração de energia¹ (p.u.) das parcelas de treinamento e de validação ficaram muito próximos, como mostra a Tabela M.2.

Tabela M.2 – Média e desvio-padrão da geração de energia por usina

	Média (p.u.)	DP (p.u.)
AUS	0,284 ± 0,001	0,254 ± 0,001
BRA	0,320 ± 0,002	0,260 ± 0,001
EUA	0,251 ± 0,003	0,240 ± 0,002

Fonte: elaboração própria (2022).

A Figura M.1 apresenta os diagramas de caixa² (*boxplot*) comparando a métrica R^2 resultante do uso de diferentes taxas de aprendizado do gradiente descendente estocástico. A linha horizontal identifica a mediana dos resultados com a taxa de aprendizado padrão sugerida pelo pacote TensorFlow (ABADI *et al.*, 2015) de 0,01, e utilizado em todos os experimentos anteriores.

Não há ganho significativo com a redução ou aumento deste hiperparâmetro de forma geral, embora pequenas vantagens de uma taxa maior ou menor possam ser vistas para as usinas individualmente. O teste com valores acima de 0,3 apresentaram provável divergência, com resultados de previsão na notação 'nan', geralmente associados às operações envolvendo valores infinitos (ou divisão por zero).

Pode-se verificar ainda, na Figura M.2, que o número de épocas necessárias até o melhor resultado – ou início do sobreajuste – cai significativamente com o aumento da taxa de aprendizado, algo esperado.

Assim, considerando uma ligeira melhora da distribuição do R^2 com o aumento da taxa de aprendizado, e sabendo-se que a época de melhor validação cruzada

¹Os dados de geração horária de energia foram normalizados por unidade (p.u.) com base no pico de capacidade de cada usina: 2,04 kWp (AUS), 95 kWp (BRA) e 140,14 kWp (EUA).

²Diagrama de caixa é uma forma de representação de uma distribuição em termos da mediana (marca interna) e dos limites inferior (1º quartil) e superior (3º quartil) da área hachurada, limites extremos inferior (inf) e superior (sup) conectados por linhas, além de – eventualmente – pontos considerados fora da curva (♦). Na ausência desses pontos, mais distanciados do Q1 ou Q3 do que uma vez e meia a distância interquartil (Q3-Q1), os limites extremos coincidem com o mínimo (inf) e máximo (sup) (DEKKING *et al.*, 2005). Adicionalmente, pode ser exibida a média (○).

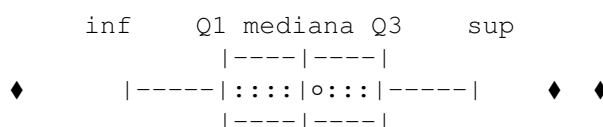
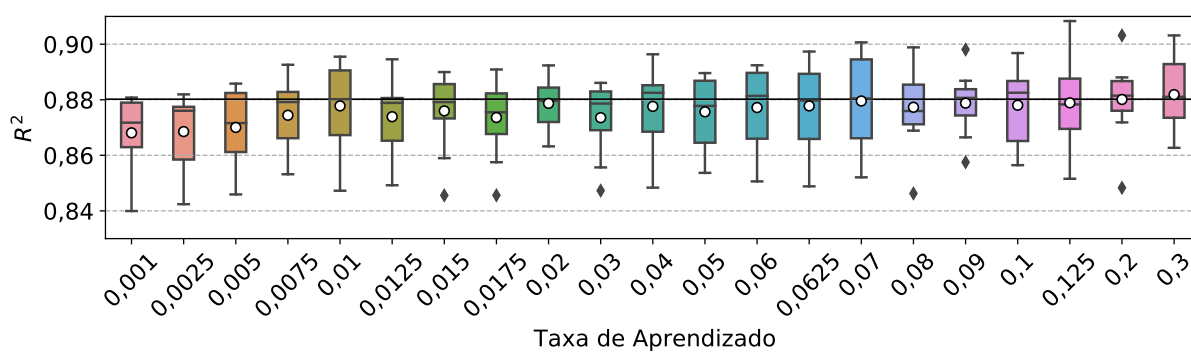
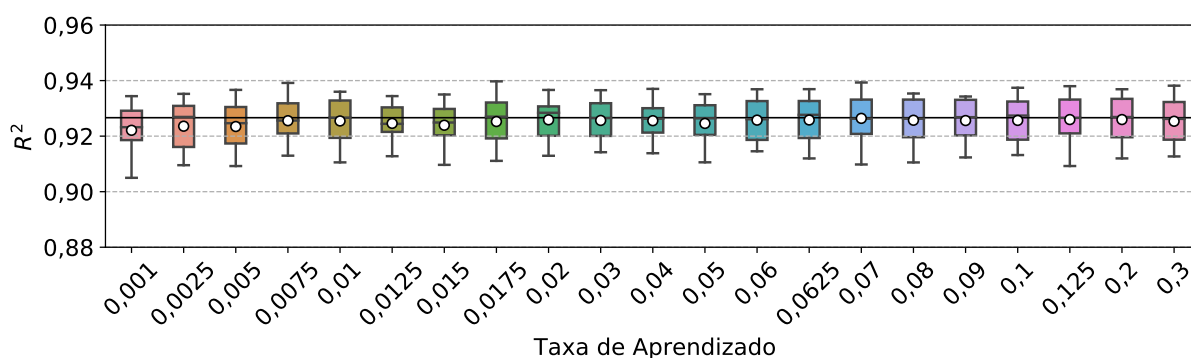


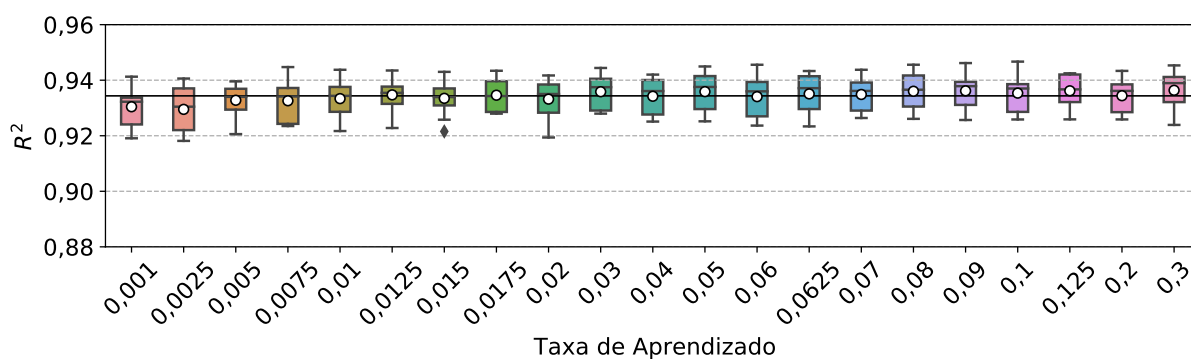
Figura M.1 – R^2 de validação cruzada para diferentes taxas de aprendizado do SGD



(a) Usina AUS



(b) Usina BRA

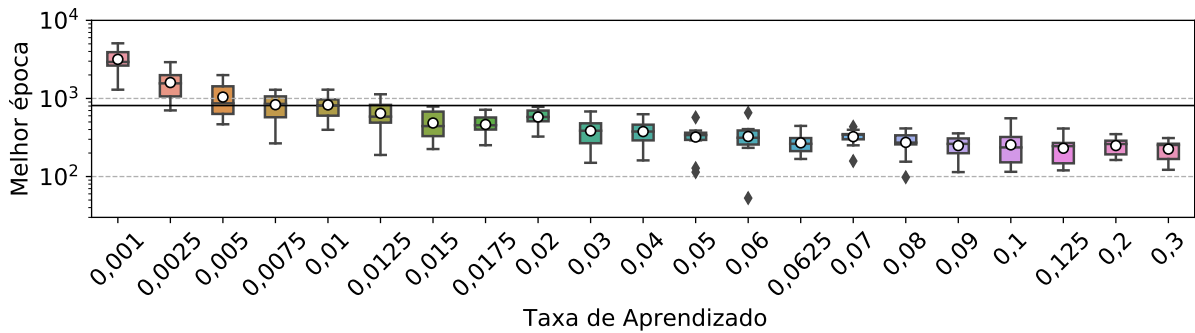


(c) Usina EUA

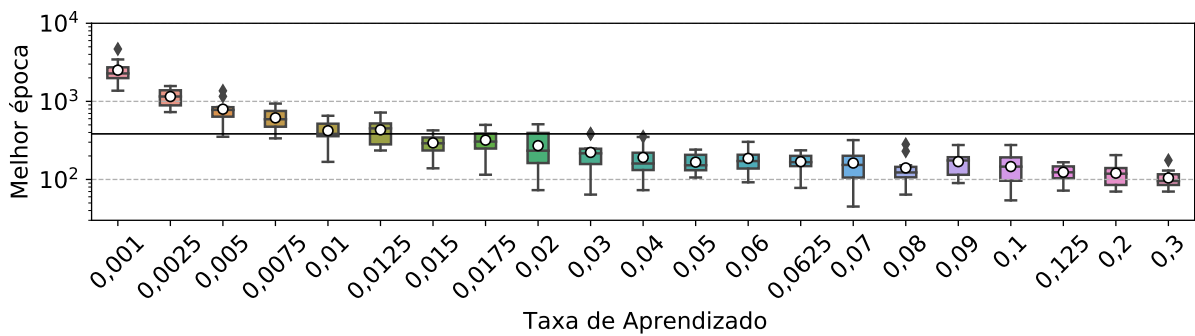
Fonte: elaboração própria (2022).

pode cair significativamente, um novo teste foi realizado para verificar se diante de taxas de aprendizado mais elevadas poderia haver alteração dos resultados obtidos com aplicação da regularização lasso ou com diferentes lotes do SGD. Os coeficientes foram testados na mesma ordem de grandeza dos experimentos anteriores.

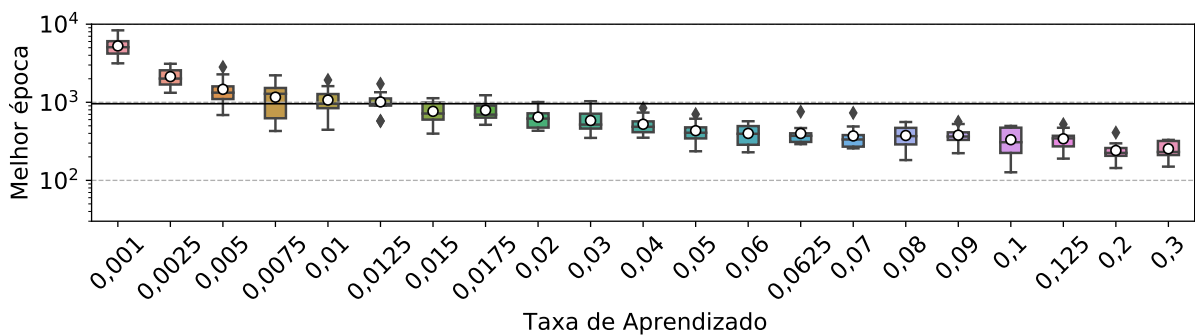
Figura M.2 – Época de melhor validação cruzada para diferentes taxas de aprendizado do SGD



(a) Usina AUS



(b) Usina BRA

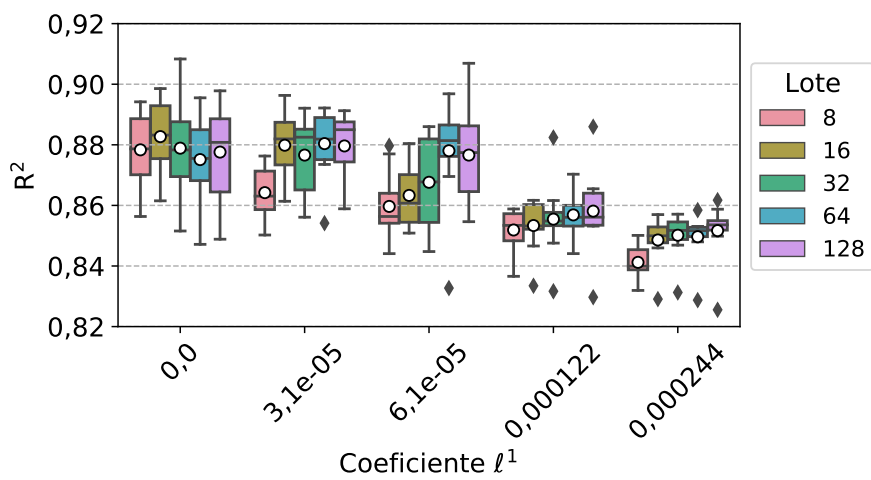


(c) Usina EUA

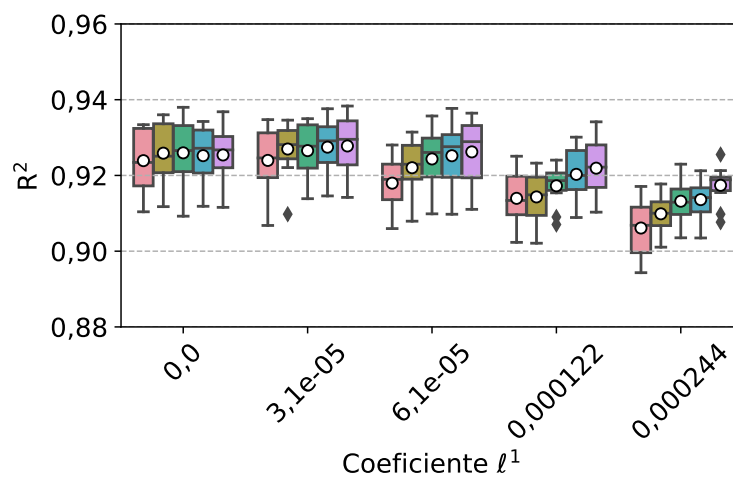
Fonte: elaboração própria (2022).

Não houve ganho com a aplicação de regularização, senão marginal, para uma usina ou outra, e determinado tamanho de lote do SGD. A Figura M.3 exemplifica os resultados obtidos, para a taxa de aprendizado de 0,125.

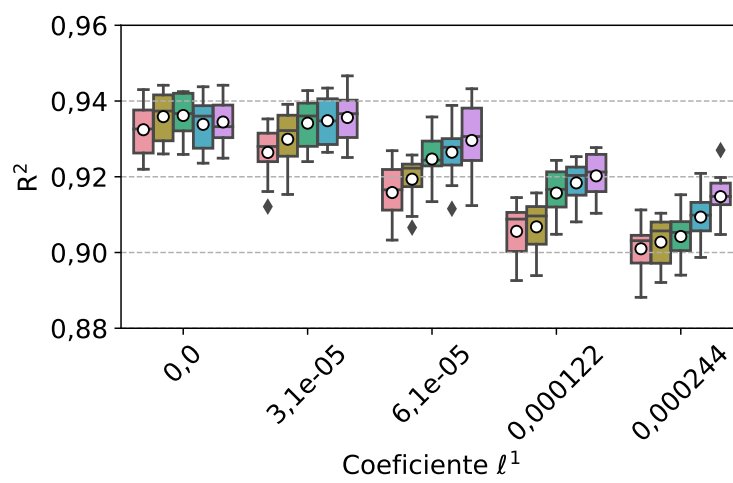
Figura M.3 – Validação cruzada para taxa de aprendizado de 0,125 e diferentes coeficientes lasso e lotes do SGD



(a) Usina AUS



(b) Usina BRA

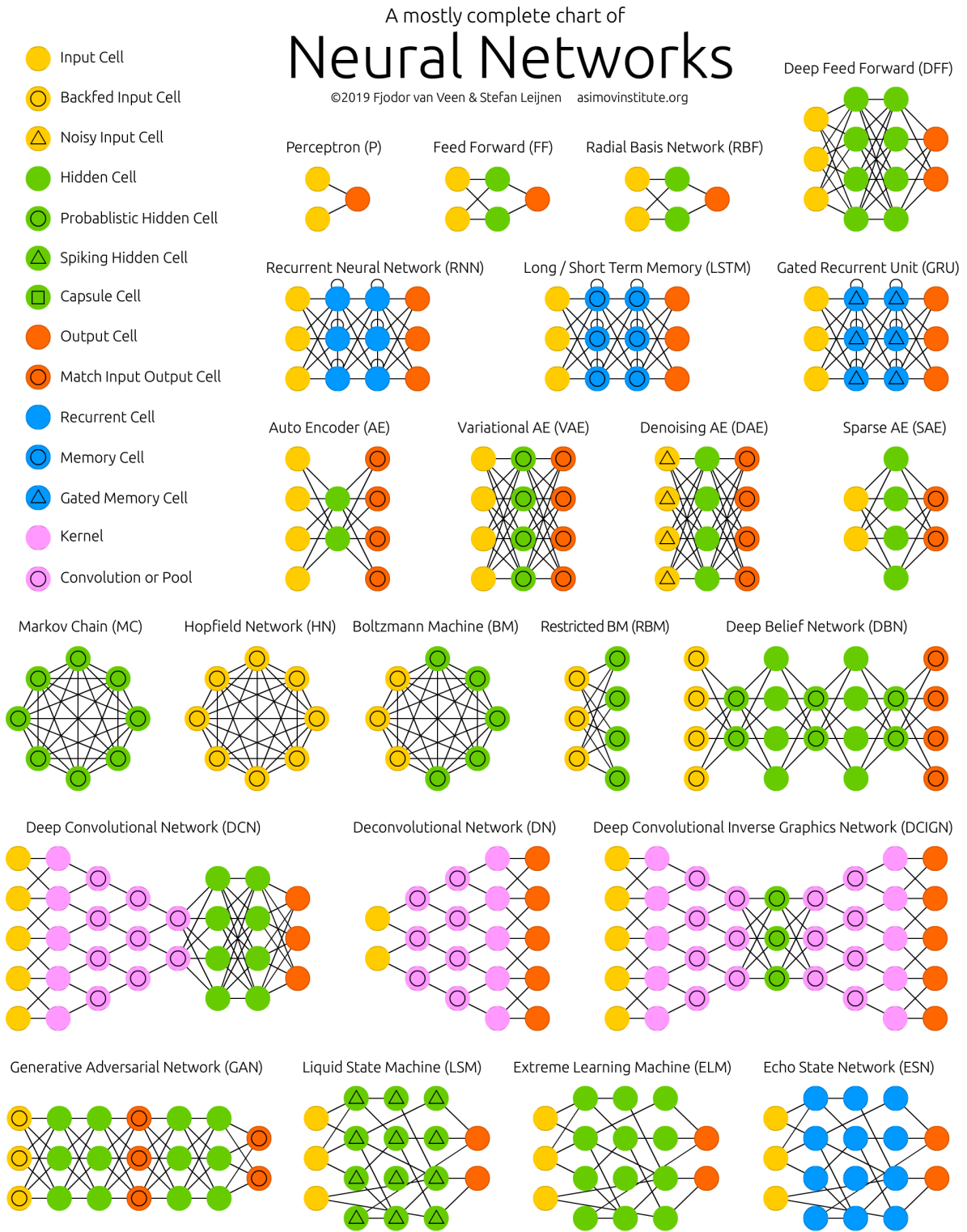


(c) Usina EUA

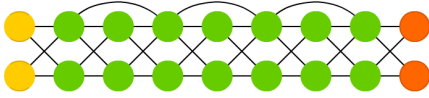
Fonte: elaboração própria (2022).

Anexos

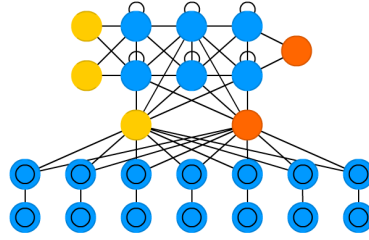
ANEXO A – DIAGRAMAS DE TIPOS DE REDES NEURAIS



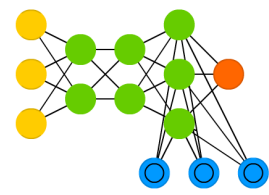
Deep Residual Network (DRN)



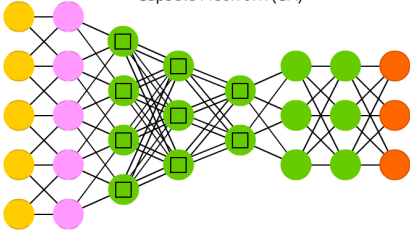
Differentiable Neural Computer (DNC)



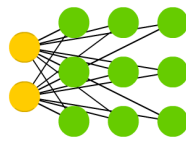
Neural Turing Machine (NTM)



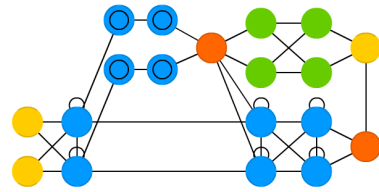
Capsule Network (CN)



Kohonen Network (KN)



Attention Network (AN)



ANEXO B – DADOS DA AUSGRID DE USINAS DA AUSTRÁLIA



Solar home electricity data - notes (August 2014) Three years of half-hour electricity data for 300 solar homes (1 July 2010 to 30 June 2013)

Purpose

We are releasing this data for use by organisations and individuals for a variety of purposes, including research, policy-making and providing information about solar photovoltaic system performance. We intend for the data to help with analysis, including by research organisations, the solar industry, government departments and regulators.

Data overview

The data has been sourced from 300 randomly selected solar customers in Ausgrid's electricity network area that were billed on a domestic tariff and had a gross metered solar system installed for the whole of the period from 1 July 2010 to 30 June 2013. The customers chosen had a full set of actual data for the period from 1 July 2010 to 30 June 2011, gathered through our meter reading processes. We also undertook some data quality checking and excluded customers on the high and low ends of household consumption and solar generation performance during the first year.

Disclaimers and use of data

Ausgrid accepts no liability for the use of this data by any organisation. The customers in this dataset may not represent a statistically relevant sample of residential customers in the Ausgrid network area, and have not been surveyed to collect household characteristics. Typically, households that install solar systems own their home and live in separate houses with the available roof space needed to install a solar power system.

Further inquiries

If you have further inquiries about this data or would like to provide feedback, including how the data is being used by your organisation and how useful it is, please email sharinginformation@ausgrid.com.au.

Data format

The solar household data is contained in a zipped file and each year of data is in a separate file in a csv file format. It contains five description columns for each line of data, followed by 48 columns of half-hour electricity meter data. The time format for the 48 columns of interval data is Eastern Standard Time (EST) and Eastern Daylight Savings Time (EDT) during the summer period.

Column	Field	Description
1	Customer	Customer ID from 1 to 300
2	Postcode	Postcode location of customer
3	Generator Capacity	Solar panel capacity recorded on the application for connection for each customer. Units are Kilowatt Peak (kWp), which is the solar panels peak power under full solar radiation and tested under standard conditions.

4	Consumption Category	Two letter code each meaning the following: GC = General Consumption for electricity supplied all the time (primary tariff, either inclining block or time of use rates), excluding solar generation and controlled load supply CL = Controlled Load Consumption (Off peak 1 or 2 tariffs) GG = Gross Generation for electricity generated by the solar system with a gross metering configuration, measured separately to household loads
5	Date	Date in DDMMYYYY format.
6	0:30	Kilowatt hours (kWh) of electrical energy consumed or generated in the half hour ending at 0:30 (eg. between 0:00 and 0:30). The value is positive regardless of whether it is consumption or generation.
7 to 53	1:00.....0:00	As above, covering every half hour of the day up until the last half hour of the day at 0:00, (eg. between 23:30 to 0:00).
54	Row Quality	(Blank) = every half hour value in the row is the actual electricity recorded by the meter in the half hour NA = Non-Actual where some or all of the half hour values in the row are estimates or substitutes of the electricity consumed or generated

Summary statistics

The table below shows some summary statistics of the dataset over the year. It is provided as a basis for comparison and checking purposes.

Year	2010-11		2011-12		2012-13	
	Mean	Median	Mean	Median	Mean	Median
Solar home customers – summary (300 sample)						
Annual consumption; kWh per year	6,980	6,362	6,596	6,017	6,387	5,862
Annual gross generation; kWh per year	2,119	1,764	2,083	1,708	2,181	1,814
Solar system size (kWp)	1.68	1.50	1.68	1.50	1.68	1.50
Annual gross generation; kWh/kWp	1,253	1,280	1,231	1,253	1,297	1,326
Ausgrid residential customers – summary* (>1.4 million)						
Annual consumption; kWh per year	6,611	-	6,224		5,954	

*Ausgrid residential customer electricity data can be found at [on our website](#).

IPART report on solar feed-in tariffs in NSW

The solar homes in this dataset were part of a sample of data supplied by Ausgrid to the NSW Independent Pricing and Regulatory Tribunal (IPART) for their review into a fair and reasonable solar feed in tariff in NSW. The final IPART report was published in March 2012 and can be found at:

http://www.ipart.nsw.gov.au/Home/Industries/Electricity/Reviews/Retail_Pricing/Solar_feed-in_tariffs/14_Mar_2012_-_Final_Report/Final_Report_-_Solar_feed-in_tariffs_-_March_2012