

Estimativa de Irradiância Direta Normal Utilizando Redes Neurais Artificiais

Jun Fujise (Votalia) - j.fujise@votalia.com

Andreas Konrad Blatt (Votalia) - a.blatt@votalia.com

Resumo:

Modelos de decomposição da irradiância global horizontal em suas componentes difusa e direta são amplamente utilizados para situações em que não se têm medições locais, tornando-se uma alternativa barata para a estimativa de geração de energia de uma usina fotovoltaica. O presente trabalho buscou desenvolver e avaliar uma metodologia de estimação de irradiância direta normal a partir de medições da componente global horizontal utilizando o método das redes neurais artificiais. Diversas configurações de rede foram analisadas, variando-se os parâmetros de entrada e a estrutura da rede. Para a validação do modelo proposto, foram escolhidos dois locais para serem analisados, calculando-se o erro médio anual e raiz do erro quadrático médio da componente estimada e comparando com os resultados de outros modelos de decomposição (ERBS, DISC e DIRINT). Os erros obtidos com a metodologia proposta foram em geral mais baixos que os modelos de comparação, indicando uma alternativa mais acurada para a estimativa da irradiância direta normal. Por fim, os dados estimados de todos os modelos foram inseridos no software PVsyst para a realização da modelagem energética. Notou-se que os modelos propostos de rede neural artificial resultaram em desvios absolutos de produção específica de energia menores que 1%, enquanto para os outros modelos, em geral, ficaram acima deste valor.

Palavras-chave: *Energia Solar, Modelo de Decomposição, Redes Neurais Artificiais.*

Área temática: *Radiação Solar*

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

ESTIMATIVA DE IRRADIÂNCIA DIRETA NORMAL UTILIZANDO REDES NEURAIS ARTIFICIAIS

Jun Fujise – j.fujise@votalia.com
Andreas Blatt – a.blatt@votalia.com
Votalia Energia do Brasil

Resumo. Modelos de decomposição da irradiância global horizontal em suas componentes difusa e direta são amplamente utilizados para situações em que não se têm medições locais, tornando-se uma alternativa barata para a estimativa de geração de energia de uma usina fotovoltaica. O presente trabalho buscou desenvolver e avaliar uma metodologia de estimação de irradiância direta normal a partir de medições da componente global horizontal utilizando o método das redes neurais artificiais. Diversas configurações de rede foram analisadas, variando-se os parâmetros de entrada e a estrutura da rede. Para a validação do modelo proposto, foram escolhidos dois locais para serem analisados, calculando-se o erro médio anual e raiz do erro quadrático médio da componente estimada e comparando com os resultados de outros modelos de decomposição (ERBS, DISC e DIRINT). Os erros obtidos com a metodologia proposta foram em geral mais baixos que os modelos de comparação, indicando uma alternativa mais acurada para a estimativa da irradiância direta normal. Por fim, os dados estimados de todos os modelos foram inseridos no software PVsyst para a realização da modelagem energética. Notou-se que os modelos propostos de rede neural artificial resultaram em desvios absolutos de produção específica de energia menores que 1%, enquanto para os outros modelos, em geral, ficaram acima deste valor.

Palavras-chave: Energia Solar, Modelo de Decomposição, Redes Neurais Artificiais.

INTRODUÇÃO

Modelos de decomposição de irradiância global horizontal (IGH) para suas componentes direta e difusa possuem papel fundamental no processo de dimensionamento de usinas fotovoltaicas (FV). Sabe-se que utilizando os mesmos dados de IGH, pode-se obter valores de irradiância direta normal (IDN) e difusa horizontal (IDH) variando até 10 p.p. utilizando diferentes modelos da literatura (Lave et al., 2015). Isto representa uma variação considerável para a geração de sistemas FV, e, assim, mostra-se importante as medições das diferentes componentes no local de implantação destas usinas. No entanto, é comum a falta de dados disponíveis de IDN e IDH, assim, os modelos de decomposição são úteis para realizar estimativas das componentes quando só há medições de IGH.

A utilização de modelos de redes neurais artificiais (RNA) vem tornando-se comum nas diversas problemáticas de regressão e classificação. As RNAs são baseadas no sistema nervoso biológico, simulando a dinâmica existente entre diversos neurônios e emulando o fluxo de sinais elétricos entre estas células. Assim, um computador é capaz de repetir tarefas que são ensinadas a ele no processo de treinamento do modelo de RNA.

O presente trabalho buscou avaliar a utilização de uma RNA para estimar a IDN e comparar os resultados obtidos com dados medidos da componente. Tentou-se desenvolver um modelo que resultasse em erros menores que os modelos tradicionais de decomposição, criando-se uma alternativa mais acurada para ser utilizada em situações em que não há medições locais e reduzindo os riscos associados à incerteza da estimativa de geração de energia de usinas fotovoltaicas.

2. REDES NEURAIS ARTIFICIAIS

As RNAs são modelos computacionais que simulam a dinâmica das células nervosas biológicas em uma rede neural. O método é baseado no aprendizado de máquina, que é o processo em que o computador aprende a executar determinada tarefa sem ser explicitamente programado para isso, fazendo com que o modelo consiga reconhecer padrões e tomar decisões a partir de observações de dados, ou experiências, que são inseridas como entrada do modelo.

Uma rede é composta por múltiplos neurônios, que são unidade de processamento simples. Eles são dispostos em camadas interligadas por conexões denominadas sinapses. O modelo é composto por relacionamentos não lineares entre as unidades de processamento, portanto, conseguem resolver problemas mais complexos, sendo amplamente utilizados em problemas de classificação e regressão (Vellasco, 2007).

Os neurônios de uma RNA são baseados no modelo de McCulloch e Pitts (Haykin, 1994), em que as entradas dos neurônios podem ser descritas como um vetor x de dimensão equivalente à quantidade de parâmetros de entrada, e y sendo a saída do neurônio, conforme ilustrado na Fig. 1. As entradas são ponderadas por pesos (W), que são como a memória da

rede, já que são as componentes ajustadas durante o processo de treinamento (Velasco, 2007). u_k é a soma ponderada das entradas pelos pesos W , e f é a função de ativação, responsável por transformar o sinal recebido de forma não linear por meio de uma função pré-determinada. Por fim, b_k é o termo polarizador, que corrige o viés que possa existir na saída da função de ativação.

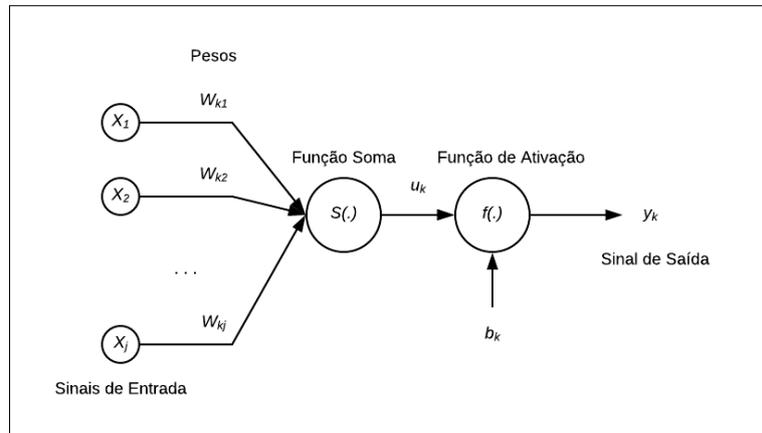


Figura 1. Modelo de neurônio de Mcculloch e Pitts (Haykin, 1994)

O fluxo de sinais que percorre um neurônio pode ser descrito pelas Eq. (1) e Eq. (2).

$$y = f(u) + b \quad (1)$$

$$u = \sum x w \quad (2)$$

Quando há apenas um neurônio na RNA, ela é chamada de *perceptron*. Já quando possui múltiplas camadas de neurônios, é nomeada *Multi Layer Perceptron* (MLP), e contém uma camada de entrada, uma ou mais camadas escondidas e uma camada de saída (Velasco, 2007).

A camada de entrada é onde os dados são recebidos pelo modelo e a camada de saída é onde é dada a resposta da rede. Por fim, as camadas escondidas são aquelas posicionadas entre a camada de entrada e a de saída, responsáveis exclusivamente pelo processamento dos sinais que fluem entre a rede, conforme mostra a Fig. 2.

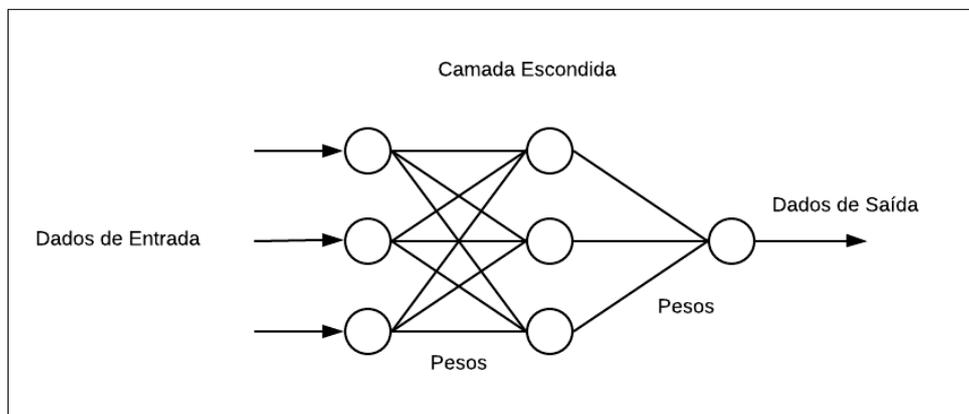


Figura 2. Modelo de RNA MLP.

Para o treinamento da rede, os pesos devem ser calibrados, utilizando-se o método de *backpropagation* (Vellasco, 2007). O processo calcula o erro na saída da rede e propaga-o pelas camadas anteriores, alterando os pesos para minimizar o erro da próxima saída. O método é baseado no algoritmo do gradiente descendente, que cria gradientes de atualizações dos pesos com base do gradiente do erro calculado.

3. METODOLOGIA

3.1 Estações Meteorológicas

Foram escolhidas cinco estações meteorológicas aleatórias no Brasil como fonte de dados para o presente estudo. Todas estão descritas na Tab. 1. Três estações foram escolhidas para compor o processo de treinamento da RNA enquanto duas foram utilizadas para a validação dos resultados. Buscou-se trabalhar com estações distantes para validar a capacidade de generalização do modelo proposto. Apenas Minas Gerais possui duas estações estudadas no presente trabalho.

Todos os pontos possuem sensores de IGH, IDH, temperatura, velocidade do vento, umidade relativa e pressão, que medem os parâmetros climatológicos integralizando-os a cada dez minutos. A IDN é obtida através das outras componentes. Foi feita uma filtragem dos dados removendo valores fisicamente inconsistentes. Para o treinamento, foram escolhidas estações com pelo menos um ano de medição.

Tabela 1. Estações meteorológicas utilizadas no presente estudo.

Estação	Local	Tempo de Medição	Uso
1	MG	20 meses	Treinamento
2	BA	17 meses	Treinamento
3	RN	15 meses	Treinamento
4	MA	9 meses	Validação
5	MG	15 meses	Validação

3.2 Composição de Modelos

Foram testados 14 arranjos de parâmetros para compor a entrada da RNA e 20 configurações de rede diferentes, variando-se o número de camadas e neurônios. Portanto, 280 combinações diferentes foram testadas a fim de otimizar o desempenho da RNA. As entradas consideradas foram:

- IGH;
- Zênite;
- Azimute;
- Componentes x e y equivalentes à hora do dia;
- Pressão;
- Temperatura ambiente;
- Velocidade do vento;
- Índice de claridade (Kt);
- Modelos de decomposição ERBS, DISC e DIRINT;

As componentes da hora do dia consideram que 24 horas equivalem a 360° , assim, pode-se converter determinada hora do dia em graus, que, posteriormente, pode ser descrito nas componentes x e y. Isso é feito pois o parâmetro Hora é cíclico, porém, seu valor é escrito de forma contínua.

O Zênite e Azimute foram estimados a partir do horário utilizando o algoritmo SPA da NREL (Reda & Andreas, 2008). Já os dados de Temperatura, Pressão e Velocidade do Vento foram medidos por sensores das estações meteorológicas analisadas. Também foram utilizados os resultados dos modelos de decomposição ERBS, DISC e DIRINT como entrada da RNA. Uma explicação mais aprofundada sobre os modelos pode ser encontrada no trabalho de Lave, Hayes, Pohl e Hansen (2015). Todos os parâmetros foram consolidados na base horária.

Os arranjos dos parâmetros de entradas estão ilustrados na Tab. 2, e os de estrutura, na Tab. 3. Na Tab. 2, cinza indica as variáveis utilizadas como entrada da rede dos diferentes modelos.

Tabela 2. Arranjos de parâmetros de entrada das RNAs analisadas.

Arranjo	IGH	Zênite	Azimute	Pressão	Umidade Relativa	Tempo(x)	Tempo (y)	Temperatura	K_t	Velocidade do Vento	ERBS	DISC	DIRINT
P1													
P2													
P3													
P4													
P5													
P6													
P7													
P8													
P9													
P10													
P11													
P12													
P13													
P14													

Tabela 3. Arranjos de estruturas das RNAs analisadas.

Estrutura	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Camada 1	10	100	100	100	100	100	10	10	10	10	30	30	30	30	30	50	50	50	50	50
Camada 2	-	-	10	100	30	50	10	100	30	50	-	10	100	30	50	-	10	100	30	50

3.2 Validação

Para validação do método, foram calculados o erro médio da estimativa anual de IDN (ME) e a raiz do erro quadrático médio da base horária de IDN em relação aos dados medidos da mesma componente (RMSE). Ambas as métricas também foram calculadas para os modelos de decomposição ERBS, DISC e DIRINT para fins comparativos.

Em seguida, os dados de IDN estimados, juntamente com dados de IGH e Temperatura Ambiental medidos, foram inseridos no software de modelagem energética PVsyst v6.84 para realizar uma avaliação das diferenças da energia estimada utilizando os diferentes modelos de decomposição. Foi considerada uma planta genérica com 29 inversores de 3,4 MW totalizando 100 MW_{ac} e 275.848 módulos de 435 W, em um total de 120 MW_p. As perdas consideradas foram fixadas para não influenciarem no resultado e a estrutura utilizada é um seguidor de um eixo na configuração 1Vx58, com orientação de azimute em 0° e inclinação a 0°.

4. RESULTADOS

Para a análise de parâmetros de entradas, os erros obtidos são ilustrados na Fig. 3 e Fig. 4. Os erros são a média de todos os modelos que utilizam o mesmo arranjo como entrada do modelo. Assim, calculou-se a média de erros entre as diferentes estruturas da rede.

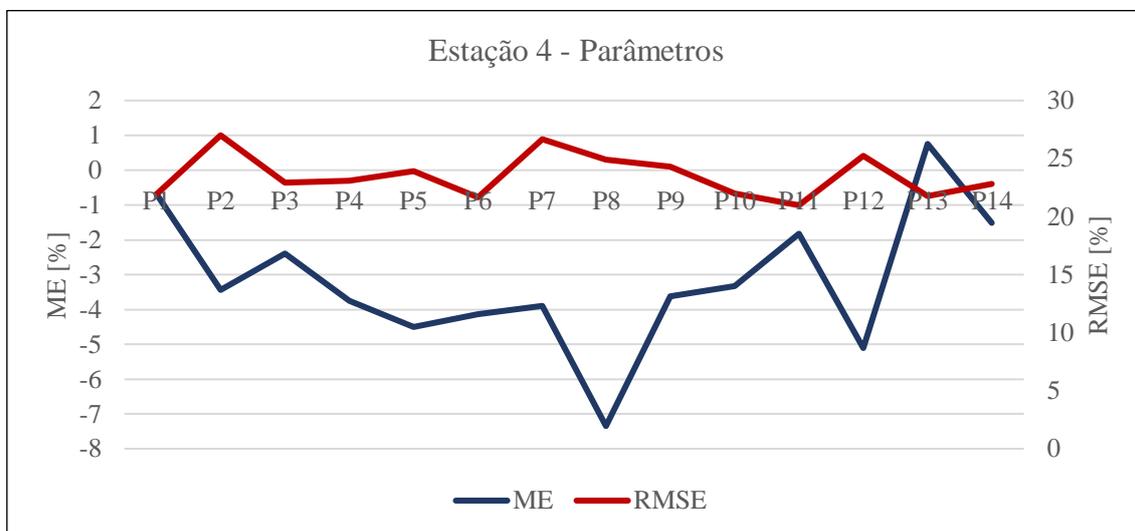


Figura 3. ME e RMSE para estação meteorológica 4 na análise de parâmetros de entrada.

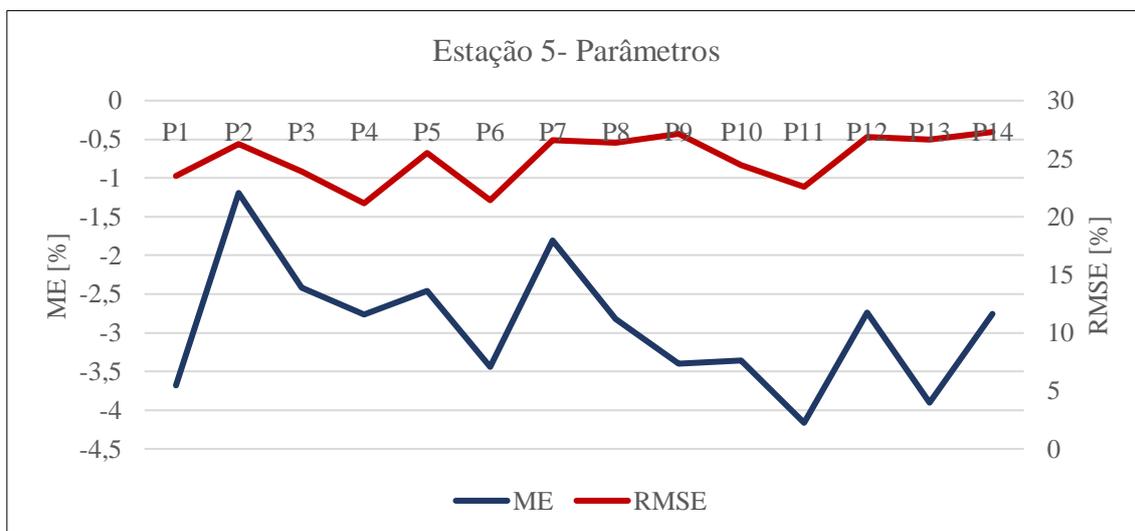


Figura 4. ME e RMSE para estação meteorológica 5 na análise de parâmetros de entrada.

Pode-se observar que o P11 foi um arranjo que obteve RMSE baixo para os dois casos, por isso, foi escolhido para as próximas análises da validação. Também foi escolhido o P3 para avaliação, já que é um arranjo mais simples que não demanda dados medidos além de IGH. Queria-se observar se um modelo mais simples teria um bom desempenho na estimativa de IDN.

Notou-se que não há uma relação forte entre arranjos de entradas e ME, já que os dois locais apresentaram resultados muito diferentes para as mesmas entradas. Já o RMSE parece ter uma correlação maior com os dados que são utilizados.

O mesmo tipo de análise foi feito para as diferentes estruturas testadas para a RNA. Os resultados são apresentados na Fig. 5 e Fig. 6.

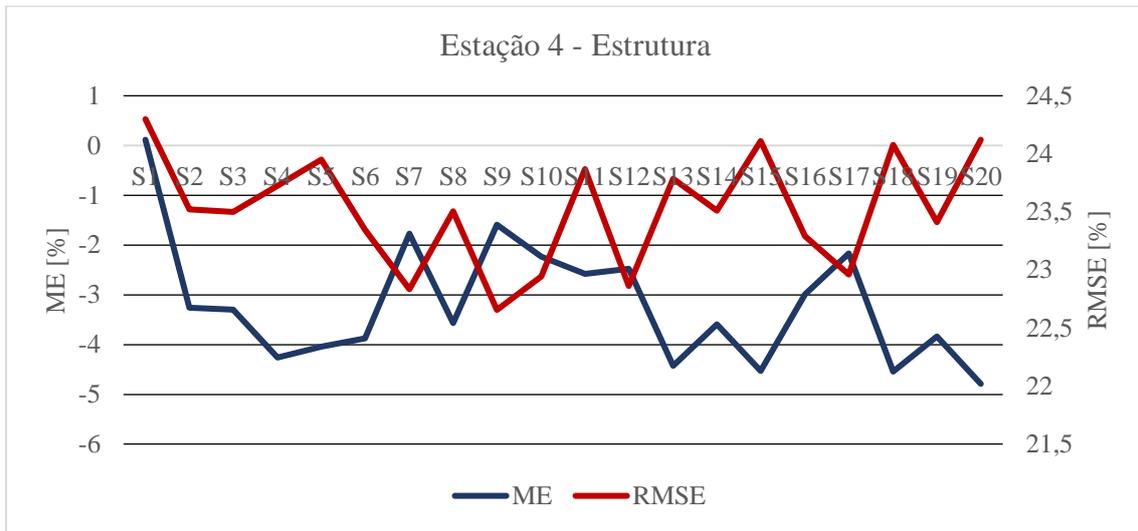


Figura 5. ME e RMSE para estação meteorológica 4 na análise de estrutura da rede.

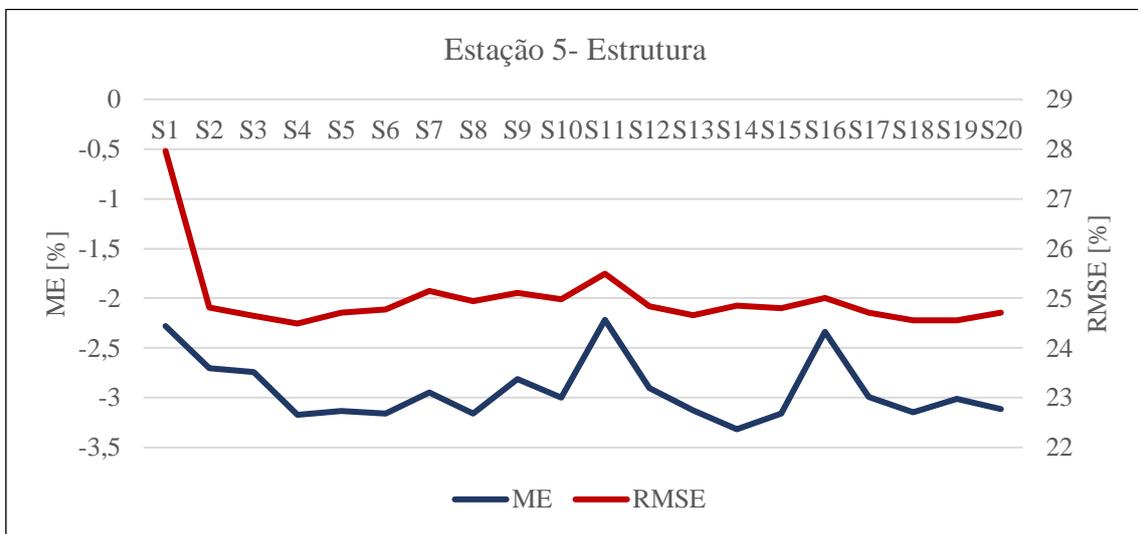


Figura 6. ME e RMSE para estação meteorológica 5 na análise de estrutura da rede.

Nota-se que a estação 4 apresentou muito mais sensibilidade ao tipo estrutura selecionada do que a estação 5. Para todas configurações, os modelos subestimaram a IDN. Foi utilizado a S9 para as próximas análises já que apresentou um dos menores RMSE para os dois casos. Assim, formou-se o Modelo A que é a combinação de P3 e S9, um modelo mais simples de RNA, e o Modelo B que é a combinação de P11 com S9.

Avaliou-se os erros entre estes dois modelos em comparação com os modelos de decomposição ERBS, DISC e DIRINT. Os resultados são apresentados na Fig. 7 e Fig. 8. Percebe-se que ambos os modelos de RNA tiveram RMSE abaixo dos outros modelos. O ME também foi baixo para estes modelos, sendo que apenas o DIRINT para estação 5 obteve um valor melhor.

Em relação a acurácia, o Modelo B teve um desempenho sutilmente melhor que o Modelo A, o que mostra que os modelos de decomposição DISC e DIRINT não tiveram um impacto significativo na melhora da performance das RNAs. Mesmo que o Modelo A só utilize o ERBS, modelo que obteve o pior desempenho analisado, estimativa de IDN teve uma melhora expressiva.

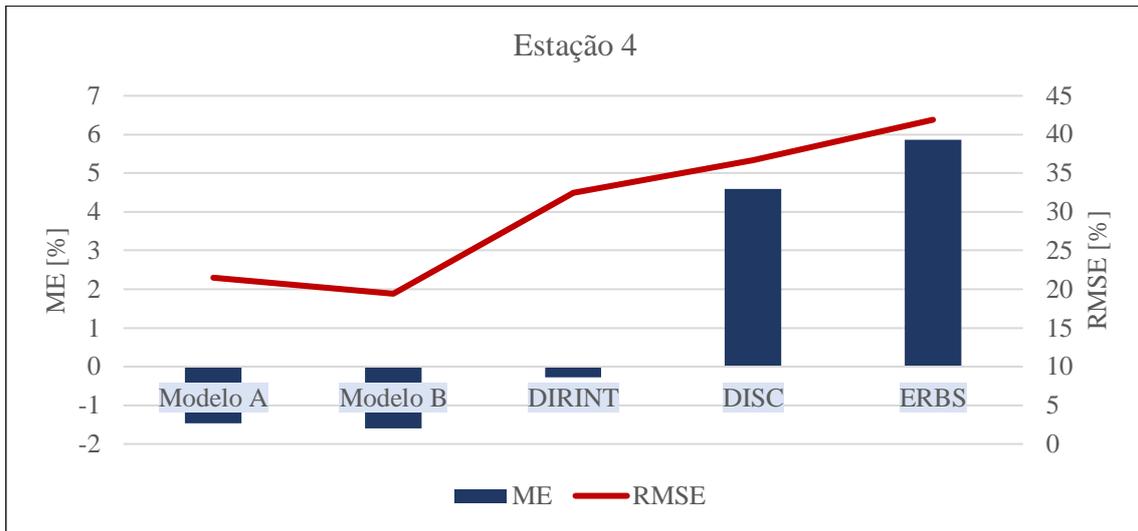


Figura 7. Comparação de ME e RMSE dos diferentes modelos de decomposição.

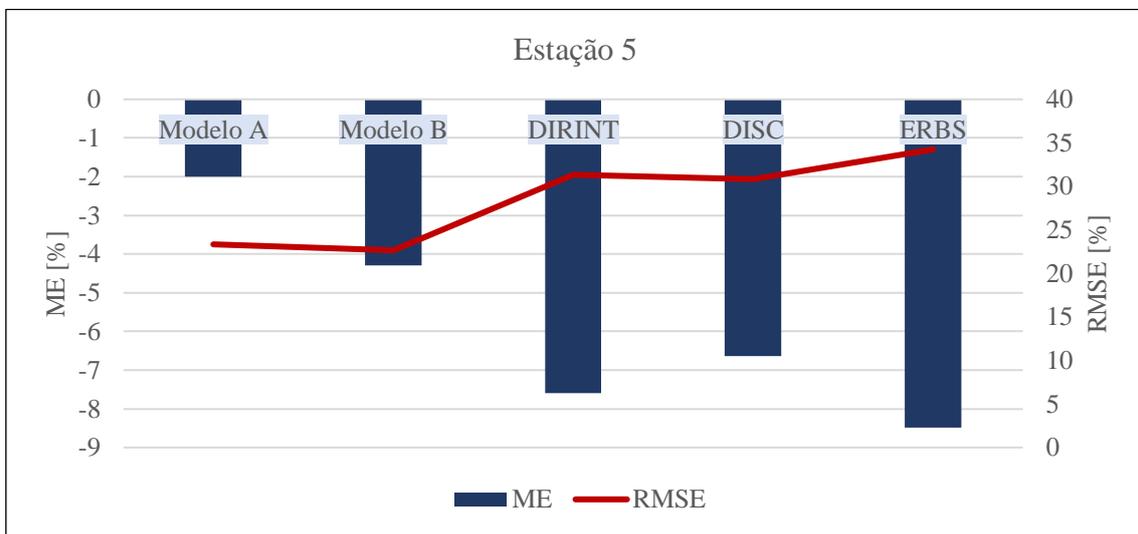


Figura 8. Comparação de ME e RMSE dos diferentes modelos de decomposição para a estação 5.

Após a análise do recurso, dados de IDN estimados, e IGH e Temperatura Ambiente medidos foram utilizados para criar diferentes arquivos meteorológicos no PVsyst. Assim, cada fonte de IDN resultou em uma simulação energética para o período concorrente, e a produção específica foi avaliada em relação a simulação feita com os dados medidos, conforme mostra a Tab. 4 e Tab. 5.

Tabela 4. Resultados das simulações no PVsyst para o local da estação 4.

Modelo	Produção Específica	Desvio da Produção Específica [%]
ERBS	1508	1.55
DIRINT	1506	1.41
DISC	1522	2.49
MODELO A	1498	0.88
MODELO B	1493	0.54
Medido	1485	-

Tabela 5. Resultados das simulações no PVsyst para o local da estação 5.

Modelo	Produção Específica [kWh/kWp/ano]	Desvio da Produção Específica [%]
ERBS	2155	-1.37
DIRINT	2163	-1.01
DISC	2182	-0.14
MODELO A	2193	0.37
MODELO B	2177	-0.37
Medido	2185	-

Os resultados mostram que os modelos propostos no presente trabalho tiveram um desvio de produção específica menores que 1% para as duas estações analisadas. O desvio obtido na estação 5, em geral, foi menor que na estação 4, podendo ser explicado pela utilização de uma estação meteorológica do mesmo estado no processo de treinamento da RNA. Isso poderia ter facilitado o treinamento da rede já que ela teria aprendido como as dinâmicas atmosféricas na região ocorrem.

Nota-se que não há relação linear entre o ME na análise do recurso solar e o desvio da produção específica. Isso ocorre já que o ME calculado é uma generalização temporal no período de um ano, enquanto a modelagem energética é dada na base horária pelo PVsyst, para só, posteriormente, consolidar-se a geração para a base anual.

5. CONCLUSÃO

O presente trabalho propôs-se a avaliar a utilização do método das RNAs para realizar a estimativa de IDN. Para tal, diversas configurações de redes foram estudadas a fim de obter-se um arranjo otimizado de parâmetros de entrada e estrutura da rede. Após avaliar 280 combinações para cada local de estudo, achou-se uma configuração que obteve boa performance para os dois casos e estabilidade nos resultados. Na maioria dos casos, os modelos foram melhores que os modelos tradicionais de decomposição.

A grande limitação na metodologia apresentada é a necessidade de dados de IDN medidos de alguns locais para realizar o treinamento da rede. No presente estudo, apenas três estações foram utilizadas, e acredita-se que caso mais estações fossem utilizadas, melhor ficaria o modelo. Uma análise de sensibilidade dos resultados com a quantidade de amostras para o treinamento pode ser feita em futuros estudos para a avaliação da incerteza desta metodologia.

Notou-se que utilizar dados climatológicos medidos não melhorou significativamente o desempenho das redes. No entanto, houve uma expressiva diferença entre o modelo ERBS e o Modelo A que pode ser considerado como um ajuste do primeiro modelo.

Portanto, apesar da necessidade da disponibilidade de IDN medida, é possível criar um modelo utilizando RNA para realizar estimações da componente para outros locais. O método proposto apresentou em geral erros menores que os modelos ERBS, DISC e DIRNT, podendo ser utilizado na realização de estudos energéticos quando não se há medições de IDN no local de análise. Recomenda-se que estudos com mais pontos de validação sejam feitos no futuro a fim de garantir-se a consistência da metodologia.

REFERÊNCIAS

- HAYKIN, S., Neural networks - a comprehensive foundation. Macmillan College Publishing Company, USA, NY – New York, 1994.
- LAVE, HAYES, POHL, HANSEN, Evaluation of global horizontal irradiance to plane-of-array irradiance models at locations across the United States. IEEE Journal of Photovoltaics, v.5, i. 2, 2015.
- REDA, ANDREAS, Solar position algorithm for solar radiation applications. Solar Energy, v. 76, n. 5, pp. 577-589, 2004.
- VELLASCO, Redes neurais artificiais, 2007. Disponível em: <<http://www2.ica.ele.puc-rio.br/Downloads/33/ICA-introdu%C3%A7%C3%A3o%20RNs.pdf>>.

ESTIMATION OF THE DIRECT NORMAL IRRADIANCE USING ARTIFICIAL NEURAL NETWORKS

Abstract. *Models decomposing the global horizontal irradiance into diffuse and direct components are widely used in situations where local measurements are not available, being a low-cost alternative for the estimation of the energy production of a photovoltaic power plant. This study developed and assessed a methodology to calculate the direct normal irradiance from the global horizontal component using artificial neural network methodologies. Several arrangements of networks were evaluated, varying the input parameters and the net structure of each one. To validate the proposed model, two on-site measurement masts were analyzed. The annual mean error and the root mean squared error of the estimated component were calculated, and the results were compared with commonly used decomposition models (ERBS, DISC and DIRINT). The obtained errors using the proposed methodology were generally lower than the errors from the compared models, which indicates an alternative estimation of the direct normal irradiance. Lastly, all data were inserted in PVsyst to simulate the generated energy in a generic PV plant using different calculations of irradiance data. It was noticed that the proposed artificial neural network models resulted in deviations of specific production of energy lower than 1% whereas the spread seen within traditional models was higher.*

Key words: *Solar Energy, Decomposition Model, Artificial Neural Network.*