

# METODOLOGIA BASEADA EM REDES NEURAIIS ARTIFICIAIS PARA PREVISÃO DA GERAÇÃO DE PLANTAS FOTOVOLTAICAS

Mario da Silva Araújo - marioaraujo@alu.ufc.br

Levi Maia Alves - levimaia2009@alu.ufc.br

Paulo Cesar Marques de Carvalho – carvalho@dee.ufc.br

Universidade Federal do Ceará, Departamento de Engenharia Elétrica

Tatiane Carolyne Carneiro – tatiane.carolyne@ufma.br

Universidade Federal do Maranhão, Curso de Bacharelado em Engenharia Ambiental

**Resumo.** *Previsões precisas da geração fotovoltaica (GFV) foram investigadas como significativas redutoras dos impactos que fontes intermitentes de geração elétrica fornecem aos sistemas elétricos. A previsão da GFV envolve incertezas que estão relacionadas às características das séries temporais do recurso solar e sua alta volatilidade devido à dependência de muitas condições do tempo. O presente artigo apresenta um estudo comparativo entre diferentes aplicações de Redes Neurais Artificiais (RNA) para a previsão horária da geração em duas plantas fotovoltaicas situadas no Laboratório de Energias Alternativas (LEA) da Universidade Federal do Ceará (UFC). No presente artigo, foram usadas duas configurações: a) RNA do tipo Perceptron com Múltiplas Camadas (MLP); e b) Rede Neural de Kohonen (SOM), ambas, treinadas com dados históricos de potência (W). Métricas de análise de desempenho foram aplicadas para avaliação e comparação das metodologias (coeficiente de correlação –  $r$ , eficiência de Nash-Sutcliffe – NSE e viés relativo – VR). Os resultados gerados pelas RNA mostraram que as aplicações com MLP apresentaram o melhor desempenho geral; as SOM obtiveram desempenho levemente abaixo e também apresentaram bom potencial para uso na previsão de geração fotovoltaica. Ambas as RNA aplicadas obtiveram correlação próxima de 90%, baixo viés e NSE em torno de 0,80 e de 0,90 em algumas aplicações.*

**Palavras-chave:** Geração fotovoltaica, Redes Neurais Artificiais, Previsão Solar

## 1. INTRODUÇÃO

As fontes de geração elétrica baseadas em fontes intermitentes têm expandido sua participação nas matrizes elétricas no mundo. Esse aumento foi impulsionado, principalmente, pelo aumento nos investimentos em pesquisa, desenvolvimento e inovação nessa área de estudo (IEA, 2020). A fonte que tem apresentado o maior incremento em termos de capacidade instalada no mundo é a fotovoltaica (FV) (IREA, 2021). Considerando o cenário brasileiro, de acordo com a ABSOLAR (2021) a contribuição FV na matriz elétrica é de 2,3%, com um total de 4.296 MW. O Nordeste brasileiro apresenta o maior potencial solar do país com valores médios diários totais de irradiação global horizontal e da componente direta normal de 5,49 kWh/m<sup>2</sup> e 5,05 kWh/m<sup>2</sup>, respectivamente (ANEEL, 2012).

A geração fotovoltaica (GFV) é caracterizada por uma intermitência que é inerente ao recurso solar e que precisa ser tratada para garantir a confiabilidade do fornecimento de eletricidade à rede (Carneiro et al., 2021). As previsões precisas da GFV podem reduzir o custo líquido da geração (Lima et al., 2020) e os impactos sobre a segurança da rede. Nos mercados de desequilíbrios de energia, a previsão solar apresenta benefícios trazendo redução na necessidade de recursos de backup e podem contribuir para a manutenção do equilíbrio entre a demanda e a oferta de energia (Wang et al., 2019, Kaur et al., 2016, Pedro et al., 2015).

Os métodos de previsão solar aprendem o comportamento das séries do recurso solar ou da GFV e implementam o conhecimento adquirido para prever valores futuros, usando dados anteriores que podem vir da própria série ou de outras séries que influenciam em seu comportamento (Carneiro et al., 2021). De acordo com Antonanzas et al., (2016) existem vários métodos aplicáveis à previsão solar, que podem ser classificados em diferentes abordagens (métodos estatísticos, métodos de aprendizado de máquina, métodos físicos e os métodos híbridos).

De acordo com MertYagli et al., (2019), os modelos baseados em aprendizado de máquina produziram uma grande quantidade de artigos na área de previsão solar nos últimos anos e se apresenta como uma abordagem popular para a previsão solar em horizonte horário. A presente aplicação tem como objetivo utilizar a abordagem de aprendizado de máquina por meio do uso e comparação de duas diferentes Redes Neurais Artificiais (RNA), os quais são o Perceptron com Múltiplas Camadas (MLP) e a Rede Neural de Kohonen (SOM), para a previsão horária da GFV de plantas instaladas na região metropolitana de Fortaleza – CE.

Além da seção introdutória, o presente artigo se organiza da seguinte forma: a Seção 2 apresenta o estado da arte em relação à aplicação de RNA para a previsão solar no mundo. A Seção 3 apresenta o embasamento teórico relacionado às RNA do tipo MLP e SOM. Seção 4 apresenta a metodologia implementada, desde a apresentação da área de estudo até o detalhamento matemático das RNA utilizadas para previsão de GFV. Seção 5 apresenta as discussões dos resultados obtidos após a implementação das RNA. Seção 6 apresenta as conclusões do estudo.

## 2. REDES NEURAIIS ARTIFICIAIS APLICADAS À PREVISÃO SOLAR

Pinheiro et al. (2017) estudaram uma RNA com treinamento *Levenberg-Marquardt* para previsão da GFV em uma planta localizada na Universidade Federal de Santa Catarina (UFSC). A RNA usou 4 entradas para receber valores de

irradiação, temperatura do módulo, temperatura ambiente e o tempo num intervalo de cinco minutos das 7h00min até 18h00min, e apresenta uma única saída a qual informa a energia produzida.

Bastos et al. (2020) utilizaram RNA recorrentes do tipo *Long-Short-Term Memory* (LSTM) para prever a corrente e a tensão da instalação composta por 4 módulos FV de 250 Wp (Universidade Federal de Alagoas). Os dados meteorológicos e de geração FV foram obtidos por uma estação solarimétrica de baixo custo, com intervalo de um segundo de uma leitura para outra. Os parâmetros utilizados para a construção desta rede foram: temperatura ambiente, temperatura do módulo, irradiância, umidade, pressão, velocidade do vento, tensão e corrente.

Barrera et al. (2020) realizaram estudos de previsão da GFV em diferentes regiões da Europa a partir de dados públicos sobre a quantidade de energia gerada, fatores solares, fatores geológicos e atmosféricos, e fatores elétricos do módulo FV. Foram utilizadas as plataformas *PVOutput* e Sistema de Informação Geográfica Fotovoltaica (PVGIS). Foram escolhidos cinco módulos FV para cada um dos seguintes países: Itália, Alemanha, Reino Unido e Bélgica. Foi coletada a localização, marca e modelo do módulo FV, dados completos durante o período de 2014 a 2017 e orientação e inclinação dos módulos FV para serem utilizados nas duas plataformas e, assim, obter os dados para treinamento da rede.

Em Vaz et al., (2016) uma estratégia de previsão para a produção de eletricidade de uma planta FV em horizonte horário e mensal utiliza uma RNA usando o modelo Não Linear Autorregressivo com Entradas Exógenas (NARX). Na aplicação, os autores utilizam dados da usina, de usinas vizinhas e dados meteorológicos. Uma modelagem para previsão da GFV para um dia e uma hora à frente é encontrada em Hossain et al. (2017) baseada na abordagem *Extreme Learning Machine* (ELM). O modelo de previsão proposto é treinado e testado utilizando dados de uma usina FV no telhado de uma Universidade na Malásia. Além dos dados de geração, os autores utilizaram parâmetros meteorológicos.

Jung et al., (2020) aplicam um modelo *recurrent neural network* (RNN) com memória de longo prazo para prever GFV na Coreia do Sul. O modelo generaliza dados de 164 usinas FV distribuídas pelo país, e um único modelo de previsão é integrado, com raiz quadrada do erro quadrático médio (RMSE) de 7,705% e 7,416% e erro percentual absoluto médio (MAPE) de 11,535% e 10,805%, os quais são calculados para as fases de treinamento e teste, respectivamente.

A Tab. 1 faz um resumo das pesquisas apresentadas neste tópico, informando os parâmetros de variáveis de entrada, modelo utilizado e variáveis de saída.

Tabela 1- Estudos apresentados os quais estão relacionados ao uso de RNA para prever dados de produção de energia solar.

Estudo	Variáveis de entrada	Modelo utilizado	Variáveis de saída
(Pinheiro et al., 2017)	Parâmetros meteorológicos incluindo o Recurso solar e temperatura do módulo FV	MLP	Energia Produzida
(Bastos et al., 2020)	Parâmetros meteorológicos incluindo o Recurso solar	LSTM	Tensão e Corrente
(Barrera et al., 2020)	Parâmetros meteorológicos incluindo o Recurso solar e dados da instalação FV	MLP	Energia Produzida
(Vaz et al., 2016)	GFV de usinas FV vizinhas e dados meteorológicos	NARX	Energia Produzida
(Hossain et al., 2017)	Parâmetros meteorológicos e GFV	ELM	Energia Produzida
(Jung et al., 2020)	GFV	LSTM	Energia Produzida

### 3. EMBASAMENTO TEÓRICO

#### 3.1 *Perceptron* com Múltiplas Camadas (MLP)

O primeiro *Perceptron* foi proposto por Frank Rosenblatt, em meados de 1950, inicialmente um classificador linear, ou seja, um algoritmo que filtrava a entrada e separava-as em duas categorias com uma reta, utilizando de várias entradas por meio da Eq. (1).

$$y = \varphi (\sum \omega i * xi + b) \quad (1)$$

Onde  $\omega$  representa o vetor de pesos de entrada,  $x$  o vetor de entrada e  $b$  como o viés. Tudo aplicado numa função de ativação não linear definido por  $\varphi$ .

Esse *Perceptron* não incluía múltiplas camadas, o que impedia de a rede neural executar tarefas um pouco mais complexas como a solução de uma simples porta XOR. A não-linearidade é inerente à maioria dos problemas encontrados na natureza, e arquiteturas com características não-lineares são importantes para a resolução desses problemas. A não-linearidade foi incorporada aos modelos de RNA por meio das funções de ativação (não-lineares) aplicada aos neurônios das redes e com a inserção, em sua estrutura, de camadas consecutivas. Uma arquitetura de RNA do tipo MLP é apresentada na Fig. 1.

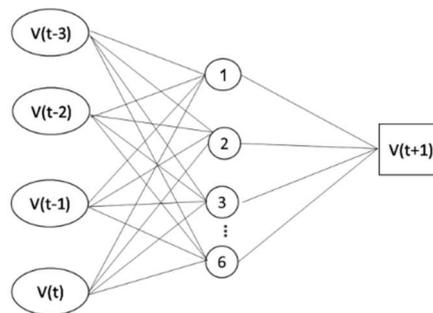


Figura 1 – Arquitetura de um *Perceptron* com duas camadas intermediárias (Carneiro et al., 2014).

Redes do tipo MLP podem ser treinadas por diversos algoritmos, mas o mais aplicado é o algoritmo de retropropagação de erros. Inicialmente, é utilizado um conjunto dos dados para treinar o modelo de acordo com os pares de entradas e saídas. Esse treinamento é composto por ajuste de pesos, vieses e parâmetros de modo a minimizar o erro (Haykin, 2001). Basicamente, a aprendizagem por retropropagação de erro é dividida em duas etapas: a) um passo para frente (*forward*): um padrão de atividade (vetor de entrada) é aplicado aos nós sensoriais da rede e seu efeito se propaga através da rede, camada por camada até gerar a saída (os pesos sinápticos da rede são todos fixos); e b) um passo para trás (*backward*): os pesos sinápticos são todos ajustados de acordo com uma regra de correção de erro. Especificamente, a resposta real da rede é subtraída de uma resposta desejada (alvo) para produzir um sinal de erro. Este sinal de erro é então propagado para trás, contra a direção das conexões sinápticas vindo daí o nome de retropropagação. Os pesos sinápticos são ajustados para fazer com que a resposta real da rede se mova para mais perto da resposta desejada, em um sentido estatístico (Haykin, 2001).

### 3.2 Redes Neurais de Kohonen

A Rede Neural de *Kohonen* ou algoritmo de *Kohonen* foi desenvolvido pelo professor e pesquisador *Teuvo Kohonen*, em 1982, com intuito de organizar dimensionalmente dados mais sofisticados em grupos menores de acordo com suas características, produzindo os mapas auto-organizáveis ou *self organizing maps (SOM)*. Os modelos SOM apresentam uma inspiração no cérebro humano, mais precisamente o córtex cerebral que possui áreas dedicadas para cada função sensorial. Os neurônios da rede de *Kohonen* também respondem a estímulos semelhantes e agrupa-os (Farias et al., 2013). A rede SOM funciona basicamente da seguinte forma: quando um padrão de entrada  $x$  é apresentado, a rede procura a unidade mais parecida com o padrão  $x$  apresentado durante o seu treinamento, a rede aumenta a semelhança do nodo escolhido e de seus vizinhos ao padrão  $x$ . Desta forma, a rede constrói mapas topológicos onde os nodos que estão topologicamente próximos respondem de forma semelhante a padrões de entrada semelhantes. (Braga, 2007)

Cada padrão de entrada apresentado à grade consiste tipicamente de uma região localizada ou "foco" de atividade contra um fundo em repouso. A localização e a natureza deste foco usualmente variam de uma realização do padrão de entrada para outra. Todos os neurônios da grade devem, portanto, ser expostos a um número suficiente de diferentes realizações do padrão de entrada para assegurar que o processo de auto-organização tenha uma chance de amadurecer apropriadamente (Haykin, 2001). O processo de aprendizagem dessa rede é baseado no aprendizado por competição e não-supervisionado (embora seja possível supervisionar), que, para cada dado de entrada, os neurônios calculam o valor da função discriminante. Essa função discriminante será o parâmetro de competição entre os neurônios de forma que o neurônio que estiver com a maior saída vencerá. Após ter sido definido o neurônio vencedor, há uma atualização nos valores dos pesos. Nesse processo, apenas o *cluster* do neurônio vencedor tem seus pesos atualizados.

O treinamento é feito seguindo os 4 processos: 1) Primeiramente é definido os pesos e parâmetros para começar o treinamento, de modo que sejam pesos diferentes; 2) Posteriormente são feitas as várias iterações em que é definido o nodo vencedor; 3) Atualiza-se os pesos desse nó e de seus vizinhos; e 4) O raio da vizinhança são decrementados até o mapa de características não mudar. A rede de *Kohonen*, portanto, apresenta a habilidade de realizar mapeamentos que agrupam os dados baseados nos padrões significativos entre os dados de entrada. É altamente recomendado para tratar problemas de escalas não lineares, como classificação de imagens e extração de características.

## 4. METODOLOGIA

### 4.3 Área de Estudo e Descrição dos Dados

As plantas FV estão localizadas no Laboratório de Energias Alternativas (LEA) da Universidade Federal do Ceará (UFC), inclinação de 10°. A primeira planta foi instalada em 2016 e possui 6 módulos FV de 250 Wp, totalizando uma potência de 1,5 kWp. Já a segunda planta possui 12 módulos de 325 Wp, distribuídos em duas *strings* com 6 módulos, totalizando uma potência de 3,9 kWp. Na Fig. 2 são apresentadas as plantas FV no LEA-UFC (Freitas Filho, 2020).



Figura 2 - Plantas FV no LEA-UFC. Fonte: Freitas Filho, 2020.

Para esta aplicação foram coletados os dados de potência FV (W) a cada 5 minutos. Os dados foram tratados com o auxílio do *MATLAB R2015b* e utilizados em horizonte horário. A operação dos módulos inicia-se a partir da presença da irradiação incidente na planta, o que acontece aproximadamente às 05h50min, e permanecem até que não tenha mais a sua presença, o que acontece aproximadamente às 18h00min. Para a aplicação os zeros são desconsiderados. Os dados da Planta 1 foram coletados para o período de dois anos (2019-2020) e para a Planta 2 os dados foram coletados em um período de um ano (outubro de 2020 – outubro de 2021).

A localidade onde o LEA está é caracterizada por uma sazonalidade bem definida, onde o primeiro semestre do ano é caracterizado pelo período mais chuvoso e o segundo semestre o período mais seco. Assim, espera-se que as plantas tenham seu pico de geração no segundo semestre do ano. As Figs. 3 e 4 apresentam representação do comportamento da potência durante o período de um dia no mês de maio/2019 (mês chuvoso) e um dia do mês de setembro/2019 (mês seco) na Planta 1. Pode-se observar que, na representação diária dos dados, o dia no mês de maio apresentou maior redução na produção diária da planta, possivelmente devido à ocorrência de chuvas.

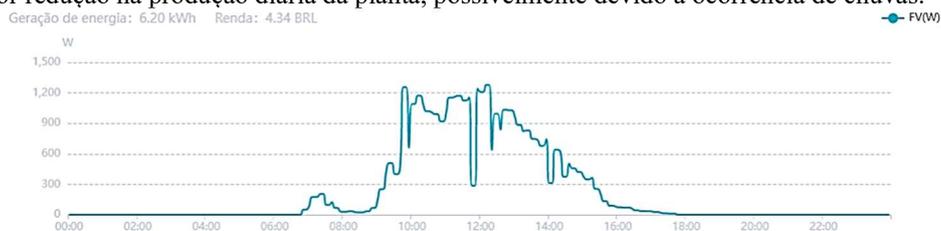


Figura 3 - Comportamento da Potência, em W, em um dia do mês de maio (Planta 1).



Figura 4 - Comportamento da Potência, em W, em um dia do mês de setembro (Planta 1).

A Figs. 5 e 6 apresentam dias representativos do comportamento da potência FV na planta 2 referentes aos meses de maio/2021 e outubro/2021. A representação diária do mês de maio demonstra maior redução da produção diária, chegando a zerar em alguns momentos do dia. A maior produção e constância são representadas pelo dia representativo do mês de outubro.



Figura 5 - Comportamento da Potência, em W, em um dia do mês de maio (Planta 2).



Figura 6 - Comportamento da Potência, em W, em um dia do mês de outubro (Planta 2).

#### 4.4 Modelo implementado

##### 4.4.1 Preparação dos dados

Os procedimentos e modelos desenvolvidos no presente artigo foram implementados usando o software *MATLAB R2015b* e também o *Microsoft excel*. Os modelos MLP e SOM são responsáveis pela estimativa de valores horários de potência (W) para as duas plantas FV mencionadas. Inicialmente, os dados são preparados; essa etapa envolve, entre outros aspectos, a forma de aquisição, o pré-processamento, a normalização e a definição dos conjuntos de treino e de teste. Os dados são normalizados utilizando a normalização linear, deixando os dados entre 0 e 1, para serem aplicados, de forma adequada, às funções de ativação. Nas duas aplicações, em ambas as plantas, os dados foram separados em 70% para o conjunto de treinamento e 30% para o conjunto de teste; os dados de teste são compostos por dados em intervalos de tempo diferentes e totalmente independente daqueles utilizados para treinamento do modelo.

Foram nomeadas as seis simulações da seguinte forma: o número 1 representa as aplicações para a planta de potência de 1,5 kWp e o número 2 representa as simulações para a planta de potência de 3,9 kWp): a) MLP#1: RNA MLP aplicada aos dados da Planta 1; b) MLP#2: RNA MLP aplicada aos dados da Planta 2; c) SOM#1: RNA SOM aplicada aos dados da Planta 1; d) SOM#2: RNA SOM aplicada aos dados da Planta 2; e) Pers#1: Persistência aplicada aos dados da Planta 1; f) Pers#2: Persistência aplicada aos dados da Planta 2.

##### 4.4.2 Perceptron com Múltiplas Camadas (MLP)

A arquitetura das duas RNA do tipo MLP é apresentada na Fig. 7, formadas por uma camada de entrada, uma camada oculta e uma de saída. Em ambos os casos, simulação MLP#1 e MLP#2, a camada inicial é composta por cinco entradas, que são as potências históricas (em W): a) P(t-4); b) P(t-3); c) P(t-2); d) P(t-1); e e) P(t). Realizando simulações com diferentes atrasos de tempo (2, 3 e 4), foram obtidos os melhores resultados com 4 atrasos de tempo, por isso a escolha de cinco entradas nas duas redes. O número de neurônios na camada oculta é obtido através do processo de tentativa e erro, sendo 6 na MLP#1 e 10 na MLP#2. A camada de saída, em ambas as simulações é composta por um único neurônio, o valor da potência P(t+1).

O treinamento da MLP foi dividido em duas etapas principais. Na etapa *forward*, os sinais são propagados da entrada para a saída, etapa onde a saída da rede é gerada e um erro é calculado. Na segunda etapa, chamada de *backward*, os erros são retropropagados da saída para a entrada e ajustados por meio da Eq. (2) (Braga et al., 2007). São utilizadas funções de ativação sigmóides para os neurônios da camada oculta e da camada de saída (Eq. (3)).

$$w_{ij}^m(t+1) = w_{ij}^m(t) - \alpha \cdot \delta_i^m \cdot y_j^{m-1} \quad (2)$$

$$\varphi[u(n)] = \frac{1}{1 + e^{-\lambda(u)}} \quad (3)$$

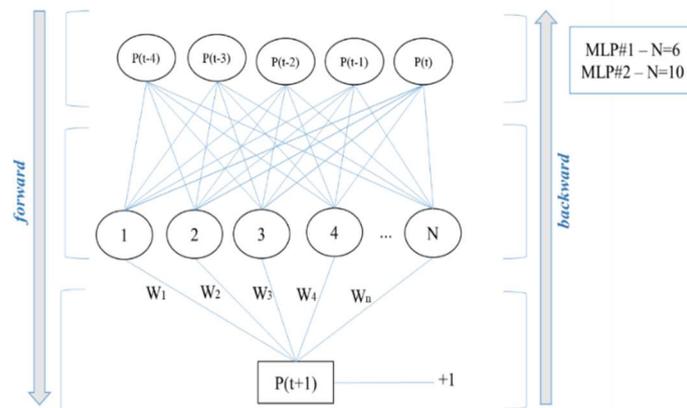


Figura 7 - Arquitetura geral da RNA MLP implementada nas simulações MLP#1 e MLP#2.

A retropropagação é feita por meio do uso do gradiente descendente para estimar o erro nas camadas intermediárias e o efeito que esses erros têm no erro da camada de saída. Após o erro da camada de saída ser calculado, na sequência, ele é retropropagado para as camadas anteriores, permitindo o ajuste dos pesos proporcionais aos valores das conexões entre as camadas (Haykin, 2001).

#### 4.4.3 Rede Neural de Kohonen ou *Self Organizing Maps* (SOM)

As Redes Neurais de Kohonen (SOM) têm como principal finalidade o agrupamento de vetores com características similares em uma mesma classe (neurônio vencedor) ou em classes semelhantes (neurônios vizinhos) (Haykin, 2001). A previsão solar realizada pela SOM é feito por meio do uso da RNA treinada como ferramenta preditiva. A modelagem completa é dividida em duas etapas principais: a) Etapa Não-supervisionada; e b) Etapa Preditiva.

Na etapa não-supervisionada, as suas arquiteturas são estabelecidas de maneira a terem uma camada de entrada (multidimensional) e uma camada competitiva ou de saída (unidimensional ou bidimensional). Na camada de saída, os neurônios concorrem de forma que, a cada iteração, um deles pode ser considerado o vencedor (a classe mais adequada para um determinado padrão de entrada). Cada entrada está conectada a todos os elementos da camada de saída, em que essas conexões são medidas por meio de pesos  $W_{ij}$  entre os neurônios de entrada  $j$  e os neurônios da camada de saída  $i$ . No presente estudo, foram realizadas duas aplicações: a SOM#1 e a SOM#2.

A SOM#1 apresenta seus vetores da camada de entrada com 6 neurônios (quatro atrasos) representando os valores de horários de potência em  $P(t-4)$ ,  $P(t-3)$ ,  $P(t-2)$ ,  $P(t-1)$ ,  $P(t)$  e  $P(t+1)$ , SOM#2 com uma camada de entrada com 5 neurônios (três atrasos), representando os valores horários de potência em  $P(t-3)$ ,  $P(t-2)$ ,  $P(t-1)$ ,  $P(t)$  e  $P(t+1)$ . Optou-se por camadas de saídas bidimensionais, sendo as camadas comportas por 100 e 25 neurônios para as SOM#1 e SOM#2, respectivamente. A Fig. 8 apresenta a estrutura que representa a rede SOM aplicada à planta FV 2. A estrutura da SOM#1 usou espaço bidimensional 10x10.

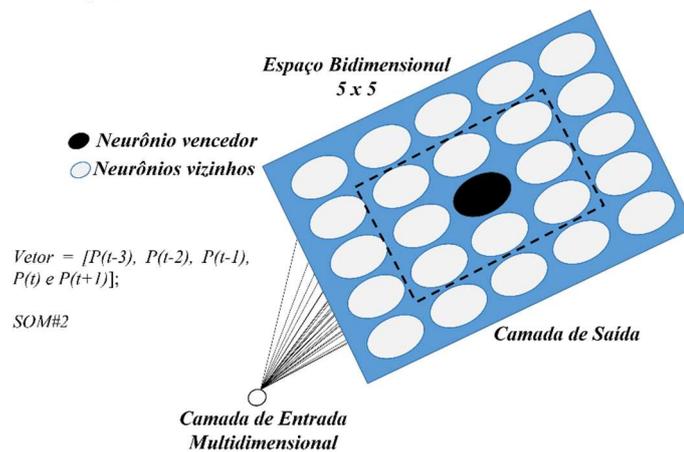


Figura 8 - Estrutura do modelo SOM e exemplo com um neurônio vencedor e seus vizinhos (arquitetura SOM#2).

O treinamento foi em modo sequencial e o conjunto de dados foi apresentado à rede neural SOM 1200 vezes. As taxas de aprendizagens utilizadas foram 0,15 e 0,25 para a SOM#1 e SOM#2, respectivamente. Na etapa de treinamento da SOM, as distâncias Euclidianas  $DI_i$  entre o vetor de entrada e os pesos ligados a cada um dos neurônios de saída é realizada por meio da Eq. (4).

$$DI_i = \sqrt{\sum_{j=1}^J (x_j - w_{ij})^2} ; \text{ para } i=1,2,\dots, M \quad (4)$$

Onde  $x_j$  é o  $j$ -ésimo componente do vetor de entrada  $x$ ;  $J$  é a dimensão do vetor de entrada  $x$ ; e  $M$  é o número total de neurônios na camada de saída.

O neurônio de saída  $i$  que possuir a menor distância Euclidiana ao vetor de entrada é considerado o neurônio vencedor. Os pesos conectados a este neurônio  $i^*$  e aos neurônios dentro de um raio de vizinhança  $V_{i^*}$  são então atualizados pela regra de Kohonen (Beale et al, 2012), conforma a Eq. (5).

$$W_{ij}(n) = W_{ij}(n-1) + \alpha \cdot [x_j(n) - W_{ij}(n-1)]; \quad (5)$$

para  $i \in V_{i^*}$  e  $j=1,2,\dots, J$

Onde  $\alpha$  é a taxa de aprendizagem e  $n$  é o índice que representa a sequência de apresentação das amostras a rede.

Essa regra faz com que os pesos ligados ao neurônio vencedor e seus vizinhos a se deslocarem na direção do vetor de entrada apresentado à rede, reduzindo a distância Euclidiana cada vez mais e fazendo com que estes neurônios aprendam a classificar vetores similares. Essa etapa, o aprendizado é não supervisionado, não

apresentando as saídas desejadas. A determinação da vizinhança pode ser realizada de diversas formas (Beale et al, 2012). Nessa aplicação, utilizamos uma função de base quadrada e a determinação da vizinhança é feita em duas etapas: a) fase de ordenação: os neurônios se organizam no espaço de entrada de forma consistente com as suas posições. Nela, o treinamento é limitado por 10 e 30% das iterações, com raio de 2 e 1 passos, respectivamente; e b) fase de ajuste (o raio de vizinhança fica na unidade): essa fase dura o restante do número de iterações (60%), ocorre a atualização somente nos pesos do neurônio vencedor. Espera-se que os pesos se modifiquem de forma relativamente uniforme no espaço de entrada, mantendo a topologia definida na fase de ordenação.

Depois de treinadas, inicia-se a etapa preditiva. As redes SOM podem ser adaptadas para sistemas de previsão, conforme aplicado em Farias et al., (2013). A previsão ocorrerá considerando o vetor de entrada com a ausência da variável a ser prevista seguindo as etapas descritas na Fig. 9.

O passo 1 do Esquema de Previsão Solar por meio de SOM treinada é possível por meio da inclusão de uma variável booleana  $m_j$  (Eq. (6)). Essa variável  $m_j$  é usada para incluir ( $m_j = 1$ ) ou excluir ( $m_j = 0$ ) a contribuição de um dado elemento  $j$  do vetor de entrada no cálculo das distâncias Euclidianas (Farias et al., 2013).

$$DI_i = \sqrt{\sum_{j=1}^J m_j (x_j - w_{ij})^2} ; \text{ para } i=1,2,\dots, M \quad (6)$$

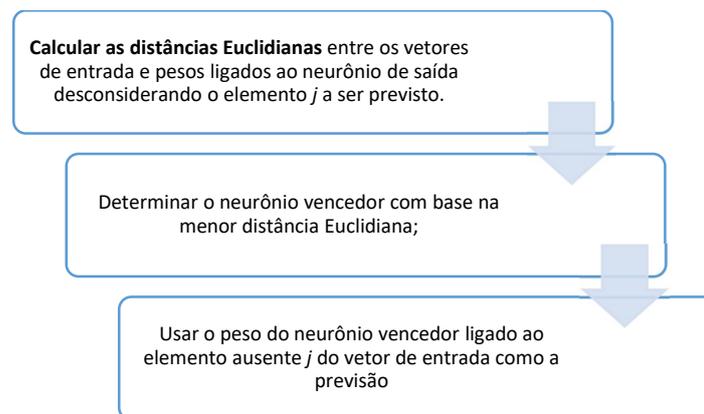


Figura 9 – Esquema da Previsão Solar por meio da SOM treinada.

As métricas de correlação ( $R$ ), viés relativo ( $VR$ ) e coeficiente de eficiência do modelo *Nash-Sutcliffe* ( $NSE$ ) estão descritas nas Eqs. de (7) a (9) (Lettenmaier & Wood, 1993) e são utilizadas para a avaliação do desempenho dos modelos implementados para estimativa da potência das duas plantas.

$$R = \frac{\sum_i^N (X_i - X_{med}) \cdot \sum_i^N (Y_i - Y_{med})}{\sqrt{\sum_i^N (X_i - X_{med})^2 \cdot \sum_i^N (Y_i - Y_{med})^2}} \quad (7)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (Y_i - X_i)^2}{\sum_{i=1}^N (X_i - X_{med})^2} \quad (8)$$

$$VR = \frac{X_{med} - Y_{med}}{Y_{med}} \quad (9)$$

Onde  $X_{med}$  é o valor médio do conjunto de dados observados,  $Y_i$  é o valor previsto,  $Y_{med}$  é o valor médio do conjunto de dados previstos e  $N$  é o número total de dados previstos. A dependência linear é medida pela  $R$ . A tendência a subestimar ou superestimar as potências observadas, se houver, é medida pelo  $VR$ . A aderência entre dados previstos e observados é definida pelo  $NSE$ . Com os valores de  $R = 1$ ,  $NSE = 1$  e  $VR = 0$ , podemos dizer que uma previsão é precisa. Assim, quanto mais próximo desses valores, melhor será a modelagem desenvolvida.

Além das RNA dos tipos SOM e MLP, para comparação, foi testado o método de Persistência, descrito em Eq. (10), o qual usa a média móvel simples em que a previsão é a média das  $N$  observações mais recentes de uma série temporal, metodologia frequentemente utilizada para previsão de curto prazo.

$$P(t) = \frac{1}{N} \sum_{i=1}^N P(t-1) \quad (10)$$

## 5. Resultados e Discussões

As Figs. 10a, 10b e 10c apresentam os resultados das previsões realizadas pelos modelos MLP#1, SOM#1 e Pers#1, respectivamente. São apresentadas as potências previstas pelos modelos e os correspondentes valores observados para a planta 1. Representa-se uma janela de 120 horas do período de testes nas referidas simulações, sendo esses dados totalmente independente daqueles utilizados para treinamento do modelo.

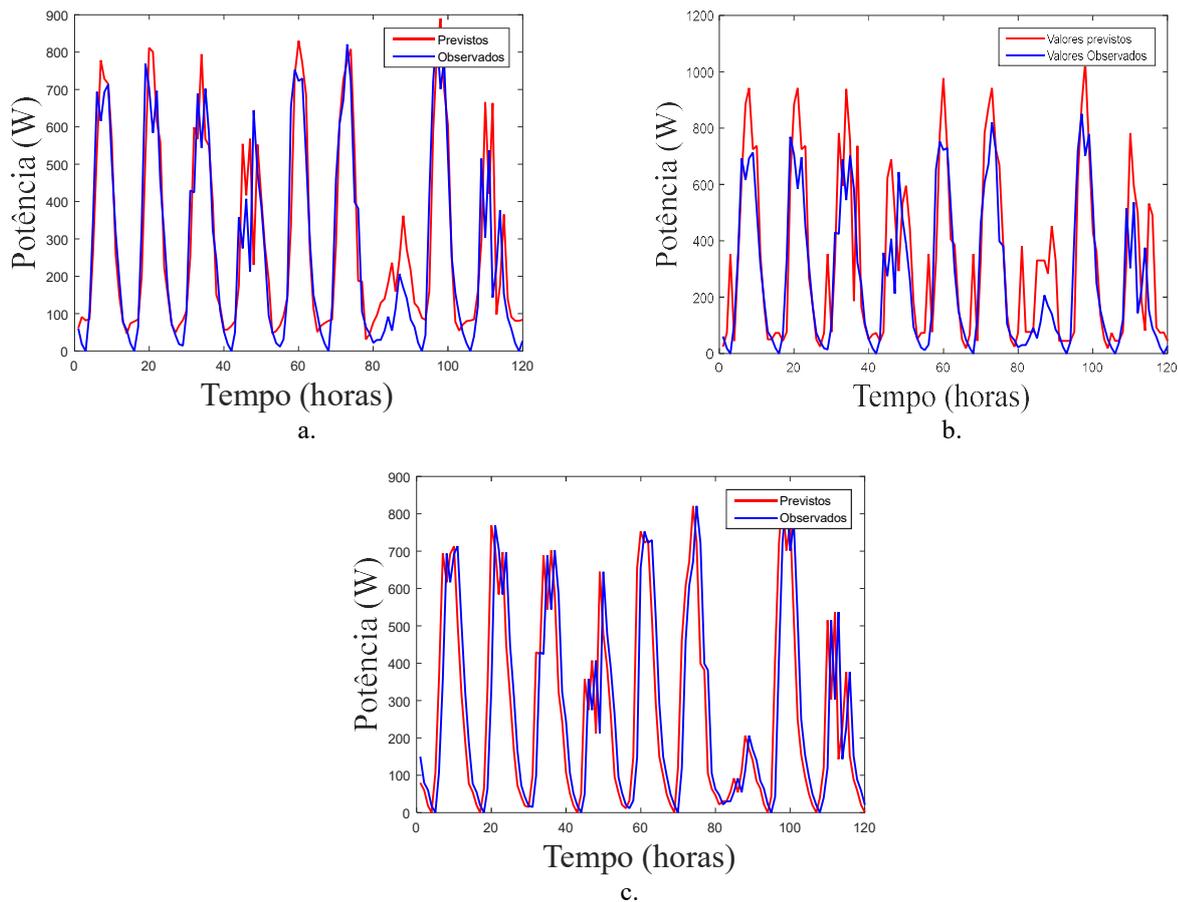
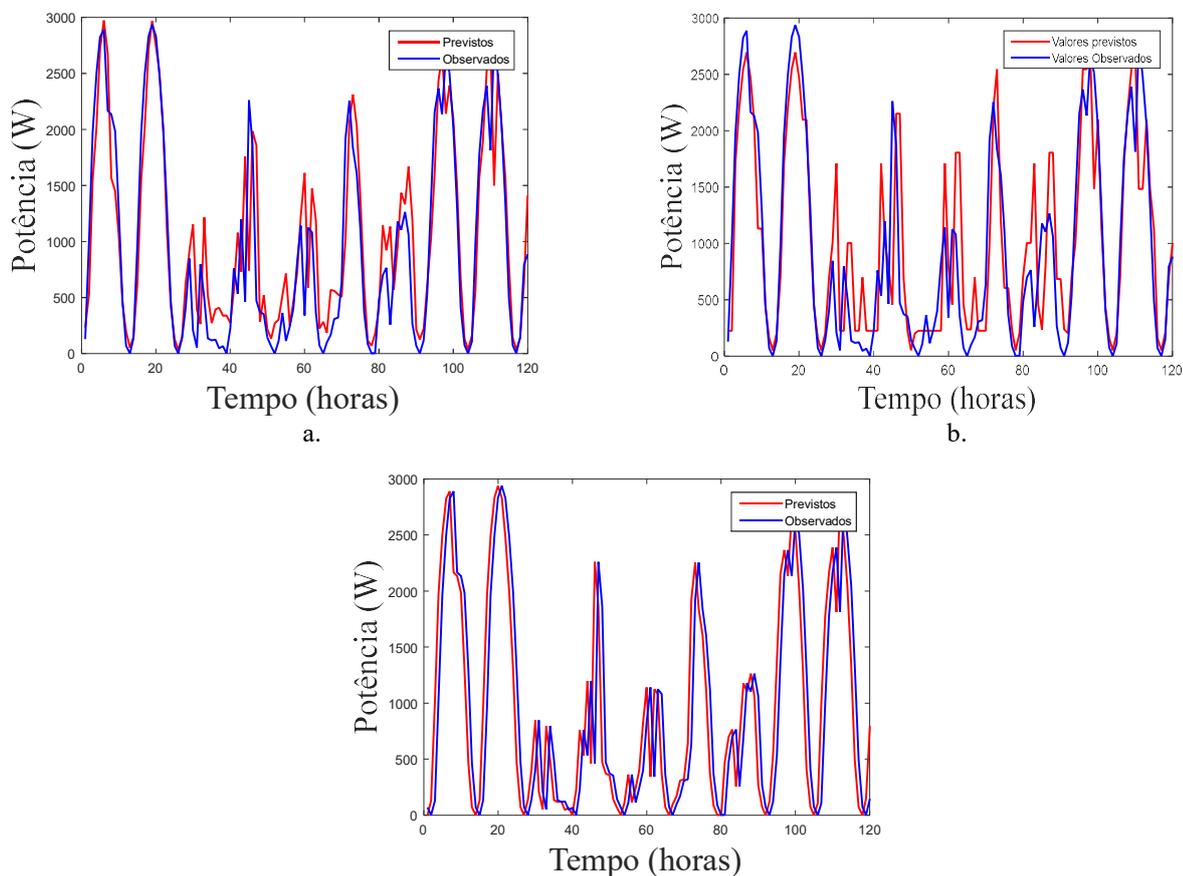


Figura 10 – Dados previstos e observados para os dados de potência da planta 1 (1,5 kWp): a. MLP#1; b. SOM#1; e c. Pers#1.

As Figs. 11a, 11b e 11c mostram um comparativo entre as potências previstas e observadas para a planta 2 e para os modelos MLP#2, SOM#2 e Pers#2, respectivamente.



c.

Figura 11 – Dados previstos e observados para os dados de potência da planta 2 (3,9 kWp): a. MLP#2; b. SOM#2; e c. Pers#2.

A Tab. 2 apresenta os resultados das métricas utilizadas para a avaliação dos resultados obtidos pelas duas aplicações com RNA do tipo MLP, SOM e pelo modelo Persistência.

Tabela 2. Métricas utilizadas para avaliação das metodologias.

	<b>NSE</b>	<b>VR</b>	<b>R</b>
<b>MLP#1</b>	0,8713	0,0104	0,9638
<b>MLP#2</b>	0,7996	0,0183	0,9249
<b>SOM#1</b>	0,8583	0,0172	0,8889
<b>SOM#2</b>	0,7828	-0,0161	0,8793
<b>Pers#1</b>	-2,2767	0,0011	-0,6390
<b>Pers#2</b>	0,6236	-2,1360	0,8310

Analisando as informações dos parâmetros R, NSE e VR, as redes MLP foram as que obtiveram melhor desempenho, mas as SOM, nas duas aplicações, também obtiveram correlação próxima de 90%, viés em torno do 0 e NSE próximo a 80%. As redes MLP já estão consolidadas na literatura e nessa aplicação demonstraram capacidade de prever com precisão dados de GFV. As RNA do tipo SOM utilizam a classificação de padrões como ferramenta preditiva e, por esse motivo, necessitam de elevada quantidade de exemplos para melhorar os seus resultados. Considerando que foram utilizados dois anos de dados para a Planta 1 (2019-2020) e apenas um ano de dados para a planta 2 (outubro/2020 – outubro/2021), isso pode justificar o rendimento da SOM#1 ter sido superior a SOM#2. No entanto, trabalhos futuros com maior quantidade de dados podem fazer as SOM superarem o desempenho das MLP. Em ambos os casos, as aplicações com o modelo tradicional Persistência apresentaram desempenho inferior às RNA aplicadas.

## 6. Conclusões

Foi proposto um estudo comparativo entre aplicações de dois tipos de RNA, a MLP e a SOM, e o modelo de Persistência para a previsão de GFV em duas plantas FV instaladas no Laboratório de Energias Alternativas (LEA), Universidade Federal do Ceará, em Fortaleza. Analisando os resultados, pode-se concluir que as RNA do tipo MLP e SOM apresentam bom potencial preditivo para os dados analisados. As MLP obtiveram melhor desempenho nas aplicações nas duas plantas, com valores de 96 e 92% de correlação, viés de 0,0184 e 0,0183 e NSE de 0,8713 e 0,7986 para as aplicações nas Plantas 1 e 2, respectivamente. As redes SOM obtiveram desempenho levemente inferior às MLP implementadas, mas também com bom desempenho, com correlação próxima de 90%, baixo viés e NSE entre 0,78 e 0,85. Analisando as aplicações por plantas, é possível observar que os resultados das simulações, independente da técnica utilizada, foram superiores quando aplicadas aos dados da planta 1, podendo ser explicados devido ao maior número de dados ou exemplos apresentados aos diferentes métodos, facilitando o aprendizado.

## Agradecimentos

Ao CNPq pela bolsa de produtividade concedida ao terceiro autor e pela bolsa de iniciação científica concedida ao primeiro autor e a participação voluntária ao segundo autor.

## REFERÊNCIAS

- AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL. Resolução Normativa ANEEL nº 482: Estabelece as condições gerais para o acesso de microgeração e minigeração distribuída aos sistemas de distribuição de energia elétrica, o sistema de compensação de energia elétrica, e dá outras providências. Brasília. 2012.
- Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de-Pison, F. J., and Antonanzas-Torres, F., 2016. Review of Photovoltaic Power Forecasting, *Solar Energy*, 136, pp. 78–111.
- Barrera J.M., Reina A., Maté A., Trujillo J.C., 2020. Solar Energy Prediction Model Based on Artificial Neural Networks and Open Data. *Sustainability*. 2020; 12(17):6915. <https://doi.org/10.3390/su12176915>
- Bastos, Ícaro G. P., De Melo, Gustavo C. G., De Araújo, Ícaro B. Q., Barboza, Erick de A., Vieira, Tiago F., Brito, Davi B., 2020. Previsão de geração fotovoltaica a partir de dados meteorológicos utilizando rede LSTM. VIII Congresso Brasileiro de Energia Solar, Fortaleza.
- Beale, M.; Hagan, M.; Demuth H. (2012). *Neural Network Toolbox 7.0.3: User’s Guide*. The MathWorks Inc, Natick, USA, 404 p.
- Braga, A. P., Carvalho, A. P. de L., Ludermir T. B., 2007. *Redes Neurais Artificiais: Teoria e Aplicações*. 2ª edição, Editora LTC, Rio de Janeiro.

- Carneiro, T. C., de Carvalho, P. C. M., Alves dos Santos, H., Lima, M. A. F. B., Braga, A. P. D. S., 2021. Review on Photovoltaic Power and Solar Resource Forecasting: Current Status and Trends. *Journal of Solar Energy Engineering*, 144(1), 010801. <https://doi.org/10.1115/1.4051652>.
- Carneiro, T. C., Santos, H. A. D., Braga, A. P. D. S., & Carvalho, P. C. M. D. (2014). Redes neurais artificiais para previsão de velocidade do vento: estudo de caso para Maracanaú-CE. *Congresso Brasileiro de Automática*.
- Farias, C. A. S.; Santos C. A. G.; Lourenço, A. M. G.; Carneiro, T. C., 2013. Kohonen Neural Networks for Rainfall-runoff Modeling: Case Study of Piancó River Basin, *Journal of Urban and Environmental Engineering*, v.7, n.1, p.176-182.
- Freitas Filho, M. F., Araujo, D. N., Carvalho, P. C. M., Sasaki J. M., 2020. Análise da composição físico-química da sujidade de plantas fotovoltaicas: estudo de caso para Fortaleza, *Revista Tecnologia*, <https://doi.org/10.5020/23180730.0.10751>.
- Haykin, S., 2001. *Redes Neurais – Princípios e Prática*, 2ª edição., Bookman.
- Hossain M., Mekhilef S., Danesh M., Olatomiwa L., Shamshirband S., 2017. Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. *Journal of Cleaner Production*; 167:395-405. <http://dx.doi.org/10.1016/j.jclepro.2017.08.081>.
- INTERNATIONAL ENERGY AGENCY (IEA). *Key World Statistics 2020*. Disponível em: <<https://www.iea.org/reports/key-world-energy-statistics-2020>>. Acesso em: 15 de julho de 2021.
- INTERNATIONAL RENEWABLE ENERGY AGENCY (IREA). *Solar Energy Data*. Disponível em: <<https://www.irena.org/solar>>. Acesso em: 15 de outubro de 2021.
- Jung Y., Jung J., Kim B., Han S., 2020. Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea. *Journal of Cleaner Production*; <https://doi.org/10.1016/j.jclepro.2019.119476>.
- Kaur, A., Nonnenmacher, L., Pedro, H. T. C., and Coimbra, C. F. M., 2016. Benefits of Solar Forecasting for Energy Imbalance Markets, *Renewable Energy*, 86, pp. 819–830.
- Lettenmaier D. P., wood, E. F. “Hydrologic Forecasting” in *Handbook of Hydrology*. Org. by Maidment, D. R., McGraw-Hill Inc., New York, USA, pp. 26.1–26.30, 1993.
- Lima, M. A. F. B., Carvalho, P. C. M., Fernández-Ramírez, L. M., and Braga, A. P. S., 2020. Improving Solar Forecasting Using Deep Learning and Portfolio Theory Integration, *Energy*, 195, p. 117016.
- MertYagli, G., Yang, D., and Srinivasan, D., 2019. Automatic Hourly Solar Forecasting Using Machine Learning Models. *Renewable Sustainable Energy Rev.*, 105, pp. 487–498.
- Pedro, H. T. C., and Coimbra, C. F. M., 2015. Nearest-Neighbor Methodology for Prediction of Intra-Hour Global Horizontal and Direct Normal Irradiances, *Renewable Energy*, 80, pp. 770–782.
- Pinheiro, Elisângela, Adalberto, Lovato, Ruther, Ricardo. Aplicabilidade de redes neurais artificiais para análise de geração de energia de um sistema fotovoltaico conectado à rede elétrica. *Revista Brasileira de Energias Renováveis*, v.7, n.5, pp. 875-892, 2017.
- Vaz A.G.R., Elsinga B., van Sark W.G.J.H.M., Brito M.C., 2016. An artificial neural network to assess the impact of neighboring photovoltaic systems in power forecasting in Utrecht, the Netherlands. *Renewable Energy*; 85:631-641. <http://dx.doi.org/10.1016/j.renene.2015.06.061>.
- Wang, G. C., Ratnam, E., Haghi, H. V., and Kleissl, J., 2019. Corrective Receding Horizon EV Charge Scheduling Using Short-Term Solar Forecasting, *Renewable Energy*, 130, pp. 1146–1158.

## METHODOLOGY BASED ON ARTIFICIAL NEURAL NETWORKS FOR PHOTOVOLTAIC PLANT GENERATION FORECAST

**Abstract.** *Precise predictions of photovoltaic generation (GPV) were investigated as significant reducers of the impacts that intermittent sources of electrical generation provide to electrical systems. The GPV forecast involves uncertainties that are related to the characteristics of the time series of solar resources and their high volatility due to their dependence on many climatic conditions. Our paper presents a comparative study between different applications of Artificial Neural Networks (ANN) applied to the hourly forecast of solar generation in two photovoltaic plants in the Alternative Energy Laboratory (LEA) of the Federal University of Ceara (UFC). In our study, two configurations were used: a) Multilayer Perceptron (MLP); and b) Kohonen Neural Network (KNN), both trained with historical power data (W). Performance analysis metrics were applied to evaluate and compare the methodologies (correlation coefficient - r, Nash-Sutcliffe efficiency - NSE and relative bias - VR). The RNA results showed that applications with MLP had the best overall performance; SOM had a slightly lower performance and also presented good potential for use in forecasting photovoltaic generation. Both ANNs applied had correlation close to 90%, low bias and NSE around 0.80 and 0.90 in some applications.*

**Key words:** *PV power, Artificial Neural Networks, Solar forecast*