

# INFLUÊNCIA DO PREDITOR ÍNDICE NIÑO OCEÂNICO (ONI) NA PREVISÃO DA IRRADIAÇÃO GLOBAL HORIZONTAL DE FORTALEZA EM DIFERENTES HORIZONTES TEMPORAIS UTILIZANDO MACHINE LEARNING

**Juliana Silva Brasil** (UFC) - julianasbra@gmail.com

**Felipe Pinto Marinho** (UFC) - fpmarinho@alu.ufc.br

**Rahysa Silva de Oliveira** (UFC) - rahysaso@gmail.com

**David Mickely Jaramillo Loayza** (UFC) - mickely.jaramillo@outlook.com

**Paulo Alexandre Costa Rocha** (UFC) - paulo.rocha@ufc.br

**Andre Valente Bueno** (UFC) - bueno@ufc.br

**Ricardo Jose Pontes Lima** (UFC) - rjponteslima@gmail.com

**Maria Eugenia Vieira da Silva** (UFC) - eugenia@ufc.br

## **Resumo:**

*A introdução da energia solar na rede elétrica vem aumentando anualmente, inclusive no Ceará. Nesse contexto, são suscitadas questões relativas à complementaridade de abastecimento, uma vez que a fonte solar é intermitente e variável com o tempo. O planejamento adequado pode ser facilitado pela previsão da irradiação utilizando métodos de aprendizagem de máquinas. O presente trabalho avaliou o desempenho de 3 modelos de previsão da irradiação global horizontal - rede neural, Boosting e modelo de persistência - para a cidade de Fortaleza, Ceará, em 8 horizontes temporais distintos, analisando a influência do El Niño e da La Niña, na forma do preditor ONI, nessas previsões. Utilizou-se, além do ONI, informações meteorológicas (temperatura ambiente, umidade relativa, velocidade do ar, direção do vento e nível de precipitação), dados de irradiação e hora e data da aquisição da informação. O banco de dados foi separado em 2 grupos, um de treinamento e outro de teste, bem como foi realizada uma seleção de parâmetros para a rede neural e para o Boosting. O cálculo da variabilidade dos dados de irradiação classificou esse preditor como de fraca variabilidade. Os resultados apontaram que houve redução do RMSE quando o preditor referente à intensidade do El Niño, ONI, foi adicionado, exceto para o horizonte de 2 min, o que indica que a adição do preditor melhorou o desempenho dos modelos de previsão. Notou-se também que o nRMSE cresceu conforme o horizonte temporal aumentava. O Boosting apresentou os menores erros dentre os modelos considerados, seus valores de nRMSE são compatíveis com aqueles encontrados na literatura, chegando a ser menor para um determinado horizonte de tempo. Para a rede neural, todavia, os nRMSE obtidos foram superiores aos encontrados em trabalhos já publicados.*

**Palavras-chave:** *Aprendizagem de máquina, Irradiação global horizontal, El Niño*

**Área temática:** *Radiação Solar*

**Subárea temática:** *Recursos Solares e Meteorologia da Radiação Solar*

# INFLUÊNCIA DO PREDITOR ÍNDICE NIÑO OCEÂNICO (ONI) NA PREVISÃO DA IRRADIAÇÃO GLOBAL HORIZONTAL DE FORTALEZA EM DIFERENTES HORIZONTES TEMPORAIS UTILIZANDO MACHINE LEARNING

**Juliana Silva Brasil** – julianasbra@gmail.com

**Felipe Pinto Marinho** – felipe.pinto.marinho@gmail.com

**Rahysa Silva de Oliveira** – rahysaso@gmail.com

**David Mickely Jaramillo Loayza** – mickely.22jara@gmail.com

**Paulo Alexandre Costa Rocha** – paulo.rocha@ufc.br

**André Valente Bueno** – bueno@ufc.br

**Ricardo José Pontes Lima** – rjponteslima@gmail.com

**Maria Eugênia Vieira da Silva** – eugenia@ufc.br

Universidade Federal do Ceará, Departamento de Engenharia Mecânica

**Resumo.** A introdução da energia solar na rede elétrica vem aumentando anualmente, inclusive no Ceará. Nesse contexto, são suscitadas questões relativas à complementaridade de abastecimento, uma vez que a fonte solar é intermitente e variável com o tempo. O planejamento adequado pode ser facilitado pela previsão da irradiação utilizando métodos de aprendizagem de máquinas. O presente trabalho avaliou o desempenho de 3 modelos de previsão da irradiação global horizontal - rede neural, Boosting e modelo de persistência – para a cidade de Fortaleza, Ceará, em 8 horizontes temporais distintos, analisando a influência do El Niño e da La Niña, na forma do preditor ONI, nessas previsões. Utilizou-se, além do ONI, informações meteorológicas (temperatura ambiente, umidade relativa, velocidade do ar, direção do vento e nível de precipitação), dados de irradiação e hora e data da aquisição da informação. O banco de dados foi separado em 2 grupos, um de treinamento e outro de teste, bem como foi realizada uma seleção de parâmetros para a rede neural e para o Boosting. O cálculo da variabilidade dos dados de irradiação classificou esse preditor como de fraca variabilidade. Os resultados apontaram que houve redução do RMSE quando o preditor referente à intensidade do El Niño, ONI, foi adicionado, exceto para o horizonte de 2 min, o que indica que a adição do preditor melhorou o desempenho dos modelos de previsão. Notou-se também que o nRMSE cresceu conforme o horizonte temporal aumentava. O Boosting apresentou os menores erros dentre os modelos considerados, seus valores de nRMSE são compatíveis com aqueles encontrados na literatura, chegando a ser menor para um determinado horizonte de tempo. Para a rede neural, todavia, os nRMSE obtidos foram superiores aos encontrados em trabalhos já publicados.

**Palavras-chave:** Aprendizagem de máquina, Irradiação global horizontal, El Niño.

## 1. INTRODUÇÃO

A inserção da energia solar na matriz energética brasileira tem crescido a cada ano e, conforme dados recentes de Soares *et al.* (2019), a fonte solar apresentou crescimento de 316,1% na geração de energia entre os anos de 2017 e 2018, apesar de representar apenas 0,5% da matriz energética. Quando se trata de mini e microgeração, todavia, essa fonte corresponde a 63,5% da geração, mostrando ainda tendência de aumento.

O estado do Ceará, na região Nordeste do Brasil, se destaca na geração de energia por fontes renováveis, dentre as quais a solar. A região é favorecida com altos índices de incidência solar, conforme Menezes *et al.* (2011). A geração solar na região representa uma alternativa à geração hidrelétrica local, que é bastante afetada pelos períodos de escassez de chuvas, e que geralmente acarretam aumento na tarifa de energia.

O crescente uso da fonte solar, e a consequente incorporação desta a uma rede elétrica, traz consigo questões relativas à complementaridade de abastecimento, uma vez que a fonte solar é intermitente e variável com o tempo segundo Notton *et al.* (2018). Nesse cenário, a capacidade de prever confiavelmente a intensidade da irradiação para minutos, horas, dias subsequentes se torna conveniente. Métodos de aprendizagem de máquinas podem ser utilizados com essa finalidade.

Previsões com erros reduzidos requerem, dentre outros fatores, uma adequada seleção de preditores. Conforme Mohammadi *et al.* (2018), informações meteorológicas tais como velocidade do vento e nível de precipitação estão relacionados com a intensidade de irradiação. Fenômenos como El Niño e La Niña alteram os regimes de chuva e vento em escala global, segundo Kayano *et al.* (2016), podendo, desse modo, influir na irradiação.

O El Niño e a La Niña são alterações na temperatura superficial do Oceano Pacífico, na região próxima a Linha do Equador. O El Niño caracteriza-se pelo aquecimento e a La Niña pelo resfriamento anormais. Esses eventos são

classificados conforme sua intensidade pelo ONI (The Oceanic Niño Index), que avalia a intensidade do El Niño e da La Niña, variando de -4 (La Niña forte) a 4 (El Niño forte).

O presente estudo consiste em avaliar a influência dos fenômenos El Niño e La Niña, na forma do preditor ONI, para a previsão da irradiação global horizontal na cidade de Fortaleza, Ceará, em diferentes horizontes de tempo, utilizando um método de árvores de regressão com *Boosting*, uma rede neural perceptron de múltiplas camadas (ANN) e um modelo de persistência.

A variabilidade dos dados de irradiação também foi avaliada, a fim de proporcionar uma comparação adequada com outros trabalhos já publicados. Em Fouilloy *et al.* (2018) foi apontado que cidades com diferentes classificações de variabilidade dos dados de índice de claridade apresentam faixas de erro de previsão distintas.

## 2. METODOLOGIA

As etapas desenvolvidas neste trabalho estão esquematizadas na Fig. 1.

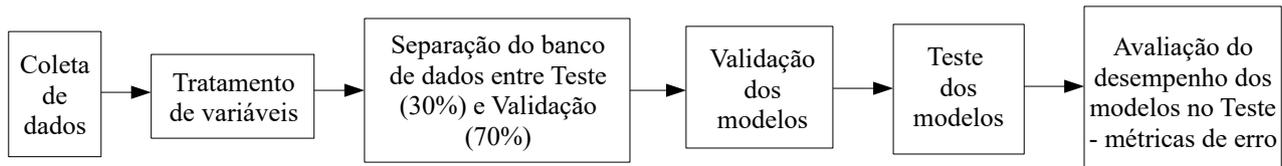


Figura 1 - Representação das etapas do trabalho.

### 2.1 Dados

Os dados foram coletados na cidade de Fortaleza, Ceará, 3° 43' 2''S, 38° 32' 35'' O, ao nível do mar, durante os anos de 2015 a 2019. Os preditores meteorológicos foram obtidos por meio da Fundação Cearense de Meteorologia e Recursos Hídricos (Funceme) e os dados de irradiação global foram coletados pelo Laboratório de Energia Solar e Gás Natural (LESGN), da Universidade Federal do Ceará, por um piranômetro Kipp & Zonen, modelo CM11, com sensibilidade de  $4,55 \cdot 10^{-6}$  V/W.m<sup>2</sup>.

Foram considerados os dados obtidos entre 5h30min da manhã e 17h30min da tarde.

As previsões foram feitas para os horizontes temporais de: 2 min, 10 min, 30 min, 1h, 2h, 6h, 24h, 7 dias. Para todos estes horizontes utilizou-se um banco de dados com observações obtidas a cada 2 min.

Foram considerados neste trabalho os preditores: ano, dia e horário da aquisição, ONI, temperatura e umidade relativa do ar, velocidade e direção do vento, precipitação, irradiação global do instante atual e irradiação global de 2, 4, 6, 8 e 10 minutos anteriores.

**Tratamento dos dados.** Os valores de irradiação, tanto a do instante atual quanto a dos instantes anteriores, obtidos em W/m<sup>2</sup>, foram transformados no Índice de Claridade,  $K_t$ , que é a razão entre a irradiação medida e a irradiação extraterrestre para aquele instante e lugar, conforme Inman *et al.* (2013). Esse índice pode ser calculado pelas Eq. (1) e Eq. (2), sugeridas em Duffie e Beckman, (2006).

$$K_t = \frac{H}{H_o} \quad (1)$$

$$H_o = \frac{24}{\pi} G_{SC} \left[ 1 + 0,033 \cos\left(\frac{360}{365} n\right) \right] \left( \cos \phi \cos \gamma \sin w_s + \frac{\pi W_s}{180} \sin \phi \sin \gamma \right) \quad (2)$$

Onde  $G_{SC}$  é a constante solar, cujo valor é de 1367 W/m<sup>2</sup>,  $\phi$  é a latitude do local considerado, em radianos,  $\gamma$  é o ângulo de declinação solar ao meio-dia em relação à linha do Equador, em radianos,  $W_s$  é o ângulo do pôr do sol, em radianos,  $n$  é o dia do ano no calendário juliano,  $H$  é a irradiação medida na Terra, em W/m<sup>2</sup>, e  $H_o$  é a irradiação extraterrestre, em W/m<sup>2</sup>.

O dia do ano também foi modificado a fim de reduzir a sua amplitude de valores, conforme sugerido por Zhu *et al.* (2019). É calculado pela Eq. (3).

$$\text{Dia do ano} = \cos\left(\frac{2\pi(DOY-15)}{365}\right) \quad (3)$$

Onde DOY é o respectivo dia do ano no calendário juliano.

**El niño e La niña.** Também foi acrescentado como preditor o Índice Niño Oceânico (ONI), que classifica os episódios de El Niño e La Niña de acordo com sua intensidade, variando conforme indicado na Tab. 1. Essa informação

foi obtida pela Administração Nacional Oceânica e Atmosférica (NOOA), órgão do governo estadunidense. O índice é obtido para cada mês, todavia só é feita a classificação de intensidade após três meses de recorrência do mesmo valor de ONI.

Tabela 1 - Categorias do ONI.

Valores de ONI			
El Niño		La Niña	
ONI	Classificação	ONI	Classificação
4	El niño bastante forte	-4	La niña bastante forte
3	El niño forte	-3	La niña forte
2	El niño moderado	-2	La niña moderada
1	El niño fraco	-1	La niña fraca
0	Ausência dos eventos	-	-

## 2.2 Modelos

A implementação dos modelos foi feita utilizando a linguagem aberta R.

Para os modelos ANN e Boosting, o banco de dados foi separado em 2 grupos disjuntos e aleatórios de igual tamanho: treinamento e teste. O aprendizado do modelo foi realizado no conjunto de treinamento e os resultados de interesse foram obtidos pela aplicação dos modelos ajustados no conjunto de teste.

**Persistência.** Consiste de um modelo para servir de referência de performance para os demais métodos. Espera-se que o modelo de persistência apresente erro maior do que os demais modelos. Nele considerou-se que o Índice de Claridade do instante atual seria igual ao do instante anterior, conforme a Eq. (4).

$$Kt_t = Kt_{t-\Delta t} \quad (4)$$

**ANN.** Perceptron de Múltiplas Camadas é um tipo de rede neural classicamente aplicada em problemas de previsão. O tipo utilizado neste trabalho possui 3 camadas, entrada, intermediária ou oculta e saída. O método é inspirado nas transmissões de impulsos nervosos entre neurônios no corpo humano. Caracteriza-se pela presença de neurônios, ou nós, que ao receberem as informações de neurônios anteriores, lhes atribuem pesos distintos e posteriormente as somam. Esse somatório passa por uma função de ativação linear, que normaliza a resposta, transmitindo a saída. As funções de ativação utilizadas podem variar de acordo com a adequação destas aos dados em estudo, conforme sugerido por Inman *et al.* (2013). O funcionamento de uma rede neural está esquematizado na Fig. 2.

Neste trabalho, avaliou-se o número de neurônios da camada intermediária, a fim de encontrar o valor que proporcionasse menor erro.

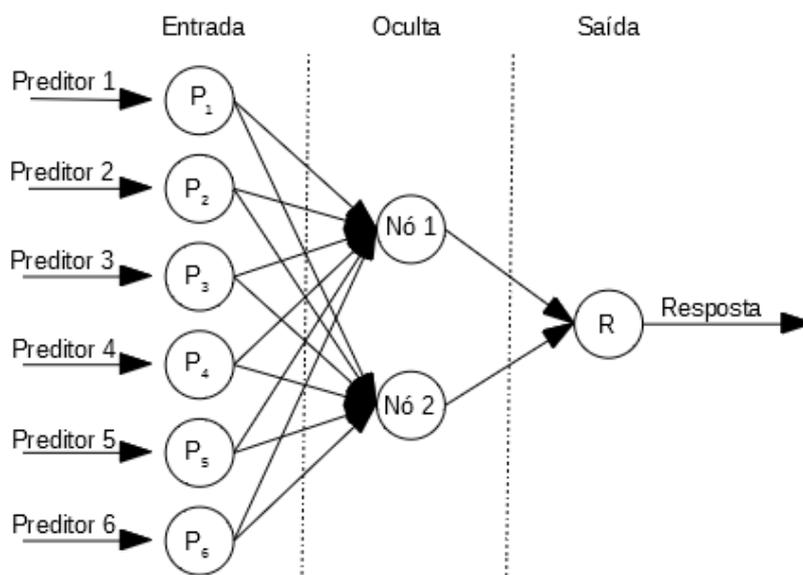


Figura 2 - Esquema simplificado da rede neural com 6 preditores de entrada e dois neurônios, os nós, na camada oculta.

**Boosting.** Consiste em um modelo *Ensemble*, que inicialmente era aplicado em problemas de classificação, mas por meio de adaptações no algoritmo, que resultaram na obtenção do método de Boosting Adaptativo (*Adaboost*), tem-se a possibilidade de utilização em problemas de regressão.

Esse modelo utiliza-se de uma função de perda que, para cada linha de treinamento do banco de dados, calcula um erro de previsão. A função utilizada pode ser linear, quadrática ou exponencial, podendo ser alterada para cada iteração. O erro obtido para cada linha é ponderado por um peso,  $D_t$ , que é maior para erros maiores e menor para erros menores. A cada iteração a distribuição de pesos é atualizada por meio da Eq. (7), a qual utiliza o fator de confiabilidade  $\beta_t$ , da Eq. (6) que é calculado com base na função de perda total,  $L_t$ , Eq. (5), que é a soma dos erros de cada linha, ponderados pelos respectivos pesos da última iteração.

$$\bar{L}_t = \sum_{i=1}^N (L_t(i) D_t(i)) \quad (5)$$

$$\beta_t = \frac{\bar{L}_t}{1 - \bar{L}_t} \quad (6)$$

$$D_{t+1}(i) = \frac{D_t(i) \beta_t^{(1-L_t(i))}}{Z_t} \quad (7)$$

Onde  $Z_t$  é um fator de normalização.

Ao final de cada iteração uma hipótese fraca é gerada. A hipótese final é calculada pelo somatório das hipóteses fracas ponderadas por seus respectivos fatores de confiabilidade. Organizam-se os valores das hipóteses fracas em ordem crescente, bem como os fatores de confiabilidade. Quando a relação da Eq. (8) for atingida, considera-se o  $f_t(x)$  correspondente ao  $\beta_t$  da última soma como a hipótese final.

$$f_{final}(x) = \inf [y \in Y : \sum_t \log \frac{1}{\beta_t} \geq \frac{1}{2} \sum_t \log \frac{1}{\beta_t}] \quad (8)$$

Esse modelo proporciona uma progressiva redução do erro no conjunto de treinamento, por isso é frequentemente utilizado com outros métodos a fim de melhorar o desempenho destes.

### 2.3 Métricas de Erro

Foram utilizadas para comparar os modelos e avaliar a adequação destes ao banco de dados trabalhado e ao problema em estudo.

**Raiz do Erro Quadrático Médio (RMSE).** O RMSE é um importante parâmetro de avaliação do ajuste de um modelo. É dado na mesma unidade da variável predita e é calculado pela Eq. (9). Nele os erros individuais são ponderados diferentemente, sendo os erros maiores ponderados por pesos maiores.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{K}t - Kt)^2}{N}} \quad (9)$$

Onde,  $\hat{K}t$  é o valor estimado de  $Kt$ , previsto pelo modelo considerado.

**Raiz do Erro Quadrático Médio Normalizado (nRMSE).** O nRMSE, é a razão entre o RMSE e a média dos valores reais da variável. É obtido na forma de porcentagem. Essa métrica de erro fornece faixas de classificação para a previsão, a saber: % nRMSE < 10% excelente, 10% < % nRMSE < 20% bom, 20% < % nRMSE < 30% razoável e % nRMSE ≥ 30% ruim, conforme sugerido por Li *et al.* (2013). É calculada pela Eq. (10).

$$nRMSE = \frac{RMSE}{\bar{K}t} \quad (10)$$

Onde,  $\bar{K}t$  é o valor médio da variável  $Kt$ .

**Erro Absoluto Médio (MAE).** O MAE, o erro médio absoluto, é a média do valor absoluto da diferença entre o valor predito e o valor real. Nesta métrica os erros individuais são ponderados igualmente pela média, diferentemente do RMSE. É obtido pela Eq. (11).

$$MAE = \frac{\sum_{i=1}^N |\hat{Kt} - Kt|}{N} \quad (11)$$

**Erro Absoluto Médio Normalizado (nMAE).** O nMAE é a razão entre o MAE e a média dos valores reais da variável, calculado pela Eq. (12).

$$nMAE = \frac{MAE}{\bar{Kt}} \quad (12)$$

**Habilidade de Previsão (S).** é a razão entre o RMSE de um determinado modelo e o RMSE do modelo de persistência. Quanto mais próximo de 1 for o seu valor, melhor é o desempenho do modelo considerado. É calculado conforme a Eq. (13).

$$S = 1 - \frac{RMSE_{\text{modelo } x}}{RMSE_{\text{persistência}}} \quad (13)$$

## 2.4 Variabilidade dos Dados

Conforme proposto por Voyant *et al.* (2015), a variabilidade do índice de claridade, P, foi calculada pela Eq. (14). A variabilidade pode ser classificada nas seguintes categorias: forte para  $P > 0,4$ , média para  $0,2 < P < 0,4$  e fraca para  $P < 0,2$ , conforme sugerido por Fouilloy *et al.* (2018).

$$P = \frac{\sum_{i=1}^N |\log Kt(i) - \log Kt(i-1)|}{N} \quad (14)$$

Onde N é a quantidade de observações disponíveis.

## 3. RESULTADOS E DISCUSSÃO

Obteve-se o valor de 0,1479 para a variabilidade do índice de claridade, classificando os dados de Kt do intervalo considerado como de fraca variabilidade.

Quanto aos parâmetros de cada modelo verificou-se menores erros para execuções da ANN com 12 neurônios na camada intermediária. Para o Boosting constatou-se um melhor desempenho para execuções considerando 400 árvores.

Os resultados obtidos, apresentados nas Tab. 2 e Tab. 3, indicam que a presença do ONI como preditor diminui o RMSE, exceto para a previsão no horizonte de 2 min com o modelo Boosting, na qual ocorreu o inverso. O modelo de melhor desempenho foi o de Árvores de Decisão aplicando o método de Boosting e tanto este modelo quanto a rede neural apresentaram valores de RMSE menores do que o do modelo de persistência.

Para os horizontes temporais de 1h e 2h, com a presença do preditor ONI no modelo Boosting, verificou-se que os valores de nRMSE estão próximos dos obtidos por Fouilloy *et al.* (2018). A previsão para 6h neste trabalho apresentou menor nRMSE do que o supracitado artigo, como indicado no Fig. 3(a). Comparando com os resultados obtidos por Benali *et al.* (2018), para a ANN, encontra-se nele nRMSE menores do que os obtidos neste trabalho, conforme indicado na Fig. 3(b), entretanto deve-se considerar que a variabilidade dos dados de índice de claridade utilizados naquele trabalho é classificada como forte, enquanto neste é fraca. Conforme verificado por Fouilloy *et al.* (2018), para locais onde a variabilidade do índice de claridade é considerada fraca, os erros de previsão são menores, enquanto que para forte variabilidade os erros obtidos são maiores.

Os valores de S estão representados na Fig. 4. Nota-se que para os horizontes de até 6h o valor do S é crescente, a partir deste, entretanto, passa a decrescer. Esse comportamento foi verificado para todas as situações consideradas neste trabalho, exceto aquela para o ANN sem o preditor ONI, que apresentou valores atípicos para 10 min e 1h. Verifica-se também que, conforme o horizonte temporal aumenta, a distância entre os gráficos de Boosting e os gráficos da ANN também cresce, indicando que a diferença entre os valores de nRMSE dos dois modelos aumenta. Nota-se também que a diferença de desempenho com e sem a variável ONI, para cada modelo, não é grande, se acentuando nas previsões de 10 min e 1h para o ANN e nas previsões para 7 dias de ambos os modelos.

Tabela 2 - Resultados (em função do Kt) considerando o preditor ONI.

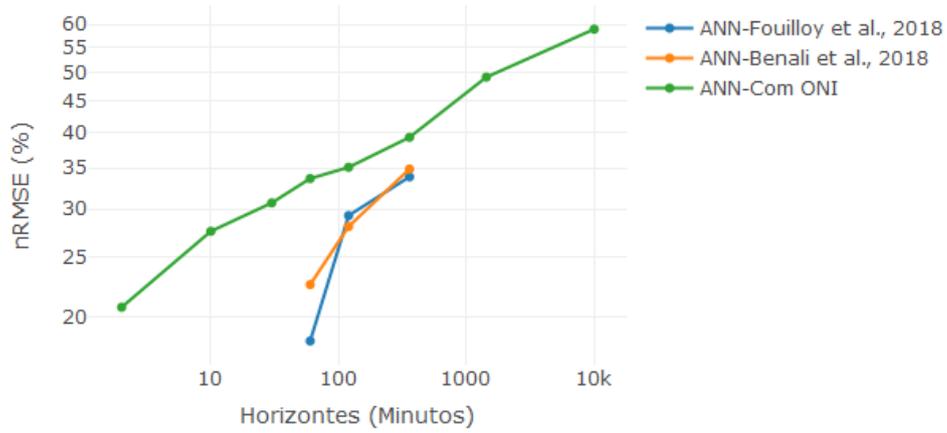
		t+2min	t+10min	t+30min	t+1h	t+2h	t+6h	t+24h	t+7dias
<b>P</b>	<b>RMSE</b>	0,1349	0,2000	0,2698	0,3099	0,4393	0,4886	0,5044	0,4914
	<b>nRMSE</b>	0,2321	0,3497	0,4721	0,5410	0,6014	0,6749	0,6860	0,6934
	<b>MAE</b>	0,0758	0,1177	0,1181	0,2010	0,3030	0,3901	0,3953	0,3713
	<b>nMAE</b>	0,1305	0,2058	0,2066	0,3509	0,4148	0,5389	0,5376	0,5240
<b>ANN</b>	<b>RMSE</b>	0,1204	0,1577	0,1754	0,1927	0,2562	0,2842	0,3619	0,4178
	<b>nRMSE</b>	0,2073	0,2758	0,3069	0,3363	0,3508	0,3926	0,4922	0,5895
	<b>MAE</b>	0,0731	0,1041	0,1181	0,1311	0,1719	0,1974	0,2582	0,3170
	<b>nMAE</b>	0,1258	0,1817	0,2066	0,2289	0,2353	0,2727	0,3511	0,4474
<b>BO</b>	<b>RMSE</b>	0,1216	0,1553	0,1630	0,1691	0,2213	0,2373	0,2724	0,3210
	<b>nRMSE</b>	0,2092	0,2715	0,2852	0,2952	0,3030	0,3279	0,3707	0,4529
	<b>MAE</b>	0,0734	0,1016	0,1093	0,1143	0,1466	0,1617	0,1850	0,2294
	<b>nMAE</b>	0,1263	0,1777	0,1913	0,1996	0,2007	0,2234	0,2517	0,3237

Tabela 3 – Resultados (em função do Kt) desconsiderando o preditor ONI.

		t+2min	t+10min	t+30min	t+1h	t+2h	t+6h	t+24h	t+7dias
<b>P</b>	<b>RMSE</b>	0,1349	0,2000	0,2698	0,3099	0,4393	0,4886	0,5044	0,4914
	<b>nRMSE</b>	0,2321	0,3497	0,4721	0,5410	0,6014	0,6749	0,6860	0,6934
	<b>MAE</b>	0,0758	0,1177	0,1181	0,2010	0,3030	0,3901	0,3953	0,3713
	<b>nMAE</b>	0,1305	0,2058	0,2066	0,3509	0,4148	0,5389	0,5376	0,5240
<b>ANN</b>	<b>RMSE</b>	0,1214	0,1804	0,1795	0,2592	0,2569	0,2885	0,3634	0,4241
	<b>nRMSE</b>	0,2088	0,3155	0,3141	0,4527	0,3518	0,3985	0,4942	0,5984
	<b>MAE</b>	0,0737	0,1110	0,1199	0,1661	0,1719	0,1998	0,2602	0,3218
	<b>nMAE</b>	0,1268	0,1941	0,2099	0,2902	0,2354	0,2760	0,3538	0,4541
<b>BO</b>	<b>RMSE</b>	0,1213	0,1554	0,1640	0,1714	0,2225	0,2392	0,2747	0,3602
	<b>nRMSE</b>	0,2085	0,2716	0,2871	0,2993	0,3046	0,3304	0,3736	0,5083
	<b>MAE</b>	0,0732	0,1014	0,1099	0,1160	0,1477	0,1633	0,1870	0,2662
	<b>nMAE</b>	0,1259	0,1773	0,1924	0,2026	0,2023	0,2256	0,2543	0,3756

Quanto ao MAE e ao nMAE, os valores dessas métricas de erro cresceram com o aumento do horizonte temporal e melhoraram com a adição do preditor ONI, exceto para as previsões de 2 min e 10 min pelo método de Boosting, no qual ocorreu o contrário.

(a) ANN



(b) Boosting

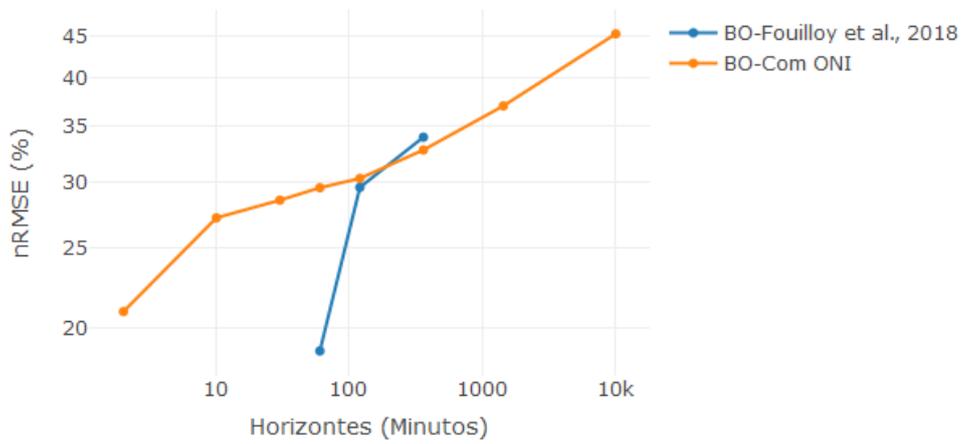


Figura 3 - Comparação dos valores de nRMSE com a literatura (eixos em escala logarítmica).

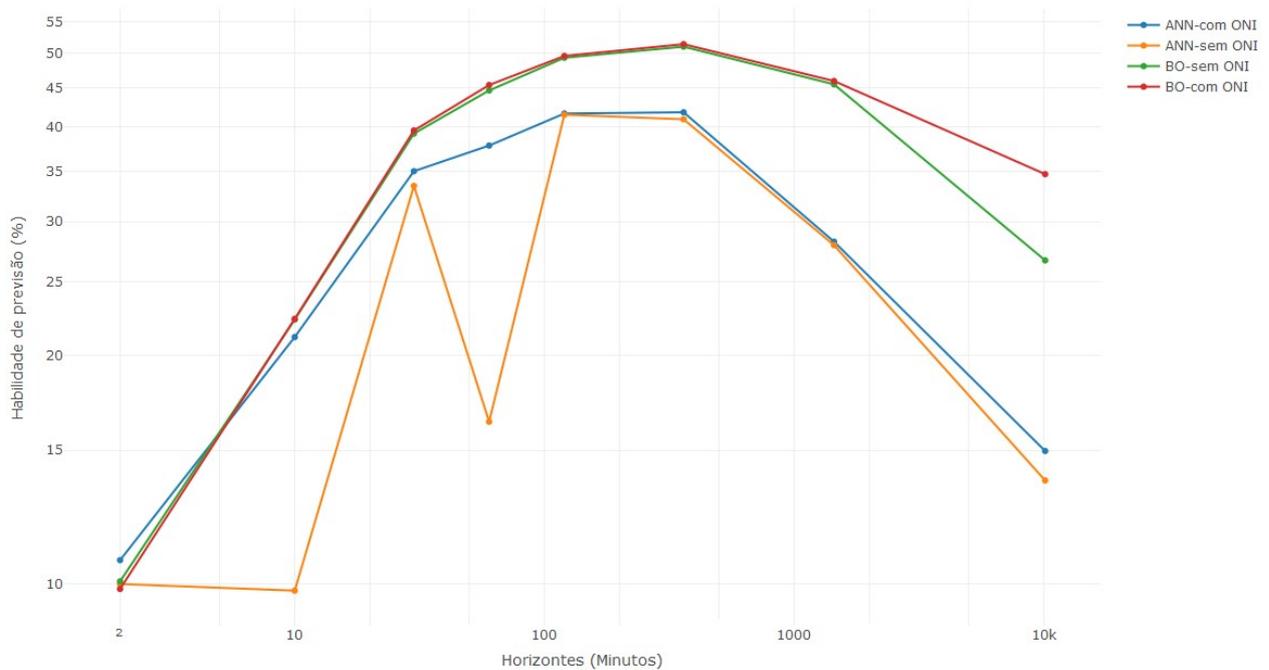


Figura 4 - Valores de S para os modelos ANN e Boosting com e sem a presença do preditor ONI (eixos em escala logarítmica).

#### 4. CONCLUSÃO

Obteve-se a variabilidade de 0,1479 para os dados de índice de claridade, classificando-os como de fraca variabilidade. Realizaram-se previsões da irradiação solar global horizontal para 8 horizontes temporais utilizando os métodos Boosting, ANN e persistência. A adição da intensidade do El Niño e La Niña, ONI, como preditor melhorou a performance dos modelos, sendo o algoritmo de Boosting o método que melhor se adequou aos dados em estudo. Tanto o Boosting quanto a ANN apresentaram erros menores do que os do modelo de persistência. Verificou-se que quanto mais distante o horizonte temporal maior o erro para todos os métodos, com e sem o preditor ONI, exceto para as previsões de 2 min com Boosting, que apresentaram piora com a adição deste preditor.

A diferença entre o erro obtido no Boosting e na ANN aumentou com a distância do horizonte temporal e o efeito da adição do preditor ONI se mostrou mais significativo para a previsão de 7 dias.

Comparando com trabalhos já publicados, verificou-se que para o Boosting as previsões de 1h e 2h apresentaram valores de nRMSE que se aproximam dos valores encontrados na literatura, alcançando um valor de erro menor do que o da literatura para a previsão de 6h. Quanto à ANN, os valores de erro obtidos neste trabalho ainda estão altos se comparados àqueles obtidos em trabalhos de outros autores. Utilizar um algoritmo de treinamento para selecionar os parâmetros da rede neural pode ser uma alternativa para reduzir os valores de erro.

#### *Agradecimentos*

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001 e com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil.

#### REFERÊNCIAS

- Benali L, Notton G, Fouilloy A, Voyant C, Dizene R. Solar radiation forecasting using artificial neural network and random forest methods: applications to normal beam, horizontal diffuse and global components. *Renewable Energy* 2019; 132:871-884.doi:10.1016/j.renene.2018.08.044.
- Duffie J.A., Beckman W.A., 2006. *Solar Engineering of Thermal Processes*. John Wiley & Son.
- Fouilloy A., Voyant C., Notton G., Motte F., Paoli C., Nivet M.L., Guillot E., Duchaud J.L., 2018. Solar irradiation prediction with machine learning: Forecasting models selection method depending on weather variability. *Energy*, vol. 165, pp. 620-629, doi:10.1016/j.energy.2018.09.116.
- Inman R.H., Pedro H.T.C., Coimbra C.F.M., 2013. Solar forecasting methods for renewable energy integration. *Progress in Energy and Combustion Science*, vol. 39, pp. 535-576, doi:10.1016/j.peccs.2013.06.002.
- Kayano M.T., Andreoli R.V., Souza R.A.F., Garcia S.R., Calheiros A.J.P., 2016. El niño e la niña nos últimos 30 anos: diferentes tipos. *Revista Climanalise*, Instituto Nacional de Pesquisas Espaciais, INPE.
- Li M.F. et al, 2013. General models for estimating daily global solar radiation for different solar radiation zones in mainland China. *Energy Conversion and Management*, vol. 70, pp. 139-148, doi:10.1016/j.enconman.2013.03.004.
- Menezes O.L., Oliveira J.L., Vasconcelos F.C., Pereira J.M.R., Sakamoto M.S., 2011. Atlas solarimétrico do Ceará 1963-2010. Fundação Cearense de Meteorologia e Recursos Hídricos-Funceme.
- Mohammadi K., Goudarzi N., 2018. Study of inter-correlations of solar radiation, wind speed and precipitation under the influence of el niño southern oscillation (ENSO) in California. *Renewable Energy*, vol. 120, pp. 190-200, doi:10.1016/j.renene.2017.12.069.
- Notton G. Nivet M.L., Voyant C., Paoli C., Darras C., Motte F., Fouilloy A., 2018. Intermittent and stochastic character of renewable energy sources: consequences, cost of intermittence and benefit of forecasting, *Renewable and Sustainable Energy Reviews*, vol. 87, pp. 96-105, doi:10.1016/j.rser.2018.02.007.
- Soares F. K., Loureiro M. H. C., Matos R. A. S., 2019, Relatório Síntese do Balanço Energético Nacional 2019, Empresa de Pesquisa Energética - EPE, Ministério de Minas e Energia, Rio de Janeiro. Disponível em: <http://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2019> (acesso em: 04. nov. 2019).
- Voyant C., Soubdahan T., Lauret P., David M., Muselli M., 2015. Statistical parameters as a means to a priori assess the accuracy of solar forecasting models. *Energy*, vol. 90, pp. 671-679, doi:10.1016/j.energy.2015.07.089.

Zhu T., Xie L., Wei H., Wang H., Zhao X., Zhang K., 2019. Clear-sky direct normal irradiance estimation based on adjustable inputs and error correction. *Journal of Renewable and Sustainable Energy*, vol. 11, doi:10.1063/1.5094808.

### **INFLUENCE OF THE PREDICTOR OCEANIC NIÑO INDEX (ONI) IN THE HORIZONTAL GLOBAL IRRADIATION FORECASTING OF FORTALEZA IN DIFFERENT TIME HORIZONS USING MACHINE LEARNING**

**Abstract.** *The introduction of solar energy into the electricity grid has been increasing annually, including in Ceará. In this context, questions are raised regarding the complementarity of supply, since the solar source is intermittent and variable over time. Proper planning can be facilitated by predicting irradiation using machine learning methods. The present work evaluated the performance of 3 forecasting models of horizontal global irradiation - neural network, Boosting and persistence model - for the city of Fortaleza, Ceará, in 8 different time horizons, analyzing the influence of El Niño and La Niña, in the form of the ONI predictor, in these forecasting. In addition to ONI, meteorological information (ambient temperature, relative humidity, wind speed, wind direction and precipitation level), irradiation data, time and date of information acquisition were used. The database was separated into 2 groups, training and testing, as well as a selection of parameters for the neural network and for the Boosting. The variability calculation of irradiation data classified this predictor as one of weak variability. The results showed that there was a reduction in the RMSE when the predictor referring to the intensity of El Niño, ONI, was added, except for the 2 min horizon, which indicates that the addition of the predictor improved the performance of the forecast models. It was also noted that the nRMSE grew as the time horizon increased. Boosting presented the lowest errors among the considered models, its values of nRMSE are compatible with those found in the literature, getting to be smaller for a determined time horizon. For the neural network, however, nRMSE values are higher than those found in works already published.*

**Key words:** *Machine Learning, Global Horizontal Irradiation, El Niño.*