

Previsão de Geração de Energia Fotovoltaica Utilizando Redes Neurais Artificiais

Roan R. Brolese*. André B. Michel*

*Centro de Ciências Exatas, da Natureza e de Tecnologia,
Universidade de Caxias do Sul (UCS), Caxias do Sul, RS, Brasil
(e-mails: rrbrolese@ucs.br, abmichel@ucs.br)

Abstract: This work has the objective of developing a photovoltaic power generation forecasting in the city of Garibaldi, Rio Grande do Sul, with a time horizon of six months, using an artificial neural network. The input data of this network is composed of climatic historical data sets in addition to a set of historical solar photovoltaic generation data, which will be used as targets. The input variable classification is performed using a stepwise regression, which lists the highest relation between input variables and the objective. The modeling is accomplished by the numerical software Matlab. To train the artificial neural network the supervised learning method of error backpropagation was used. This method applies the training function knowing as bayesian regularization, aiming to reduce the forecast error. The artificial neural network reached a 12.97% MAPE. The comparisons of neural network performance for different time horizons are also presented in this context. The results achieved are in agreement with those found in the bibliography.

Resumo: O presente trabalho tem o objetivo de realizar uma previsão de geração fotovoltaica para a cidade de Garibaldi, Rio Grande do Sul, para um horizonte de seis meses a frente, por meio da modelagem de uma rede neural artificial. Os dados de entrada desta rede são compostos por conjuntos de dados históricos climáticos em complemento a um conjunto de dados históricos de geração solar fotovoltaica que serão usados como alvos. A classificação das variáveis de entrada é realizada por meio da regressão *stepwise*, que elenca as variáveis com maior relação com a resposta. A modelagem conta com o auxílio do software numérico Matlab, para a construção do algoritmo e a classificação dos dados. Para o treinamento desta rede neural artificial é empregado o método de aprendizado supervisionado de retropropagação do erro, utilizando a função de treinamento *bayesian regularization*, visando reduzir o erro da previsão. A rede neural artificial com melhor desempenho atingiu um MAPE de 12,97 %. Também neste contexto são apresentadas as comparações de desempenho das redes neurais para diferentes horizontes, com resultados alcançados estando de acordo com os encontrados na literatura.

Keywords: Photovoltaic power generation, Forecasting, Artificial neural network, Machine learning.

Palavras-chaves: Geração fotovoltaica, Previsão, Rede neural artificial, Aprendizado de máquina.

1. INTRODUÇÃO

A geração de energia elétrica por meio da energia solar vem crescendo globalmente. Este crescimento se deve em grande parte pela busca por energias renováveis em substituição ao uso de combustíveis fósseis para a geração elétrica. (ALZHRANI et al., 2017). Segundo Sansa et al., (2014), o aumento vertiginoso na utilização de fontes de energias renováveis no mundo acompanha o crescimento da economia juntamente com o aumento da demanda.

De acordo com Chaturvedi e Isha (2016), até 2050 a demanda de energia elétrica no mundo será maior que o dobro da demanda atual. Em virtude desta previsão é fundamental a ampliação de geração de energia elétrica por meio de fontes renováveis. Em 2012 por meio da norma regulamentadora número 482 a ANEEL, Agência Nacional de Energia Elétrica, elaborou critérios para a integração de sistemas de micro e minigeração distribuída. (ANEEL, 2012).

A realização de uma previsão precisa de geração pode auxiliar no controle de operações e nas manutenções relacionadas a sistemas com fontes de geração fotovoltaicas a ele conectados. Majumder, Behera e Nayak (2017), também relatam a importância das previsões para projetos de novas usinas, analisando dados climáticos de uma determinada região, sendo assim possível planejar a inserção destas novas usinas através de estimativas prévias. Fatores climáticos como temperatura ambiente, umidade relativa, radiação solar e nebulosidade são alguns exemplos de condições influentes na geração fotovoltaica. Estes dados podem ser analisados por meio de métodos de aprendizado de máquina, que procuram identificar padrões entre os dados analisados no intuito de aprender com estes dados. (ABEDIN et al., 2017). Para Liu e Zhang (2016) há inúmeras previsões realizadas utilizando estes métodos de aprendizado, sendo os modelos mais utilizados os auto-regressivos, como o ARMA, *Auto Regressive Moving Average*, as redes neurais artificiais e as máquinas de vetores de suporte.

2. PROPOSTA

Em vista das questões levantadas na seção anterior, o propósito deste trabalho é realizar a previsão de geração fotovoltaica de longo prazo, na região de Garibaldi, Rio Grande do Sul, para os próximos 6 meses, com um horizonte de aproximadamente 184 dias a frente utilizando dados históricos climáticos adquiridos em bancos de dados especializados e de geração fotovoltaica adquiridos no local. O fator primordial para precisão em uma previsão está na qualidade dos dados históricos obtidos. O período em que determinado parâmetro é monitorado também tem influência na previsão. (CHATURVEDI E ISHA, 2016). Desta forma, este estudo visou a aquisição de dados com o maior período de monitoramento na região. Devido a quantidade de estudos relacionados a previsão de energia que utilizam redes neurais artificiais para suas previsões, aliada a capacidade das redes neurais de lidar com diferentes tipos de dados e aspectos não lineares. (ABEDIN et al., 2017). Este algoritmo será o utilizado nesta previsão.

3. METODOLOGIA

3.1 Métricas de Verificação

Na realização de qualquer previsão é fundamental avaliar a qualidade dos resultados obtidos, para isso, existem técnicas que são executadas para a análise destes resultados que calculam o erro entre os valores previstos e os valores atuais. Entre as técnicas utilizadas para esta análise se destacam o MAPE, erro percentual absoluto médio, e o RSME, erro da raiz média quadrática. (LAHOUAR, MEJRI E SLAMA, 2017). O MAPE está representado em (1) e o RMSE, definido em (2). Em Alzahrani (2017), outra técnica utilizada é o erro da média quadrática, MSE, expresso em (3).

$$MAPE = 100 \times \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (1)$$

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

$$MSE = \frac{1}{n} \times \sum_{i=1}^n (y_i' - y_i)^2 \quad (3)$$

Onde:

- \hat{y}_i é o valor de saída previsto;
- y_i é o valor utilizado com alvo;
- n é o número de observações.

3.2 Redes Neurais

a) Estrutura

O principal elemento de uma rede neural artificial é o neurônio. Os neurônios são combinadores lineares que interagem com os dados de entrada, uma vez que a ponderação desses dados é definida pelos seus pesos sinápticos. (ABEDIN et al., 2017). As RNAs, redes neurais artificiais, são compostas por camadas, sendo que cada uma dessas camadas contém um número de neurônios. Na primeira camada, conhecida como camada de entrada, são alocados os dados de entrada que serão utilizados para o aprendizado do algoritmo, nesta camada cada neurônio é responsável por um tipo de dado, ou seja, se para uma previsão são utilizados dois tipos de dados, na camada de entrada existirá apenas dois neurônios. Na segunda camada, chamada de camada oculta, acontece o tratamento dos dados, essa camada é composta pelos neurônios ocultos.

Para a definição da quantidade ótima de neurônios a serem inseridos na camada oculta de uma rede neural artificial, existem diferentes abordagens. Este trabalho utilizou a técnica de determinação da quantidade de neurônios ocultos a partir do teorema de Kolmogorov, representada em (4) que relaciona a quantidade de neurônios ocultos com o número de neurônios de entradas e de saída. (SILVA, 2005).

$$N_{ocultos} = 2 \times N_{entradas} \quad (4)$$

Onde:

- $N_{ocultos}$ é o número de neurônios ocultos;
- $N_{entradas}$ é o número de neurônios de entrada.

b) Pré-processamento de Dados

Durante a etapa de pré-processamento dos dados, as variáveis preditoras passam por um processo de classificação. O método de classificação de variáveis utilizado neste trabalho foi o *stepwise regression*, que busca dentre o conjunto total, as variáveis com maior importância estatística. Este processo busca avaliar cada uma das possíveis variáveis de entrada de um modelo de previsão, neste caso, as variáveis de entrada de uma rede neural, afim de eliminar variáveis redundantes ou que possam prejudicar a eficiência da previsão. (GHUNEM; ASSALEH; EL-HAG, 2012).

c) Função de Ativação

Os neurônios são somadores ponderados pelos pesos sinápticos. Os resultados desta soma ponderada são aplicados a uma função de ativação. Estas funções são normalmente funções não-lineares e são responsáveis pela resposta do neurônio de saída. (SILVA, 2005). A função de ativação utilizada na modelagem da rede neural foi a função sigmoide, caracterizada pela sua curva em S. A função sigmoide tangente hiperbólica assume valores contínuos entre -1 e 1, representada em (5). (HAYKIN, 2001).

$$\varphi = \tanh(v) \quad (5)$$

Onde:

- V é o valor que será ativado pela função.

d) Função de Treinamento

A função de treinamento é a responsável pelo aprendizado da rede. No modelo construído, a função de treinamento a ser utilizada será a *bayesian regularization*, que permite atualizar os pesos sinápticos ao longo do treinamento garantindo uma generalização da rede neural e prevenindo que a ocorrência de *overfitting*. O objetivo de uma rede neural é produzir o menor erro possível no treino e no teste, procurando manter o mesmo erro em ambos. O algoritmo desta função de treinamento foi desenvolvido afim de reduzir a soma do erro quadrático da saída da rede em relação ao seu alvo. A técnica de regularização entre a soma destes erros e a soma dos erros quadráticos dos pesos sinápticos é apresentada em (6). (SAINI 2007).

$$F = \gamma SSE + (1 - \gamma) SSE_w \quad (6)$$

Onde:

- SSE é a soma do erro quadrático;
- SSE_w é a soma do erro quadrático dos pesos;
- γ é o parâmetro de desempenho que está diretamente relacionado com o ritmo do treinamento.

e) Validação Cruzada

Na técnica de Validação Cruzada o treinamento de uma rede neural é realizado com um conjunto de dados diferente do conjunto utilizado no teste. O conjunto de dados de treinamento é dividido em dois subconjuntos, sendo treinamento e validação. O objetivo deste método está em encontrar uma generalização ótima evitando o *overfitting*, quando a rede neural é excessivamente treinada, perdendo generalização. (HAYKIN, 2001).

3.3 Aquisição de dados Climáticos

A NASA, *National Aeronautics and Space Administration*, mantém um projeto de apoio a energias renováveis e afins, o *Power Data Access Viewer*, um sistema de visualização de dados solares e meteorológicos. Este acervo foi construído por meio de dados fornecidos por estações meteorológicas e por imagens de satélites, com medições realizadas desde 1981. Este sistema é disponibilizado gratuitamente afim de facilitar e proporcionar o incremento de novas fontes de geração renováveis. (STACKHOUSE et al., 2018).

A prefeitura municipal de Garibaldi disponibiliza em, Garibaldi (2019), um canal chamado Portal Clima. Este canal possui um acervo de dados meteorológicos registrados em períodos horários. A fonte citada disponibiliza dados de duas estações meteorológicas localizadas em Garibaldi, a Estação Aeroclub e a Estação São Gotardo.

3.4 Dados Históricos de Geração Fotovoltaica

Por ser um aprendizado supervisionado é importante ter conhecimento dos dados de geração para ensinar a rede neural e ajudar na validação dos resultados obtidos, nos quais são aplicadas as métricas de validação que compararam o erro entre os valores previstos e os valores reais de geração.

Para este trabalho, os dados de geração utilizados serão os dados mensurados na região de Garibaldi fornecidos por uma empresa da região que atua no ramo de geração de energia solar, com um período temporal de dois anos e meio de medição, sendo possível, até o momento da previsão, o acúmulo de três anos de dados históricos. Estes dados de geração serão utilizados como alvos durante a etapa de treinamento da rede neural. Após realizado o treinamento, os dados de geração serão empregados para a validação da previsão, em que os dados reais serão comparados com os dados previstos.

3.5 Parametrização dos Dados

A função de ativação que será utilizada é a função sigmoide tangente hiperbólica, função que é geralmente usada em modelos de redes neurais com múltiplas camadas. Neste trabalho, será realizada a modelagem de uma rede neural com três camadas, a camada de entrada, uma camada oculta e a camada de saída. Outro fator que justifica a escolha deste tipo de função é a possibilidade de ser treinada pelo algoritmo de retropropagação, por ser uma função diferenciável. (SINGH, 2015).

Para o treinamento da rede neural se faz necessário o uso de algum método de parametrização dos dados, método este responsável por ajustar os pesos sinápticos de cada dado. Neste trabalho será utilizado o algoritmo de aprendizado com retropropagação do erro. Antes do processo de treinamento e ajustes dos pesos sinápticos ser iniciado, o conjunto de dados foi dividido.

Seguindo os métodos de validação cruzada, o conjunto utilizado no treinamento será diferente dos utilizados no teste. Usualmente, cerca de 10% do conjunto total de dados é separado para a validação do modelo. Portanto, os dados restantes serão divididos em dois outros conjuntos, o primeiro para o treinamento com 70% do total e o segundo para teste, com os 20% restantes. (HAYKIN, 2001).

A Fig. 1 exibe o fluxograma do modelo de RNA proposto, sendo que após a aquisição dos dados históricos é realizado o pré-processamento dos dados. Aplicando o conceito da validação cruzada, a coleção de dados é dividida. Na sequência é modelada a RNA, configurando suas características como: a quantidade de neurônios ocultos, o método de treinamento e a função de ativação. Com o modelo pronto inicia-se o processo de treinamento da rede, na qual o algoritmo percorrerá os dados e buscará por padrões.

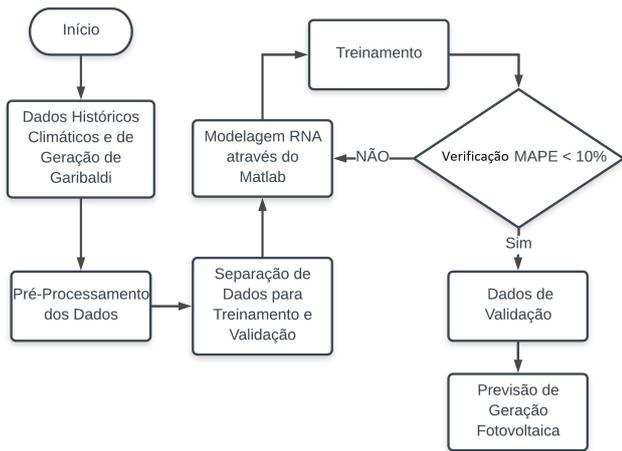


Fig. 1 Fluxograma do Modelo de RNA Proposto.

Uma vez que o treinamento realizado exibe a sua melhor performance, isto é, quando o menor MAPE é alcançado, o treinamento é encerrado. O restante da coleção de dados que não foi utilizado para o treinamento será inserida no algoritmo afim de realizar a previsão. A previsão será avaliada por meio da verificação do MAPE entre os valores previstos e valores reais na região de Garibaldi, analisando os níveis de precisão entre a previsão e os dados reais.

3.6 Modelagem das RNAs

Para a previsão serão configuradas duas redes neurais, uma utilizando os dados do *Power Data Access Viewer*, que aqui será denominada Rede Neural PDAV e outra que mescla os dados com maior relevância dentre as duas fontes de dados climáticos, denominada Rede Neural Mista. Os dados utilizados para o treinamento e validação das redes neurais pertencem ao intervalo de 1052 dias, de 20 de abril de 2016 até 06 de março de 2019.

4. RESULTADOS

Neste capítulo serão abordados os resultados obtidos conforme a metodologia proposta. Serão expostas as previsões realizadas por cada rede com seus respectivos MAPE. Também serão expostas previsões realizadas para meses de inverno e verão, afim de verificar a variação do MAPE. Ainda neste capítulo, será comparada a acuracidade de cada previsão através das métricas de avaliação apresentadas na seção 3.1, com os erros obtidos no processo de treinamento e na validação. Por fim, os resultados atingidos neste texto são comparados com os trabalhos apresentados na literatura.

4.1 Previsões Rede Neural PDAV

Antes da inserção dos dados na rede neural é realizado o pré-processamento dos dados. Para a regressão *stepwise* foram inseridas 24 variáveis candidatas. O método classificatório selecionou as variáveis que melhor modelam a saída da rede neural, sendo assim, as variáveis selecionadas foram:

Precipitação, umidade específica, faixa de temperatura a 2 metros, irradiação incidente sobre uma superfície horizontal, fluxo térmico radiativo infravermelho, irradiação de topo da atmosfera e índice de irradiação sobre céu claro.

Esta rede foi então modelada com 15 neurônios ocultos, segundo o teorema de Kolmogorov, com função de ativação sigmoide tangente hiperbólica e função de treinamento *bayesian regularization*. Um esquemático desta rede é apresentado pela Fig. 2. Nesta representação, b representa a constante que é adicionada ao produto da entrada pelo seu respectivo peso sináptico w .

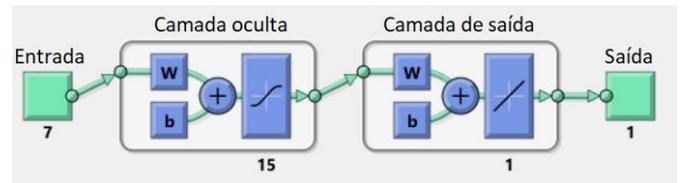


Fig. 2 Esquemático Rede Neural.

Existem outras funções de treinamento diferentes da *bayesian regularization*, que podem minimizar o erro de treinamento. Em testes realizados com estas funções, que não possuem regularização, ocorreu excesso de treinamento e a rede generalizou em todos os casos testados. Com isso, as redes testadas apresentavam um baixo erro de treinamento, porém um alto erro na validação.

Os dados foram separados conforme o método de validação cruzada. O treinamento foi realizado até o MAPE atingir o menor valor possível. A regressão deste treinamento apresentou uma acuracidade de 0,9637 entre os dados utilizados como alvo, e os dados de saída do treinamento, como mostra a Fig. 3.

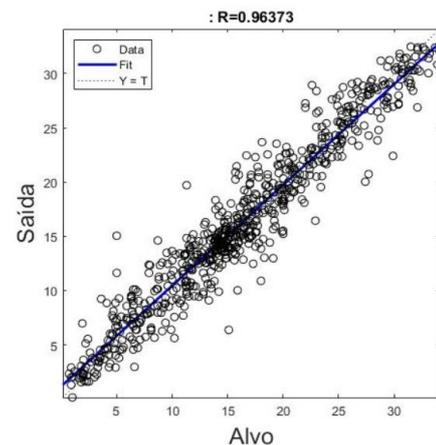


Fig. 3 Regressão Treinamento PDAV

O MAPE atingido durante o treinamento foi de 14,86%. Após o treinamento da rede neural, os dados de validação são inseridos como entrada da rede neural. Com a rede neural treinada a saída da rede é por aprendizagem, e por meio dos dados inseridos na entrada a rede prevê os dados de saída. A Fig. 4, mostra a regressão entre a saída e os dados de comparação, com uma acuracidade de 0,9513.

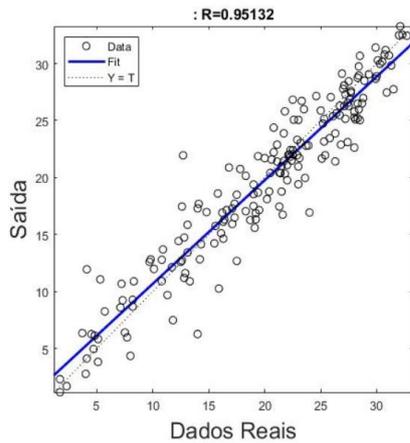


Fig. 4 Regressão Previsão PDAV.

Com um MAPE calculado em 12,97%, a Fig. 5 ilustra a previsão realizada para 184 dias a frente. Os dados em azul representam os dados reais de geração obtidos em campo e utilizados aqui somente como comparação e os dados em vermelho representam a saída da rede.

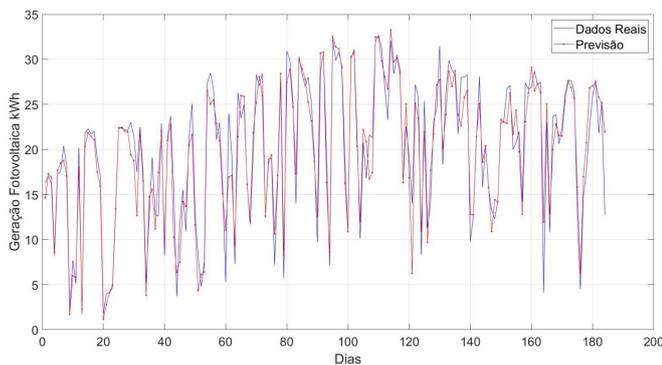


Fig. 5 Saída Previsão PDAV.

4.2 Previsões Rede Neural Mista

Nesta rede neural foram incluídas as 9 variáveis obtidas por meio do Portal Clima em conjunto com as 4 componentes de irradiação disponíveis na Rede Neural PDAV. A classificação eliminou 6 variáveis, portanto, esta rede neural teve 7 variáveis de entrada: umidade relativa, umidade absoluta, pressão atmosférica absoluta, precipitação, irradiação incidente sobre uma superfície horizontal, fluxo térmico radiativo infravermelho, irradiação de topo da atmosfera.

A RNA foi modelada com 15 neurônios ocultos, com função de ativação sigmoide tangente hiperbólica e função de treinamento *bayesian regularization*. A Fig. 6 ilustra a regressão entre as saídas do treinamento e seus respectivos alvos. Pode ser observado que a acuracidade desta regressão alcançou 0,9684.

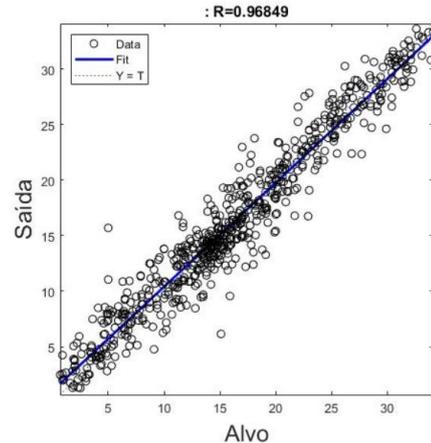


Fig. 6 Regressão Treinamento Mista

Neste treinamento a rede neural Mista apresentou um erro de 15,37 %. Com a rede neural treinada são inseridos os dados para a previsão. Na Fig. 7 é apresentado o resultado da regressão da rede neural Mista entre a saída da previsão e os dados reais utilizados para a comparação. Esta regressão apresentou uma acuracidade de regressão 0,9374. A Fig. 8 ilustra a previsão realizada.

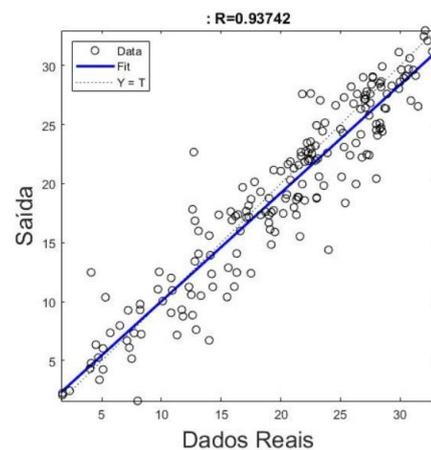


Fig. 7 Regressão Previsão Mista.

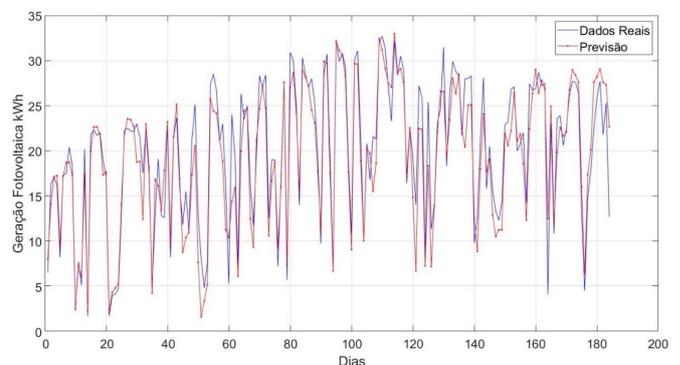


Fig. 8 Saída Previsão Mista.

Sendo, na Fig. 8, os dados em vermelho a previsão e os dados em azul os valores reais utilizados para comparação. Nesta

previsão um MAPE de 13,87 % foi calculado, resultado 0,9% inferior a rede anterior.

4.3 Análise Comparativa de Precisão entre as Redes Neurais

A Tabela 1 apresenta um comparativo dos erros apresentados durante a fase de treinamento. As métricas utilizadas foram o MAPE, o RMSE e o MSE. A última coluna mostra o valor R da regressão entre a saída e seus respectivos alvos.

Tabela 1. Comparativo dos erros de treinamento entre as duas RNAs.

	MAPE (%)	RMSE (%)	MSE (%)	R
PDAV	14,86	2,13	4,57	0,96
MISTA	15,37	2,00	4,02	0,96

A Tabela 2 mostra um comparativo dos erros apresentados na previsão para 184 dias para as duas redes neurais utilizadas.

Tabela 2. Comparativo dos erros de previsão entre as duas RNAs.

	MAPE (%)	RMSE (%)	MSE (%)	R
PDAV	12,97	2,42	5,90	0,95
MISTA	13,87	2,85	8,14	0,93

As redes neurais apresentaram valores de MAPE, de 14,86 e 15,37 % no treinamento e de 12,97 e 13,87 % na previsão.

4.4 Previsões Realizadas para Diferentes Temporadas

Com o objetivo de verificar as variações entre o erro nas previsões para diferentes épocas do ano, por meio da verificação do MAPE, foram realizadas previsões para temporadas de inverno e verão. Sendo assim, no primeiro caso observou-se o comportamento para previsões realizadas para os meses de inverno, com dados de validação para 92 dias de inverno a partir de 21 de junho de 2018. No segundo, para a temporada de verão, foram utilizadas 92 amostras de dias quentes, partindo-se de 17 de novembro. Posteriormente, foi abordada a previsão para um único mês, sendo utilizados como validação os meses de julho, novembro e dezembro de 2018. A Tabela 3 mostra o comparativo entre os erros nas previsões realizadas utilizando as três redes neurais.

Tabela 3. Tabela Comparativa de MAPE.

	PDAV MAPE %	MISTA MAPE %	INTERVALO
INVERNO	21,6665	19,8252	92 dias
VERÃO	12,4300	14,5704	92 dias
JULHO	22,9639	28,2650	31 dias
NOVEMBRO	9,0993	9,1215	30 dias

	PDAV MAPE %	MISTA MAPE %	INTERVALO
DEZEMBRO	10,9700	11,51930	31 dias

Com estas previsões é perceptível o aumento no MAPE em situações de baixa geração de energia. Como o erro é calculado por meio da diferença absoluta entre o valor real e o previsto dividido pelo valor real, quando o valor real é pequeno o erro tende a aumentar. Na Fig. 9 é ilustrada a previsão realizada para o mês de julho, com MAPE de 22,96 % e a Fig. 10 mostra a previsão realizada para o mês de novembro com MAPE de 9,09 %.

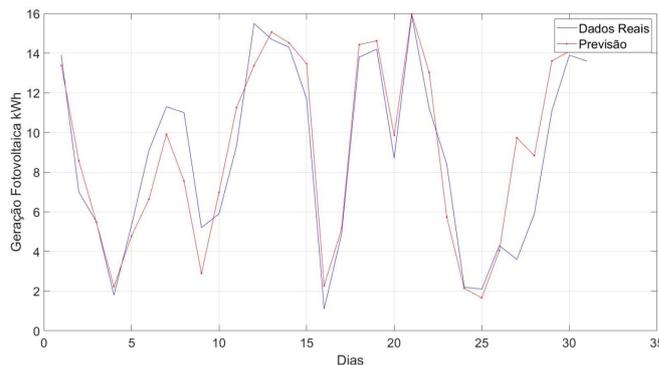


Fig. 9 Previsão Realizada para julho, MAPE 22,96 %.

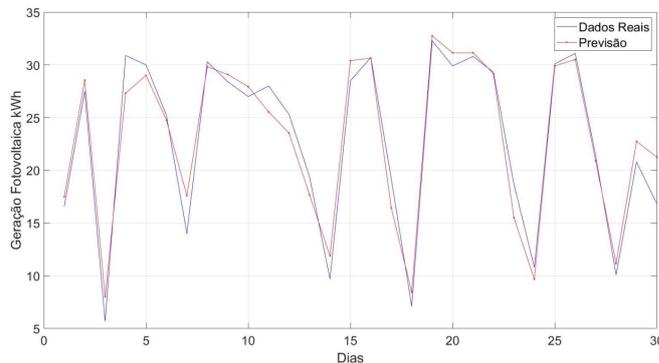


Fig. 10 Previsão Realizada para novembro, MAPE 9,09 %.

Como o horizonte de dias para a validação da previsão era menor, para cada horizonte foi realizado um novo treinamento. Este treinamento permitiu aumentar o número de dados amostrais utilizados. Os erros durante o processo de treinamento permaneceram semelhantes, entre 14,37 e 15 %, como é apresentado na Tabela 4.

Tabela 4. MAPE Calculado nos Diferentes Treinamentos.

	PDAV MAPE %	MISTA MAPE %
VERÃO	14,8947	14,6365
INVERNO	14,3787	14,6040
JULHO	14,6470	14,4951
NOVEMBRO	15,0096	14,3349

	PDAV MAPE %	MISTA MAPE %
DEZEMBRO	14,7533	14,1208
6 MESES	14,8694	15,3763

4.5 Comparação com a Literatura

As Tabelas 5 e 6 mostram as diferentes previsões encontradas na bibliografia comparadas com as melhores previsões das redes elaboradas neste trabalho. Estas previsões foram todas realizadas através de redes neurais artificiais, sendo o MAPE a métrica de validação empregada. Em Lahouar, Mejri e Slama (2017) são realizadas previsões com dados obtidos de uma região da Austrália. Em Cheng, GE e Cao (2012) as previsões são realizadas para uma região na China e as previsões de Haque, Nehrir e Mandal (2013) são para a cidade de Ashland no estado do Óregon, Estados Unidos.

Tabela 5. Comparação de MAPE para Previsão a Médio Prazo.

Autores	Intervalo	MAPE %
Lahouar, Mejri e Slama (2017)	outubro	10,26
Lahouar, Mejri e Slama (2017)	janeiro	19,16
Lahouar, Mejri e Slama (2017)	abril	28,98
Lahouar, Mejri e Slama (2017)	julho	17,45
O autor (2019)	julho	22,23
O autor (2019)	novembro	13,20
O autor (2019)	dezembro	9,78

Tabela 6. Comparação de MAPE para Previsão a Longo Prazo.

Autores	Intervalo	MAPE %
Cheng, GE e Cao (2012)	1º 90 dias	20,24
Cheng, GE e Cao (2012)	2º 91 dias	19,31
Cheng, GE e Cao (2012)	3º 92 dias	15,52
Cheng, GE e Cao (2012)	4º 92 dias	21,98
Haque, Nehrir e Mandal (2013)	inverno	15,77
Haque, Nehrir e Mandal (2013)	primavera	16,08
Haque, Nehrir e Mandal (2013)	verão	15,41
Haque, Nehrir e Mandal (2013)	outono	19,17
O autor	inv. 92 dias	19,82
O autor	ver. 92 dias	12,43
O autor	6 meses	12,97

Nas previsões realizadas para médio prazo por Lahouar, Mejri e Slama (2017), assim como os resultados, aqui apresentados, mostram a variação no MAPE calculado para as diferentes épocas do ano, as quais foram realizadas as previsões. Nas previsões realizadas para longo prazo esta variação também ocorre. Os autores Cheng, GE e Cao (2012) e Haque, Nehrir e Mandal (2013), realizaram previsões para diferentes épocas do ano, o primeiro realizou suas previsões para os quatro trimestres do ano analisado, o segundo segregou suas previsões pelas estações do ano. Em ambas as previsões, como também nas previsões apresentadas neste texto, o MAPE calculado varia de acordo com a época.

É preciso ressaltar o desempenho da métrica de validação MAPE aplicada em situações de baixos valores reais, pois quando há uma divisão por um dado real de baixo valor, o erro tende a aumentar. Além disso, as redes neurais mostraram-se realmente capazes de lidar com as intermitências impostas pela geração fotovoltaica. A não linearidade na geração de energia foi superada pela rede neural, que trouxe resultados e variações semelhantes aos dados reais comparados.

O emprego de aprendizado de máquina, neste caso as redes neurais artificiais, qualificaram características climáticas próprias para a quantidade de energia fotovoltaica gerada. Em especial às componentes derivadas da irradiação solar, como a irradiação sobre uma superfície horizontal. A modelagem das redes neurais aqui apresentadas, buscou por meio do processo de classificação de variáveis *stepwise* avaliar dentre as candidatas, aquelas que mais correspondem com a resposta esperada. Esta seleção reforçou as variáveis presentes com maior frequência na bibliografia, dentre as quais podem ser citadas a temperatura máxima do ambiente, a umidade relativa do ar e a irradiação sobre uma superfície horizontal. O processo de aprendizagem atingiu um MAPE entre 14 e 15% no treinamento. Este procedimento foi realizado buscando preservar a capacidade de aprendizagem da rede neural, aplicando uma função de treinamento que prevenisse o excesso de treinamento. O objetivo de realizar uma previsão de 6 meses de geração fotovoltaica foi alcançado com resultados satisfatórios, com um MAPE de 12,97 %, erro menor do que as previsões a longo prazo presentes na bibliografia que utilizam redes neurais artificiais.

5. CONCLUSÕES

As redes neurais artificiais foram escolhidas como método de aprendizado de máquina para estudo, com ponto central nas suas aplicações relacionadas à previsão de geração fotovoltaica. O emprego da regressão *stepwise* colaborou na classificação das variáveis que melhor representavam o sistema, reduzindo o número de variáveis de entrada da rede neural e consequentemente reduzindo o tempo de treinamento em conjunto com o aumento da acuracidade do modelo.

As variáveis componentes da irradiação solar mostraram grande influência para a precisão da previsão, quando estas foram incluídas em conjunto com os dados climáticos do Portal Clima. Quando os valores reais utilizados para comparação são baixos, o MAPE calculado tende a ser maior. As previsões realizadas para os meses de inverno e para o

mês de julho apresentaram MAPE elevado de 19,82 % e 22,96 % respectivamente. No entanto, previsões realizadas para os meses de verão, para o mês de novembro e para o mês de dezembro apresentaram MAPE de 12,43 %, 9,09 % e 10,97 % respectivamente. Com estes resultados, foi possível observar um aumento no MAPE nos dias de baixa geração.

O melhor resultado do objetivo principal deste estudo atingiu um MAPE de 12,97 %, para o horizonte de seis meses, estando estes resultados de acordo com os encontrados na bibliografia. A não linearidade nos dados de geração fotovoltaica analisados e dias com baixa geração culminaram no acréscimo do MAPE nas previsões realizadas. Em trabalhos futuros será possível empregar esta mesma rede neural validada, para estimar a geração de outros locais. Há também a possibilidade de empregar outras técnicas de aprendizado de máquina citadas neste texto, como por exemplo máquina de vetores de suporte. Também podem ser analisadas previsões de curto prazo, para um ou dois dias a frente, utilizando conjuntos de dados segmentados em minutos.

REFERÊNCIAS

- Abedin, Zainal et al. A model for prediction of monthly solar radiation of different meteorological locations of Bangladesh using artificial neural network data mining tool. 2017. Disponível em: <<https://ieeexplore.ieee.org/document/7912993/>>. Acesso em: 25 ago. 2018.
- Alzahrani, Ahmad et al. Solar Irradiance Forecasting Using Deep Recurrent Neural Networks. 2017. Disponível em: <<https://ieeexplore.ieee.org/document/8191206/>>. Acesso em: 23 ago. 2018.
- ANEEL (Brasil). Nº 482: RESOLUÇÃO NORMATIVA Nº 482. 2012. 12 p. Disponível em: <<http://www2.aneel.gov.br/cedoc/ren2012482.pdf>>. Acesso em: 02 nov. 2018.
- Chaturvedi, D. K.; Isha. Solar Power Forecasting: A Review. 2016. Disponível em: <<https://pdfs.semanticscholar.org/a679/0993e18db1b5d488b161194fa07a3d5c139f.pdf>>. Acesso em: 20 ago. 2018.
- Cheng, Hang; Ge, Peng-Jiang; Cao, Wu-Shun. Forecasting Research of Long-term Solar Irradiance and Output Power for Photovoltaic generation system. 2012. Disponível em: <<https://ieeexplore.ieee.org/document/6301338/>>. Acesso em: 03 set. 2018.
- Garibaldi. Prefeitura Garibaldi. Portal Clima. 2019. Disponível em: <http://clima.garibaldi.rs.gov.br/historico.aspx?EST_ID=1>. Acesso em: 07 mar. 2019.
- Ghunem, Refat A.; Assaleh, Khaled; El-Hag, Ayman H.. Artificial Neural Networks with Stepwise Regression for Predicting Transformer Oil Furan Content. 2012. Disponível em: <<https://ieeexplore.ieee.org/document/6180233>>. Acesso em: 15 abr. 2019.
- Haque, Ashraf U.; Nehrir, M. Hashem; Mandal, Paras. Solar PV Power Generation Forecast Using a Hybrid Intelligent Approach. 2013. Disponível em: <<https://ieeexplore.ieee.org/document/6672634>>. Acesso em: 26 out. 2018.
- Haykin, Simon. Redes Neurais - Princípios e práticas. 2. ed. Porto Alegre: Bookman, 2001. 908 p. Disponível em: <http://www.ncdd.com.br/livros/redes_neurais_simon_haykin.pdf>. Acesso em: 06 out. 2018.
- Lahouar, Ali; Mejri, Amal; Slama, Jaleleddine Ben Hadj. Importance Based Selection Method for Day-ahead Photovoltaic Power Forecast Using Random Forests. 2017. Disponível em: <<https://ieeexplore.ieee.org/document/8066171/>>. Acesso em: 26 ago. 2018.
- Liu, Zhao; Zhang, Ziang. Solar Forecasting by K-Nearest Neighbors Method with Weather Classification and Physical Model. 2016. Disponível em: <<https://ieeexplore.ieee.org/document/7747859/>>. Acesso em: 25 ago. 2018.
- Majumder, Irani; Behera, Manoj Kumar; Nayak, Niranjan. Solar Power Forecasting Using a Hybrid EMD-ELM Method. 2017. Disponível em: <<https://ieeexplore.ieee.org/document/8074179/>>. Acesso em: 03 set. 2018.
- NASA (Estados Unidos). POWER Data Access Viewer. 2018. Disponível em: <<https://power.larc.nasa.gov/data-access-viewer/>>. Acesso em: 25 set. 2018.
- Saini, Lalit Mohan. Peak load forecasting using Bayesian regularization, Resilient and adaptive backpropagation learning based artificial neural networks. 2007. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0378779607002258>>. Acesso em: 01 maio 2019.
- Sansa, Ines et al. PV power forecasting using different artificial neural networks strategies. 2014. Disponível em: <<https://ieeexplore.ieee.org/document/6835397/>>. Acesso em: 11 set. 2018.
- Silva, Renato Maia. Redes Neurais Artificiais aplicadas à Detecção de Intrusão em Redes TCP/IP. 2005. 144 f. Dissertação (Mestrado) - Curso de Engenharia Elétrica, Centro Técnico Científica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2005. Cap. 4. Disponível em: <https://www.maxwell.vrac.puc-rio.br/7335/7335_5.PDF>. Acesso em: 06 out. 2018.
- Singh, Vikas Pratap et al. Forecasting of 5MW Solar Photovoltaic Power Plant Generation Using Generalized Neural Network. 2015. Disponível em: <<https://ieeexplore.ieee.org/document/7489107/>>. Acesso em: 19 set. 2018.
- Stackhouse, Paul W. et al. POWER Release 8 (with GIS Applications) Methodology (Data Parameters, Sources, & Validation). 2018. Disponível em: <https://power.larc.nasa.gov/documents/POWER_Data_v8_methodology.pdf>. Acesso em: 25 set. 2018.