



Previsão de Geração no Médio Prazo em um Parque Eólico no Rio Grande do Sul utilizando o GFS e Redes Neurais Artificiais

**João Gabriel de Farias¹, Reinaldo Hass², Júlio César Passos¹, Eduardo Victor Dias³,
Mayara Miqueletti de Lima³**

1 Departamento de Engenharia Mecânica
Universidade Federal de Santa Catarina - UFSC

2 Departamento de Física
Universidade Federal de Santa Catarina - UFSC

3 Setor de Pesquisa e Desenvolvimento
Atlantic Energias Renováveis S.A.
Campus Reitor João David Ferreira Lima, Florianópolis - SC - Brasil
joaogfarias@hotmail.com

ABSTRACT

Wind power prediction is important for the reliability of the electrical system, and can help the O&M planning for wind farms. This paper presents the performance evaluation of two methodologies to forecast the wind power generation with the time resolution of 30 min and the horizon of 120 hours ahead for the Aura Mirim II wind farm. The installed capacity of this wind farm is 30 MW and it is part of the Santa Vitória do Palmar wind farm complex, located in the city of Santa Vitória do Palmar, in the southernmost coastline of Rio Grande do Sul. The two methodologies were created applying Artificial Neural Networks: the first methodology uses the generation dataset with the sum of the power generated by the wind farm and the second obtain the wind farm power generation from the power generated by an equivalent wind turbine multiplied by the number of wind turbines in operation. Based on the results of cross-validation, the second methodology was the one that presented the lowest value of NRMSE, 22.2%, with a standard deviation of 6.0%.

Palavras-chave: *Redes Neurais Artificiais, Previsão de Geração, Santa Vitória do Palmar, Energia Eólica.*



1. INTRODUÇÃO

O mundo se direciona para a redução do consumo de combustíveis fósseis, conforme compromisso firmado por 186 países durante a COP 21 (Conferência da ONU sobre Mudanças Climáticas, em Paris). Durante o encontro estabeleceu-se que até 2050 o setor de geração de energia será livre de emissões. A energia eólica lidera essa transição para um planeta mais sustentável, com mais de 50 GW em capacidade instalada por ano desde 2014 [1]. Ao final de 2017, o Brasil possuía aproximadamente 12,76 GW de capacidade instalada acumulada, o que fez com que o país subisse para 8º lugar no ranking mundial. Esse valor cresceu para aproximadamente 14,76 GW, tal como observou-se em consulta ao Banco de Informação de Geração [2], no dia quatorze de fevereiro de dois mil e dezenove. Aproximadamente 9,0 % da matriz energética brasileira é composta por energia eólica. E essa parcela crescerá continuamente, conforme exposto no Plano Decenal de Expansão de Energia 2027 para que o aproveitamento do potencial eólico seja contínuo e de forma conjunta com a perspectiva da indústria nacional [3].

Tudo indica que a energia eólica terá uma participação cada vez maior no Sistema Interligado Nacional (SIN). Para manter a confiabilidade do sistema, visto que os ventos são intermitentes e a maior parcela de energia eólica aumenta o risco para o sistema de potência [4], torna-se necessário desenvolver e melhorar modelos de previsão de energia eólica [5]. Prever com certa confiabilidade permite reduzir o uso de termoeletricas e outras fontes despacháveis para manter a segurança do sistema. No Brasil, o Operador Nacional do Sistema Elétrico (ONS) demanda dos agentes geradores eólicos o envio de uma previsão com horizonte de 120 horas e resolução temporal de 30 min, conforme indicado no Submódulo 8.1 da ONS [6]. Além de auxiliar na maior segurança do sistema de potência, uma melhor previsão de geração eólica pode ajudar no planejamento da operação e manutenção dos empreendimentos.

A previsão de geração eólica pode ser classificada conforme o horizonte de tempo de previsão, sendo: muito curto prazo (*very short term*), o horizonte de minutos até algumas horas; curto prazo (*short term*), o horizonte de previsão vai do limite superior de previsão de muito curto prazo até 72 horas; médio prazo (*medium term*), que vai do limite de previsão de curto prazo até sete dias, o qual é viabilizado por modelos de previsão do tempo [7]; e longo prazo, que engloba



acima de uma semana até um ano ou mais [8]. Outra forma de classificar um modelo de previsão de geração é quanto aos seus dados de entrada [9], sendo: modelos estatísticos, os que empregam dados do SCADA ou dados de modelos globais de previsão meteorológica; modelos físicos, os que empregam previsões meteorológicas e características do terreno, por meio do método de *downscaling*, para prever a velocidade do vento, além de outras variáveis e então utilizar a curva de potência. Por fim, os modelos híbridos, que empregam dados do SCADA, previsões meteorológicas e características do terreno.

Há diversos estudos que visam a analisar modelos de previsão de geração, podendo-se citar [10] em que são comparados três modelos que utilizam Redes Neurais Artificiais (RNA) com diferentes dados de entrada. [10] empregaram como dados de entrada o modelo de previsão global *Global Forecast System* (GFS) e o modelo de mesoescala *Weather Research and Forecasting* (WRF). Os modelos foram testados em quatro parques eólicos com um horizonte de 72 horas, resultando em diferentes raízes do erro quadrático médio normalizado (NRMSE) para o melhor modelo de cada parque, os quais foram, 14,2% para o parque 01, 18,1% para o parque 02, 15,6% para o parque 03 e 12,4% para o parque 04. Essas diferenças no NRMSE estão de acordo com os resultados de [5], que indicam que a performance de um modelo de previsão depende de cada sítio.

No presente artigo, utiliza-se o modelo de previsão global GFS como dado de entrada para prever, por meio de redes neurais artificiais e treinamento supervisionado, a geração de um parque localizado no extremo Sul do Brasil, com um horizonte de 120 horas e resolução temporal de 30 min. Trata-se, portanto, de uma análise cujo modelo é classificado como estatístico ou de inteligência artificial para o médio prazo. Esse trabalho faz parte de um projeto de Pesquisa e Desenvolvimento (P&D) entre a Universidade Federal de Santa Catarina (UFSC), o Instituto Federal de Santa Catarina (IFSC) e a empresa Atlantic Energias Renováveis S.A. O principal objetivo do estudo aqui apresentado é avaliar dois procedimentos de previsão de geração do parque: o primeiro prevê diretamente a soma de geração do parque inteiro com o auxílio de um vetor de indisponibilidade; enquanto o segundo prevê a geração de um aerogerador equivalente que multiplicada pelo número de aerogeradores funcionando resulta na previsão de geração do parque.

2. PROCEDIMENTO DE CÁLCULO

Nesta seção é descrito o sítio eólico e fornecidas informações sobre os bancos de dados de geração, o modelo global de previsão meteorológico escolhido, o procedimento de seleção de variáveis e as diferentes topologias de redes neurais artificiais utilizadas.

2.1 Descrição do Sítio

Aura Mirim II (MIR-2) é o parque eólico, de propriedade da empresa Atlantic Energias Renováveis S.A, escolhido para execução deste estudo que faz parte do Complexo Eólico de Santa Vitória do Palmar (CESVP). Aura Mirim II possui 10 aerogeradores, modelo AW 125/3000 fabricado pela Acciona Wind Power, de 3 MW, o que totaliza 30 MW de capacidade instalada. O CESVP está localizado no litoral do extremo Sul do Rio Grande do Sul, no município de Santa Vitória do Palmar, nas seguintes coordenadas geográficas: $-33,15^\circ$ de Latitude e $-52,89^\circ$ de Longitude. Os aerogeradores possuem 120 m de altura do cubo (*hub height*) e 125 m de diâmetro do rotor. A torre de medição possui três anemômetros de copo a 57,5 m, 118,0 m e 120,0 m de altura, duas veletas (*windvane*), uma a 56,0 m e a outra a 118,0 m de altura, um termômetro, um higrômetro e um barômetro a 110,0 m de altura. Os valores de velocidades e direções, resultantes de médias de 10 min, são adquiridos a intervalos de 10 min. Na Figura 1, é mostrada a localização do parque no CESVP.

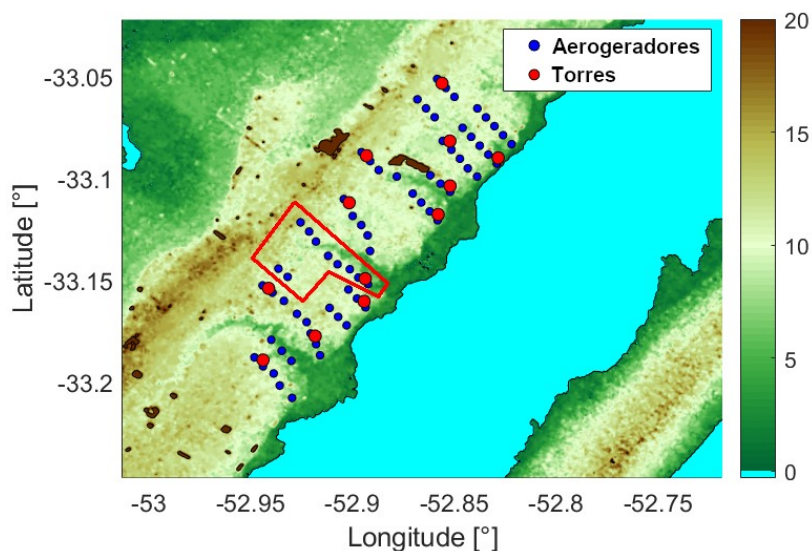


Figura 1 – CESVP com parque MIR-2 destacado por linha vermelha.

2.2 Banco de Dados de Geração

Os bancos de dados de geração utilizados neste estudo são o da soma da potência gerada pelo parque e o da potência gerada pelo aerogerador equivalente. Obteve-se o banco de dados da soma de geração do parque, potência média total gerada, pela soma da potência de cada aerogerador no mesmo instante. Já o banco de dados do aerogerador equivalente representa a média da geração do parque pelo número de aerogeradores funcionando. Foi criado o vetor de indisponibilidade que serve como uma variável de entrada para a rede neural que prevê a soma de geração do parque diretamente. Nas Figuras 2a e 2b, são mostradas as curvas de potência normalizadas, em relação ao valor mínimo e máximo de geração, do parque e do aerogerador equivalente, respectivamente, em função da velocidade do vento medida na torre de referência, lembrando que a resolução temporal dessa imagem é de 10 min.

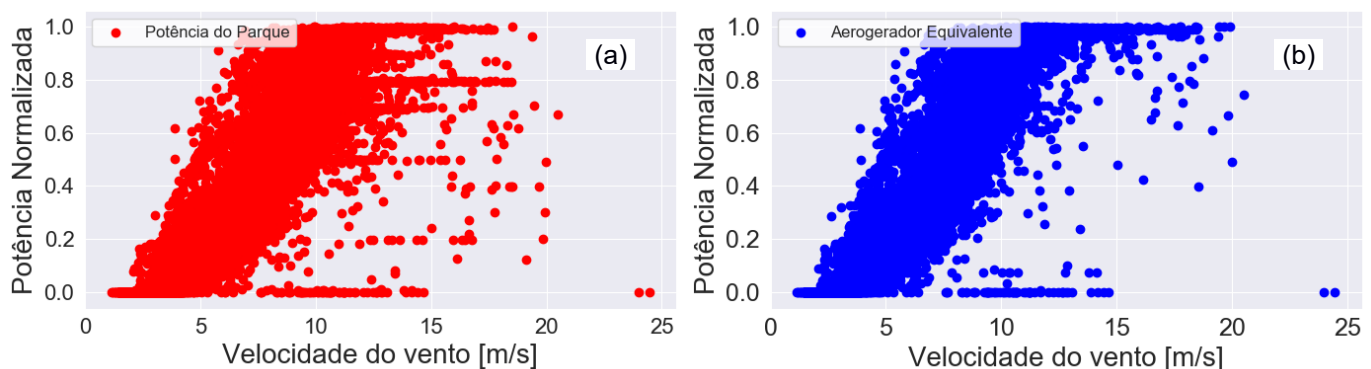


Figura 2 – Curva de potência normalizada do parque (a) e do aerogerador equivalente (b).

2.3 Sistema Global de Previsão – GFS (*Global Forecast System*)

O GFS é um modelo de previsão global, ou seja, tem por objetivo obter resultados para eventos atmosféricos para o planeta inteiro. A instituição responsável pelo GFS é o *National Center for Environmental Prediction* (NCEP). Esse modelo não consegue representar os processos meteorológicos de mesoescala. Utilizou-se o GFS de resolução espacial de $0,25^\circ$, cerca de 28 km, com saída horária entre 0 e 120 horas e com saída de 3 em 3 horas de 120 até 240 horas. Considerando-se os diferentes níveis de altura, o GFS $0,25^\circ$ tem em torno de 417 variáveis. O NCEP disponibiliza rodadas do modelo a cada 6 horas. Neste trabalho optou-se pelas rodadas disponibilizadas às 00 h do Horário do Meridiano de Greenwich (GMT). Interpolou-se

linearmente no espaço para obter o resultado das variáveis no ponto da torre de referência do parque MIR-2. Houve também a necessidade de interpolar linearmente no tempo para transformar a resolução temporal em 30 min. Além disso, transformou-se a hora de GMT para a hora local. Assim, para iniciar à meia-noite da hora local, pegou-se em cada rodada do GFS o resultado de 3 h até 123 h para totalizar as 120 horas em hora local. Com base nas componentes U e V do vetor velocidade para as alturas de 10, 80 e 100 m acima do solo, foram calculados o módulo e a direção do vetor velocidade.

2.4 Seleção de Variáveis

A seleção de variáveis é uma etapa importante para o desenvolvimento de modelos que empregam aprendizagem de máquina (*machine learning*), pois variáveis de entrada podem influenciar o desempenho das redes neurais artificiais. O emprego desta técnica a sistemas de geração com fontes renováveis de energia, implica na seleção de variáveis relacionadas a sistemas de previsão do vento, solar e de marés [11]. Para este trabalho optou-se pelo emprego de um método que envolveu o coeficiente de correlação de Pearson com as seguintes etapas: 1) cálculo da correlação de Pearson entre todas as variáveis de entrada disponíveis e o vetor de geração eólica; 2) retenção apenas das variáveis com correlação em módulo maior que 0,20 com a geração; 3) cálculo da correlação entre as variáveis retidas na etapa 2 e eliminação de uma das variáveis de conjuntos que apresentassem correlação maior que 0,99 entre si. Esse procedimento foi executado para o banco de dados da soma de geração do parque e o aerogerador equivalente.

2.5 Redes Neurais Artificiais

As redes neurais artificiais (RNA) fazem parte do ramo de aprendizado de máquina (*machine learning*) dentro da inteligência artificial, ou seja, as RNAs aprendem com os dados, por meio da etapa de treinamento, a resolver problemas. Elas foram inspiradas pelas redes neurais biológicas, ou seja, no sistema nervoso central de um cérebro animal. De forma simplista pode-se dizer que uma RNA é composta por neurônios, camadas e funções de ativação. Há neurônios nas camadas de entrada, escondidas e de saída. Os neurônios da camada de entrada transmitem a informação para dentro da rede neural, os neurônios das camadas escondidas

processam essa informação por meio de “sinapses” e os neurônios da camada de saída indicam o resultado desse processo.

Para entender melhor RNA é necessário saber como funciona sua unidade básica, a qual é representada pelo neurônio artificial. Este multiplica os dados de entrada que recebe pelos valores atribuídos aos pesos para cada uma dessas conexões. Em seguida, esses resultados são somados em conjunto com um valor de bias (geralmente de valor igual a 1) e essa soma vai para a função de ativação que passa seu resultado adiante [12]. Devido a essa característica de passar adiante, há as redes neurais *feedforward*, em que a informação para a previsão segue o fluxo da camada de entrada, depois para as camadas escondidas e posteriormente para a camada de saída. Os pesos dos neurônios são ajustados durante o treinamento por meio do processo de *backpropagation*. Esse processo consiste em avaliar a contribuição de cada neurônio para o erro entre o que a rede neural previu e o valor verdadeiro para, posteriormente, ajustar os pesos e melhorar a saída da RNA.

As camadas podem estar totalmente conectadas ou não. Uma RNA com as camadas escondidas totalmente conectadas é chamada de Multicamada Perceptron (MLP) [13]. A Figura 3a ilustra um esquema de uma RNA. As redes neurais podem ser superficiais (SNN), com apenas uma camada escondida, ou profundas (DNN), com duas ou mais camadas escondidas.

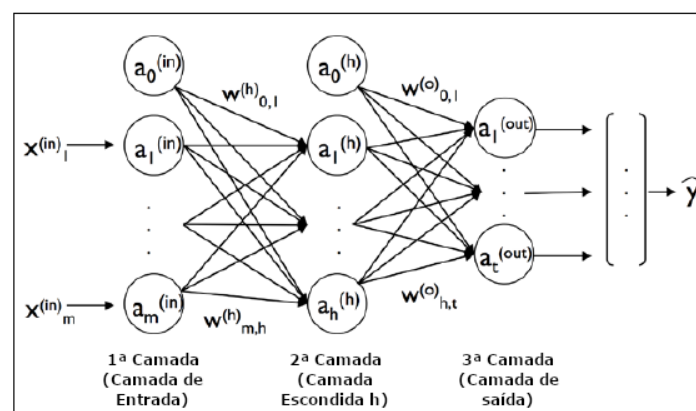


Figura 3 – (a) Esquema de uma rede neural artificial MLP. Adaptado de [13].

2.6 Avaliação do Desempenho da Rede Neural

O erro do modelo de previsão é classicamente definido como a diferença entre o valor medido e o predito. No presente trabalho avaliou-se o modelo pela raiz do erro quadrático médio (RMSE),

raiz do erro quadrático médio normalizado (NRMSE) e validação cruzada (*cross-validation*) de série temporal. A normalização do NRMSE é calculada em função da potência do parque ou do aerogerador equivalente. A RNA é treinada com um conjunto de dados e testada com outro conjunto de dados independente. Validação cruzada de série temporal é uma forma de testar a RNA diversas vezes. Isso é feito por meio da divisão do banco de dados total em diversos bancos de dados de treinamento e teste [14]. Avalia-se o modelo de previsão em cada um desses conjuntos e o desempenho resultante do modelo é a média dos desempenhos dos conjuntos. Com isso é possível calcular o desvio padrão do erro da avaliação. A Figura 4 ilustra o procedimento de validação cruzada de série temporal adotada neste trabalho, também chamado, em inglês, de *walk forward validation*.

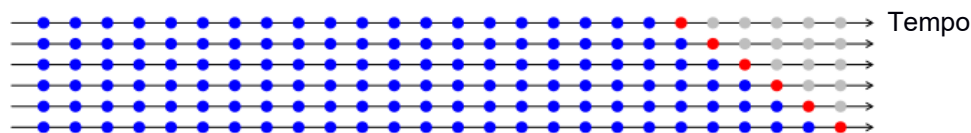


Figura 4 – Representação da validação cruzada de uma série temporal. Adaptado de [14].

2.7 Procedimento

O trabalho consiste em utilizar as variáveis obtidas na seleção de variáveis e testar diferentes modelos com auxílio de validação cruzada para o banco de dados da soma total de geração do parque e do aerogerador equivalente. O período de dados utilizado para este trabalho foi de 21/12/2017, à zero hora, até o final do dia 30/03/2018, o que totaliza 100 dias. Deste banco de dados inicializou-se a validação cruzada da série temporal com os 30 primeiros dias e posteriormente incrementou-se de 5 em 5 dias, de 120 em 120 horas. O primeiro banco de dados de treinamento possui 30 dias e o de teste possui 5 dias, posteriormente esse banco de dados de teste é adicionado ao de treinamento, isto resulta em 35 dias de treinamento e 5 dias de teste, o que se repete continuamente até chegar ao fim do banco de dados total. Essa abordagem permite uma melhor avaliação do modelo de previsão e o cálculo do desvio padrão do RMSE entre os diferentes testes.

A linguagem livre de programação *Python* e a biblioteca de inteligência artificial *Keras* foram utilizadas na execução deste trabalho. Os dados de entrada, antes de entrarem na RNA, passam por uma padronização por meio da obtenção de Z. Foram testadas redes neurais MLP

com uma, duas e três camadas escondidas. Os valores dos hiperparâmetros para cada topologia de RNA encontram-se na Tabela 1.

Tabela 1 - Hiperparâmetros dos modelos de previsão de geração a serem testados

Hiperparâmetros	MLP 1 camada escondida	MLP 2 camada escondidas	MLP 3 camada escondidas
Nº de neurônios por camada escondida	25, 50, 100	25, 50	
Função de Ativação das camadas escondidas	Sigmoid	1ª – Sigmoid 2ª - ReLU	1ª – Sigmoid 2ª e 3ª– ReLU
Função de ativação para camada de saída	Linear	Linear	Linear
Otimizador	Adam	Adam	Adam
Taxa de aprendizagem	0,001	0,0001	0,0001
Batch Size	1	16	16
Nº de épocas	500	1000	1000

3. RESULTADOS

A seguir são apresentados os erros e previsões decorrentes dos testes com as diferentes topologias e hiperparâmetros das RNAs, informações presentes na Tabela 1, para cada um dos procedimentos de previsão de geração do parque. Os valores de geração presentes nos gráficos indicam a predição após cada período de treinamento, ou seja, o período de teste de cada conjunto de treinamento da validação cruzada. Por exemplo, os 5 primeiros dias são para a RNA treinada com 30 dias, já entre o quinto e décimo dia são os resultados para a RNA treinada com 35 dias, e assim por diante. O desvio padrão é abreviado pela letra σ .

3.1 Avaliações com o banco de dados da soma da geração do parque

O valor médio do RMSE e NRMSE, assim como seu desvio padrão, obtidos por meio da validação cruzada da série temporal são mostrados na Tabela 2 para cada topologia de RNA e o banco de dados da soma de geração do parque. O número de variáveis de entrada selecionadas foi de 14, conforme procedimento descrito na seção de procedimento de cálculo. O vetor disponibilidade, rajadas de vento e módulo de velocidade a 100 m de altura estão entre essas variáveis.

Tabela 2 - Resultados para o banco de dados com a soma de geração do parque

Modelo	Nº de camadas escondidas	Nº de neurônios por camada escondida	Código	RMSE [kW]	σ RMSE [kW]	NRMSE [%]	σ NRMSE [%]
MLP	1	25	MLP1-25	7230	1538	24,1	5,1
		50	MLP1-50	6854	1610	22,8	5,4
		100	MLP1-100	6871	1793	22,9	6,0
	2	25	MLP2-25	6771	1529	22,6	5,1
		50	MLP2-50	6833	1680	22,8	5,6
	3	25	MLP3-25	7182	1947	23,9	6,5
		50	MLP3-50	7194	1987	24,0	6,6

Para a configuração MLP com uma camada escondida, nota-se que a melhor configuração é com 50 neurônios, pois apresenta na validação cruzada um RMSE de 6854 kW com desvio padrão (σ) de 1610 kW, que normalizado em relação à capacidade instalada no parque equivale a um valor de NRMSE de 22,8% com desvio padrão de 5,4%. Na MLP com duas camadas escondidas, a RNA com 25 neurônios apresentou os menores valores de RMSE e desvio padrão em relação ao modelo com 50 neurônios em cada camada. A RNA com 25 neurônios em cada uma de suas duas camadas escondidas obteve um RMSE de 6771 kW e desvio padrão de 1529 kW, o que representa um NRMSE de 22,6% e desvio padrão de 5,1%. Os modelos de Multicamada Percetron com três camadas escondidas tiveram um resultado similar entre si, porém o que apresentou menor erro foi o modelo com 25 neurônios em cada camada. Este modelo obteve um NRMSE de 23,9% e desvio padrão de 6,5%. Nota-se que o aumento do número de neurônios para a MLP não foi determinante para o melhor desempenho da RNA que prevê a soma da geração. Outro fato a se destacar é que a RNA com maior número de camadas não foi a que se teve o melhor desempenho. Dentre os modelos de RNA para a previsão direta da geração média total do parque em resolução temporal de 30 min, o que melhor desempenhou durante a validação cruzada foi o MLP com duas camadas escondidas e 25 neurônios. Este modelo apresentou o menor valor de NRMSE e o menor desvio padrão.

A Figura 5 mostra os resultados de RMSE e NRMSE durante a validação cruzada para cada tamanho de banco de dados de treinamento, lembrando que o tamanho de banco de dados de teste era fixo em 5 dias (120 h). Os menores valores de RMSE acontecem com o banco de dados de treinamento composto por 45 dias.

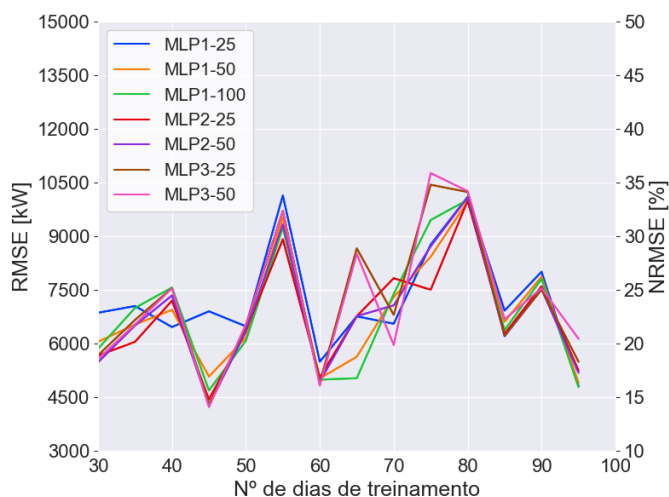


Figura 5 – RMSE e NRMSE da validação cruzada com a soma de geração do parque.

A geração predita pelas redes neurais treinadas diretamente com a soma de geração do parque encontra-se juntamente com a geração real do parque na Figura 6. Nota-se que a MLP com uma camada escondida e 25 neurônios não apresenta bom resultado ao longo da série temporal. Os demais modelos acompanharam melhor o comportamento da geração. Entretanto, nenhum dos modelos de previsão de geração operou bem com 65 dias de treinamento. Tal comportamento durante o teste, entre 24/02/2018 às 00:00 e 01/03/2018 às 00:00, foi decorrente de problema com uma das variáveis de entrada, tema que será abordado na seção 3.4.



Figura 6 – Geração do parque real e prevista utilizando a soma de geração do parque.

3.2 Avaliações com o banco de dados do aerogerador equivalente

As RNAs para o banco de dados do aerogerador equivalente tiveram como dados de entrada 10 variáveis, escolhidas conforme o procedimento de seleção de variáveis. A Tabela 3 apresenta os valores de RMSE, desvio padrão do RMSE, NRMSE e desvio padrão do NRMSE para as RNAs que preveem a potência gerada para o aerogerador equivalente.

Tabela 3 - Resultados para o banco de dados com o aerogerador equivalente

Modelo	Nº de camadas escondidas	Nº de neurônios por camada escondida	Código	RMSE [kW]	σ RMSE [kW]	NRMSE [%]	σ NRMSE [%]
MLP	1	25	MLP1-25	777	187	25,9	6,2
		50	MLP1-50	791	187	26,4	6,2
		100	MLP1-100	806	185	26,9	6,2
	2	25	MLP2-25	774	183	25,8	6,1
		50	MLP2-50	784	192	26,1	6,4
	3	25	MLP3-25	795	196	26,5	6,5
		50	MLP3-50	879	271	29,3	9,0

O modelo de previsão MLP com uma camada escondida com menor erro foi o de 25 neurônios com RMSE de 777 kW e desvio padrão (σ) de 187 kW, o que representa um NRMSE de 25,9% e desvio padrão de 6,1%. A MLP com duas camadas e 25 neurônios obteve NRMSE de 25,8% e desvio padrão de 6,1%, o que é melhor do que o obtido pela RNA com 50 neurônios em cada uma das duas camadas escondidas. O NRMSE de 26,5% e desvio padrão de 6,5% foi o resultado da MLP de 3 camadas escondidas com 25 neurônios. A RNA com três camadas escondidas e 50 neurônios, em cada, obteve o pior resultado dentre as demais. O aumento do número de camadas e neurônios não representou uma melhor performance para as redes neurais MLP. O melhor modelo em relação ao RMSE médio resultante da validação cruzada e o desvio padrão foi o Multicamada Perceptron com duas camadas escondidas e 25 neurônios

Os valores de RMSE e NRMSE para cada treinamento durante a validação cruzada das redes neurais que preveem diretamente a potência do aerogerador equivalente encontram-se na Figura 7. Novamente o menor erro aconteceu com o banco de treinamento de 45 dias.

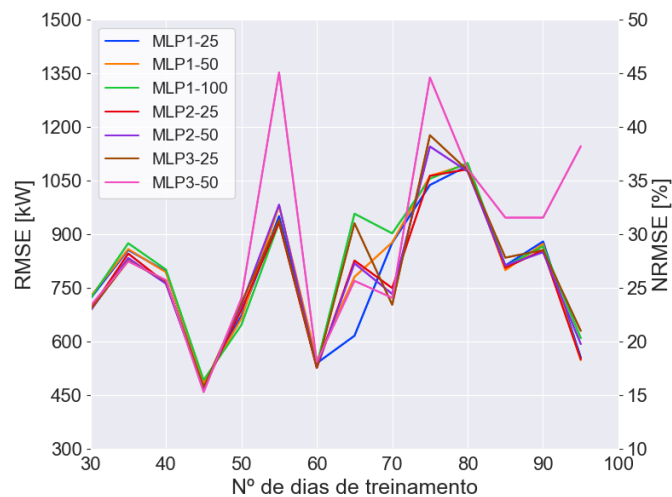


Figura 7 – RMSE e NRMSE da validação cruzada com aerogerador equivalente.

A Figura 8 apresenta a previsão de geração para o aerogerador equivalente e a previsão real para os testes durante a validação cruzada. O modelo MLP com uma camada escondida e 50 neurônios apresenta valores negativos de geração e maiores que 3000 kW médios em 30 min na série temporal prevista, o que não condiz com a geração real. Nenhum modelo conseguiu prever bem a geração do aerogerador equivalente entre 24/02/2018 às 00:00 e 01/03/2018 às 00:00, pelo mesmo motivo apresentado na seção anterior, visto que o mesmo comportamento ocorre.

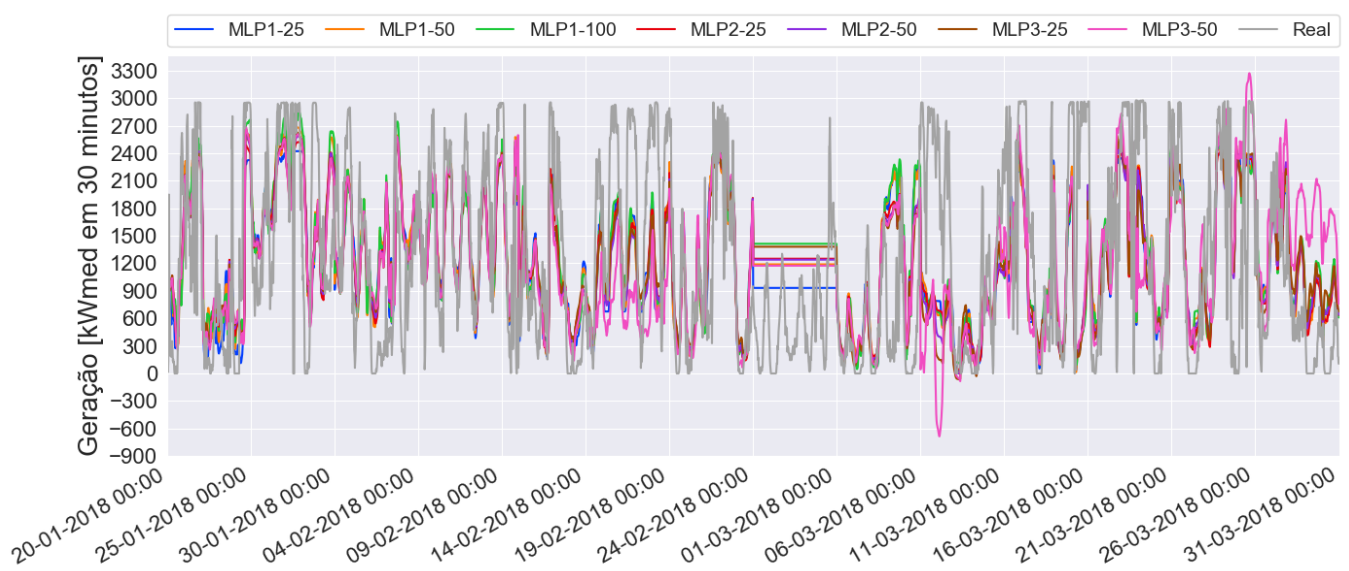


Figura 8 – Geração real e prevista do aerogerador equivalente.

A geração total do parque é obtida pela multiplicação da energia do aerogerador equivalente pelo número de turbinas funcionando no parque no momento. O RMSE, NRMSE e os desvios padrões desses resultados em relação à geração real do parque são apresentados na Tabela 4. Nota-se que as previsões por esse método apresentam RMSE médios entre 22,2% e 25,3% e desvio padrão entre 6,0% e 9,1%. A MLP com duas camadas escondidas e 25 neurônios em cada obteve o melhor resultado, o qual foi de 6669 kW de RMSE e desvio padrão de 1807 kW, que representa um NRMSE de 22,2% e desvio padrão de 6,0%.

Tabela 4 - RMSE e NRMSE para geração total do parque com o aerogerador equivalente

Modelo	Nº de camadas escondidas	Nº de neurônios por camada escondida	Código	RMSE [MW]	σ RMSE [MW]	NRMSE [%]	σ NRMSE [%]
MLP	1	25	MLP1-25	6687	1887	22,3	6,3
		50	MLP1-50	6822	1881	22,7	6,3
		100	MLP1-100	6942	1841	23,1	6,1
	2	25	MLP2-25	6669	1807	22,2	6,0
		50	MLP2-50	6745	1937	22,5	6,5
	3	25	MLP3-25	6822	1971	22,7	6,6
		50	MLP3-50	7581	2726	25,3	9,1

O RMSE e NRMSE para a geração total do parque obtida por meio das previsões do aerogerador equivalente para cada etapa da validação cruzada são mostrados na Figura 9.

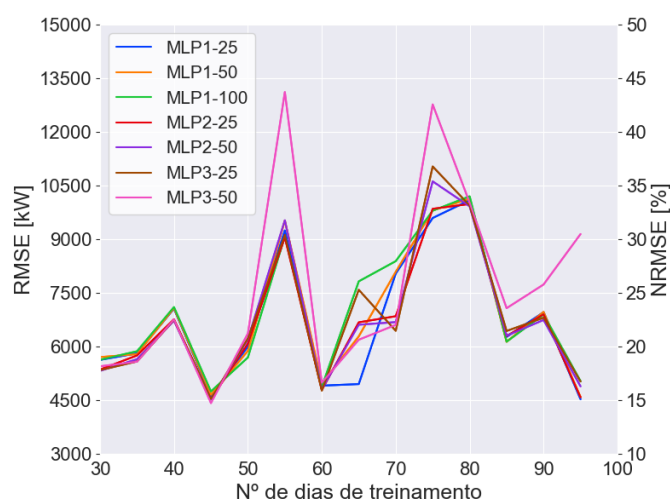


Figura 9 – RMSE e NRMSE da geração total do parque com aerogerador equivalente.

A Figura 10 mostra a geração total do parque obtida pela multiplicação da previsão de geração do aerogerador equivalente pelo número de aerogeradores em funcionamento e a geração real. Observa-se que entre os dias 24/02/2018 e 01/03/2018 os resultados acompanharam de certa forma a geração real, isso ocorre pela multiplicação com os aerogeradores em funcionamento. Porém erros nos dados de entrada, GFS, podem ter prejudicado essa etapa da validação.

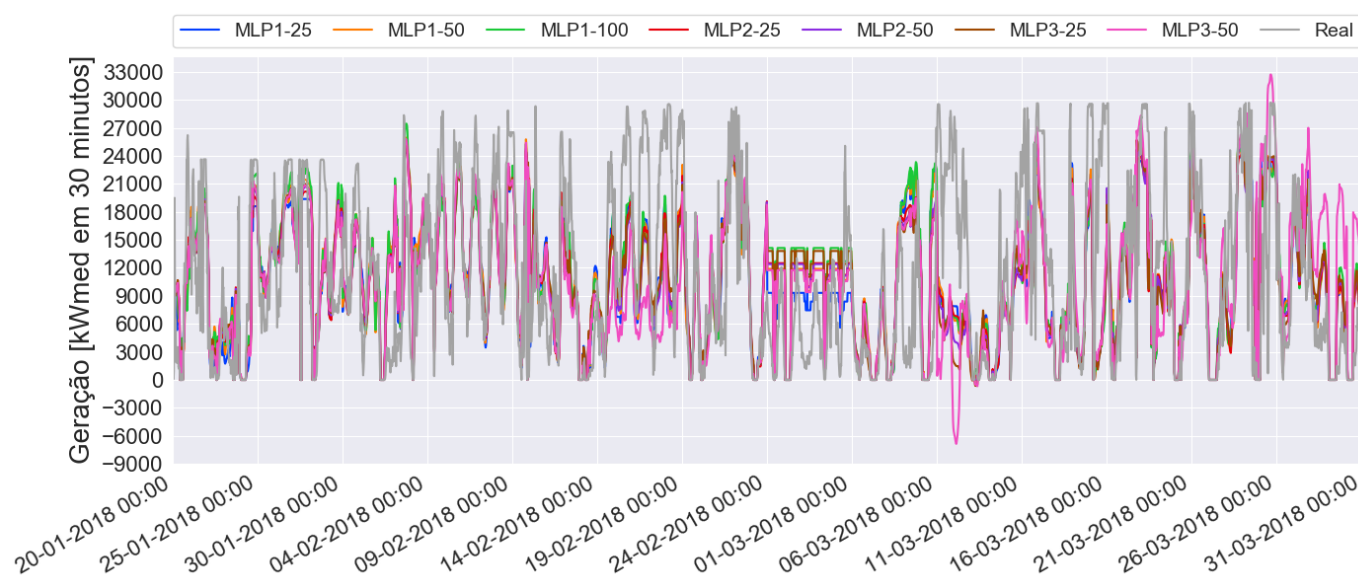


Figura 10 – Geração real e predita do parque utilizando o aerogerador equivalente.

3.3 Comparação

Analisando a seção 3.1, chega-se ao modelo Multicamada Perceptron com duas camadas escondidas e 25 neurônios como o que apresentou melhor performance dentre as RNAs que previam diretamente a soma de geração do parque, sendo chamado agora de MLP2-25-SomaParq. Por meio dos dados apresentados na seção 3.2, nota-se que o melhor modelo para prever a geração do parque, pela multiplicação dos aerogeradores em funcionamento pela geração do aerogerador equivalente, foi também a MLP com duas camadas escondidas e 25 neurônios, sendo denominada MLP2-25-AeroEq. A Figura 11 apresenta o RMSE e NRMSE para essas duas abordagens de previsão da geração do parque ao longo da validação cruzada. O RMSE da abordagem de previsão da geração do parque pelo aerogerador equivalente foi superior apenas no teste da RNA com período de 75 dias de treinamento. Da Tabela 2 e 4 tiramos que a MLP2-25-SomaParq obteve um RMSE médio de 6771 kW e desvio padrão de

1529 kW, já a MLP2-25-AeroEq obteve RMSE de 6669 kW e desvio padrão de 1807 kW. O RMSE médio da validação cruzada é menor para MLP2-25-AeroEq, porém seu desvio padrão é maior. As previsões realizadas durante os testes da validação cruzada estão mostradas na Figura 12. Observa-se que a MLP2-25-AeroEq acompanha melhor o comportamento da geração real do parque ao longo da série temporal.

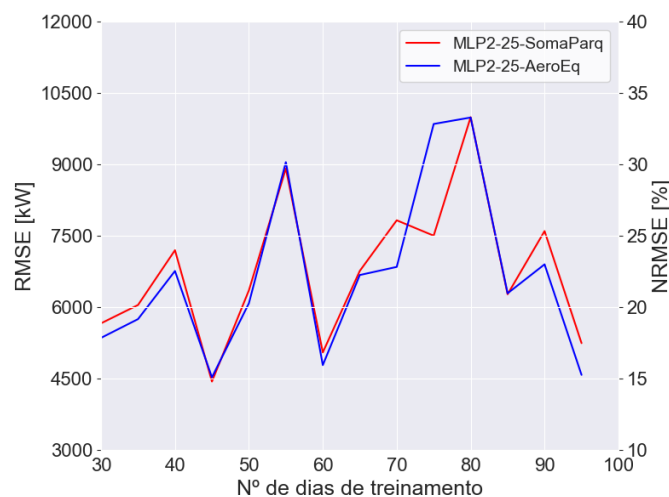


Figura 11 – RMSE e NRMSE da geração total do parque com os melhores modelos.

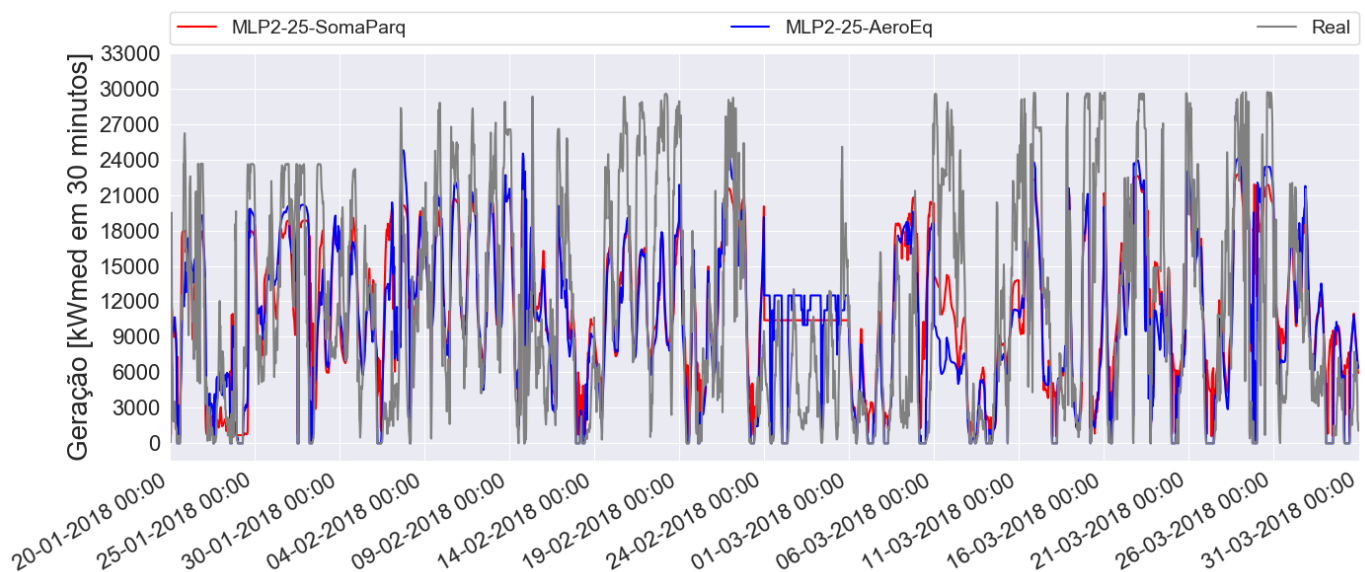


Figura 12 – Geração real e predita do parque utilizando os melhores modelos.

3.4 Análise do Erro de Previsão entre 24/02/2018 e 01/03/2018

Conforme observado nas Figuras 8, 10 e 12, não houve um bom desempenho das previsões entre 24/02/2018 às 00:00 e 01/03/2018 às 00:00. Este período de teste corresponde às RNAs treinadas com 65 dias. A anomalia na previsão foi causada pela variação abrupta (equivalente à função degrau) de uma das variáveis de entrada no período de testes. A Figura 13 mostra os resultados das previsões de geração do parque para as RNAs com e sem a variável causadora da anomalia. Na legenda, Proc 1, Proc 2, Com Var e Sem Var indicam, respectivamente, primeiro procedimento, segundo procedimento, com e sem a variável causadora da anomalia. Com essa modificação, no conjunto das variáveis de entrada, os Proc 1 e Proc 2 possibilitaram previsões mais próximas da geração do parque. Como se pode observar, o Proc 2 está mais próximo do real, particularmente no intervalo de 12 h, entre 0 e 12 h do dia 25/02. Os valores de RMSE, para os Proc 1 e Proc 2, no período das 120 h, sem a variável causadora da anomalia, foi de 2882 e 3353 kW (médias em 30 min), respectivamente.

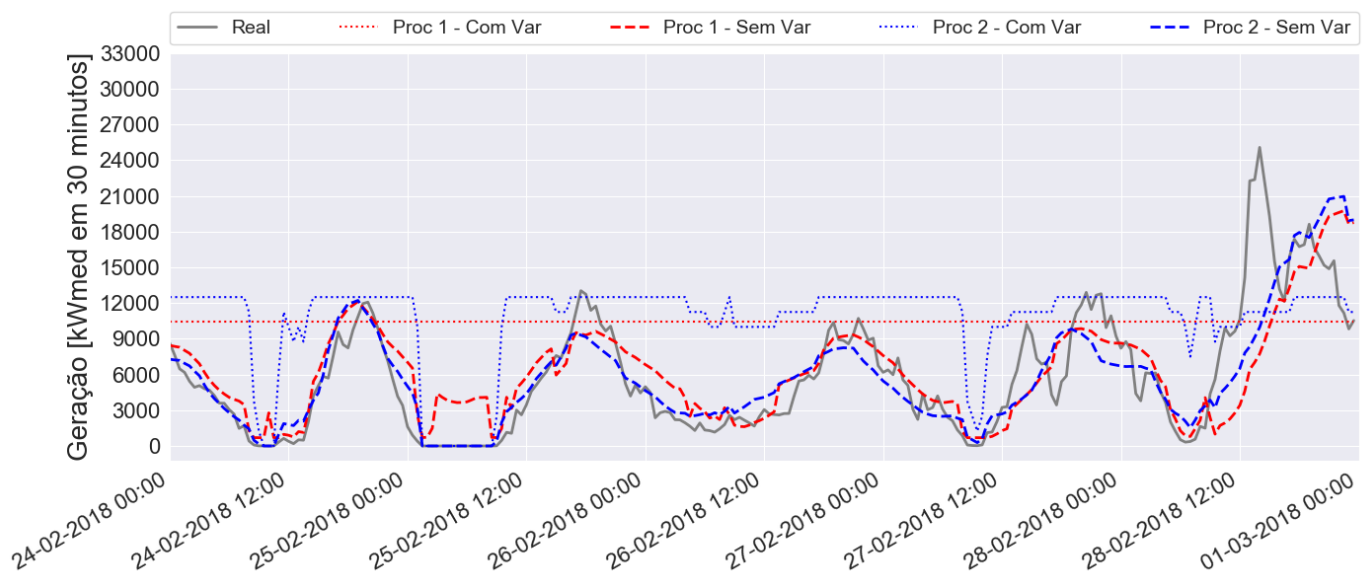


Figura 13 – Geração do parque, real e prevista, após a retirada da variável de entrada causadora das anomalias nas previsões entre 24/02/2018 às 00:00 e 01/03/2018 às 00:00.



4. CONCLUSÃO

A previsão de geração do parque eólico Aura Mirim II, localizado em Santa Vitória do Palmar, no Rio Grande do Sul, por meio da técnica de redes neurais artificiais, com médias de 30 min e horizonte de previsão de 120 horas, foi feita utilizando dois procedimentos: o primeiro utiliza o banco de dados de geração com a soma da geração do parque e o segundo calcula a geração do parque a partir do que foi gerado por um aerogerador equivalente multiplicado pelo número de aerogeradores em funcionamento. Por meio de uma validação cruzada, o segundo procedimento foi o que apresentou o menor valor de NRMSE, 22,2%, com um desvio padrão de 6,0%, enquanto que a abordagem pela previsão da soma de geração do parque como saída direta da RNA obteve NRMSE de 22,6 % e desvio padrão de 5,1 %. Além de um menor NRMSE médio na validação cruzada, a abordagem com o aerogerador equivalente acompanhou melhor o comportamento real do parque. A melhor topologia de rede neural, para ambos os procedimentos de previsão de geração do parque, foi a Multicamada Perceptron com duas camadas escondidas e 25 neurônios em cada camada. Outro fator a se destacar é que a utilização do aerogerador equivalente facilita a construção de um modelo operacional de previsão com um horizonte de 120 horas, pois prescinde da necessidade do vetor indisponibilidade, como variável de entrada da rede neural. Este vetor não é facilmente definido com antecedência para um cenário real. A pesquisa prevê a análise de outros modelos como dados de entrada das redes neurais, podendo-se citar previsões de outros modelos globais (GDPS) e modelos de mesoescala (WRF). O uso de diversos modelos como dados de entrada para a RNA reduz a dependência de um único modelo, ou seja, o erro de um modelo pode ser compensado pelo acerto do outro.

Outro ponto a ser estudado é a quantidade ótima de dias de treinamento para as RNAs. Para o parque analisado, os menores RMSE ocorreram com 45 dias de treinamento.

AGRADECIMENTOS

Os autores agradecem o apoio do CNPq, POSMEC-UFSC e Atlantic Energia Renováveis S.A.



REFERÊNCIAS

- [1] Global Wind Energy Council. 2018. Global Wind Report: Annual Market Update 2017.
- [2] Agência Nacional de Energia Elétrica (ANEEL). 2019. Banco de Informação de Geração (BIG). Disponível em: <<http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/capacidadebrasil.cfm>> Acesso em: 14/02/2019.
- [3] Ministério de Minas e Energia (MME), Empresa de Pesquisa Energética (EPE). 2018. Plano Decenal de Expansão de Energia 2027.
- [4] SHEN, Z.; RITTR, M. 2016. Forecasting volatility of wind power production. Applied Energy.
- [5] JUNG, J., BROADWATER, R. P. 2014. Current status and future advances for wind speed and power forecasting. Renewable and Sustainable Energy Reviews.
- [6] Operador Nacional do Sistema Elétrico (ONS). 2016. Submódulo 8.1 Programação diária da operação eletroenergética.
- [7] MONTEIRO, C. et al. 2009. Wind Power Forecasting: State-of-the-Art 2009. Argonne National Laboratory.
- [8] CHANG, W.-Y. 2014. A Literature Review of Wind Forecasting Methods. Journal of Power and Energy Engineering.
- [9] GIEBEL, G., KARIONIOTAKIS, G. N., DRAXL, C. 2011. The State-of-the-Art in Short-Term Prediction of Wind Power A Literature Overview, 2nd Edition. ANEMOS.
- [10] ARTIPOLI, G., DURANTE, F. 2014. Physical Modeling in Wind Energy Forecasting. DEWIN Magazin.
- [11] Salcedo-Sanz, S., Cornejo-Bueno, L., Prieto, L., Paredes, D., García-Herrera, R. 2018. Feature selection in machine learning prediction systems for renewable energy applications. Renewable and Sustainable Energy Reviews.
- [12] Albon, C. 2018. Machine Learning with Python Cookbook, 1st Edition. O'Reilly Media, Inc.
- [13] Raschka, S., Mirjalili, V. 2017. Python Machine Learning, 2nd Edition. Packt Publishing.



[14] Hyndman, R. Athanasopoulos, G. 2018. Forecasting: Principles and Practice, 2nd Edition. Otexts. Disponível em: <<https://otexts.com/fpp2/accuracy.html>>. Acessado em: 10/02/2019.

BIOGRAFIA

João Gabriel de Farias – Nascido em Palhoça-SC, em 21 de setembro de 1991. Estudante do mestrado em Engenharia Mecânica na UFSC, universidade na qual se graduou Engenheiro Mecânico em 2016. Pelo programa Ciência Sem Fronteiras estudou por um ano na *The Pennsylvania State University*. Tem interesse em energias renováveis e inteligência artificial.

Júlio César Passos – É pesquisador do CNPq e professor titular no Departamento de Engenharia Mecânica e do Programa de Pós-Graduação em Engenharia Mecânica na UFSC. Ebulição, condensação e energias renováveis são seus interesses, na área de pesquisa e desenvolvimento.

Reinaldo Haas – Natural de Vidal Ramos-SC. É professor do Depto de Física da UFSC. Possui interesse em modelagem numérica da atmosfera, climatologia e energias renováveis

Eduardo Victor Dias – Engenheiro Mecânico pela UFSC (2008) e representante da Atlantic Energias Renováveis S.A.

Mayara Miqueletti de Lima – Engenheira Industrial Elétrica pela UTFPR (2014) e representante da empresa Atlantic Energias Renováveis S.A.