

Comparação entre Redes Neurais Artificiais e Regressão Linear Múltipla para estimar radiação solar no município de Seropédica - RJ

Comparison between Artificial Neural Networks and Multiple Linear Regression to estimate solar radiation in the city of Seropédica - RJ

Erylaine Reis Rubim Moreira Araujo¹, Robson Mariano da Silva²

RESUMO

O presente estudo propõe o desenvolvimento e avaliação de uma metodologia utilizando Redes Neurais Artificiais (RNAs) para estimar a incidência de radiação solar em Seropédica - RJ. Primeiramente foi avaliada a necessidade de se utilizar todos os dados disponibilizados pelo Instituto Nacional de Meteorologia. Em seguida foram realizados experimentos variando o número de neurônios na(s) camada(s) escondida(s). Diferentes parâmetros estatísticos foram utilizados para avaliar o desempenho dos modelos (r , MAE, RMSE, D, R^2 , C e skill). Os modelos de RNAs foram comparados com Regressão Linear Múltipla (RLM) para verificar qual método seria satisfatório. Como resultado, foi possível constatar que não há necessidade de utilizar todas as variáveis disponibilizadas pela estação. Analisando a média das 50 simulações realizadas, foi possível verificar também que a RNA cujo desempenho foi superior contava com apenas uma camada escondida, apresentando índice de confiança médio (D) de 88% e coeficiente de determinação médio (R^2) de 85%. Mesmo mostrando-se superiores, os modelos de RNAs não apresentaram ganhos significativos frente aos modelos de RLM. Assim, foi possível concluir que RNAs é uma ferramenta adequada para estimar a incidência de radiação solar.

Palavras-chave: Radiação solar. Redes Neurais Artificiais. Regressão Linear Múltipla.

ABSTRACT

The present study aims the development and evaluation of a methodology using Artificial Neural Networks (ANNs) to estimate the incidence of solar radiation in Seropédica - RJ. Firstly, the need of using all the data provided by the National Institute of Meteorology was assessed. Then experiments were performed by varying the number of neurons in the hidden layer. Different statistical parameters were used to evaluate the performance of the models (r , MAE, RMSE, D, R^2 , C and skill). The RNA models were compared with Multiple Linear Regression (MLR) to verify which method would be satisfactory. As a result, it was possible to find that there is no need to use all the variables available. Analyzing the average of the 50 simulations performed, it was also possible to verify that the RNA whose performance was superior had only one hidden layer, with an average confidence index (D) of 88% and an average determination coefficient (R^2) of 85%. Even though they were superior, the RNA models did not show significant gains compared to MLR models. Thus, it was possible to conclude that ANNs are a suitable tool for solar radiation estimation.

Keywords: Solar Radiation. Artificial Neural Networks. Multiple Linear Regression.

¹ Mestranda em Modelagem Matemática e Computacional. Universidade Federal Rural do Rio de Janeiro

E-mail: erylaine@hotmail.com

² Doutor em Engenharia. Biomédica. Universidade Federal do Rio de Janeiro

1. INTRODUÇÃO

Nos últimos anos a demanda por energia elétrica aumentou drasticamente, pois se tornou fundamental para a vida cotidiana e para a economia. A Energia elétrica proveniente de combustíveis fósseis ou de hidrelétricas gera um grande impacto ambiental, seja pelo agravamento do efeito estufa, a poluição do ar, ou por causar alagamento de grandes áreas, interferindo na vegetação e fauna local, entre outros (COSTA, 2016). Na busca por formas de geração de energia que menos agridam ao meio ambiente, que reduzam as emissões de gases de efeito de estufa, além da perspectiva de esgotamento do petróleo nas próximas décadas (Bentley, 2002) pode-se citar a energia solar como uma alternativa de energia renovável. Existe também a necessidade de expandir a credibilidade e a segurança do setor energético de uma nação e para isso, diversificar a matriz energética é uma boa forma de precaver possíveis imprevistos, como por exemplo, a geração de energia hidroelétrica possui inconstância em decorrência dos longos períodos de estiagem comuns no Brasil (FIORIN ET.AL., 2011).

Metodologias aplicando Redes Neurais Artificiais (RNAs) estão sendo desenvolvidas e utilizadas para avaliação e previsão da disponibilidade de recursos de energia solar, como é o caso de Coutinho (2016) que fez uso de Redes Neurais Artificiais do tipo Perceptrons de Múltiplas Camadas (MLP) e Função de Base Radial (RBF) para realizar a predição de um passo à frente em séries de dados meteorológicos. Além disso, realizou a comparação de tais métodos com modelos de Regressão Linear Múltipla (RLM) onde comprovou que as RNAs MLP e RBF são mais adequadas para tal estudo, pois suas estimativas foram consideravelmente próximas aos dados reais. Mendes et.al. (2017) utilizou as medições de dados meteorológicos provenientes do Sistema de Organização Nacional de Dados Ambientais (SONDA) para predizer a radiação de energia solar global através das Redes Neurais. Como resultado obteve que as Redes Neurais são uma boa ferramenta para previsão de dados de modelos com características não lineares e obteve êxito na previsão de radiação solar global.

Devido o Brasil estar localizado em uma região tropical e subtropical, recebe elevada irradiação solar durante todo o ano, o que propicia a utilização desse tipo de energia. O potencial de geração fotovoltaica no Brasil é tão grande que, no local com menos incidência de radiação solar no país, é possível gerar mais eletricidade solar do que no local com maior incidência de radiação solar da Alemanha, por exemplo. (PEREIRA *et. al.*, 2017).

De acordo com o Instituto de Pesquisa Econômica Aplicada (IPEA, 2019) a capacidade instalada no Brasil aumentou consideravelmente nos últimos dois anos, crescendo de 0,1% para 1,4% de 2016 a 2018. Entre esses dois anos, aproximadamente 41 mil novas usinas de energia solar foram instaladas no Brasil. Houve um aumento significativo no número de estabelecimentos comerciais, industriais e residenciais que passaram a produzir e consumir dessa fonte de energia. Este aumento se dá em virtude de vários fatores, como a diminuição nos custos de fabricação, o aumento na eficiência dos módulos fotovoltaicos e a facilidade na implementação de sistemas de pequeno e grande porte, no entanto o país ainda está em seus passos iniciais na geração de energia solar fotovoltaica. (ABDALA, 2019)

Seropédica é um município do estado do Rio de Janeiro que se localiza aproximadamente 70 km da capital. De acordo com o Instituto Brasileiro de Geografia e Estatística (IBGE, 2017) sua área é de 283,762km². No último censo em 2010 sua população foi de 78.186 pessoas, garantindo a 31^a posição no ranking dos 92 municípios do estado do Rio de Janeiro e a 370^a posição entre os 5.565 municípios do Brasil. A estimativa populacional para 2018 foi de 86.743 pessoas e sua densidade demográfica é 275,53 hab/km².



Figura 1. Mapa do Rio de Janeiro com a região de estudo

As principais atividades econômicas da região são o comércio, a extração de mineral e a indústria (PMS, 2019). Em 2016 o PIB per capita de Seropédica foi de R\$24.602,97. (IBGE, 2017).

De acordo com o Instituto Nacional de Meteorologia (INMET) a temperatura média compensada anual em Seropédica é de 23,80°C. Para a Fundação Instituto de

Desenvolvimento Econômico e Social do Rio de Janeiro (FIDERJ, 1978, apud Castro et.al. 2009) o clima é subúmido e mesotérmico com calor bem distribuído ao longo do ano.

Devido ao fato da população e da economia local estar em crescimento, conseqüentemente haverá o aumento no consumo de energia elétrica. A região é propícia a exploração da energia solar, fato este que justifica a escolha da área de estudo para o presente trabalho.

Os modelos numéricos que são utilizados em previsão meteorológica possuem códigos computacionais para parametrização dos processos radiativos na atmosfera e podem ser utilizados para previsão da irradiação solar na superfície. Porém, as previsões fornecidas por tais modelos, apresentam desvios elevados em relação a valores medidos em superfície para um ou dois dias de antecedência (HEINEMANN, 2004). Estudos indicam como principal motivo desses desvios elevados a dependência da irradiação solar em relação às condições meteorológicas que envolvem intrinsecamente processos físicos não-lineares de difícil parametrização (MESINGER, 2005).

Assim, o presente trabalho é um estudo sobre previsão de radiação solar incidente diária no Município de Seropédica, objetivando principalmente fornecer informações que atendam as atividades ligadas à geração de energia. Para isso utilizou-se o método computacional conhecido como Redes Neurais Artificiais (RNAs) pela sua notória capacidade de reconhecimento de padrões, além de gerar bons resultados para problemas que apresentam alto grau de não-linearidade. Foi feito também uma comparação com o método de Regressão Linear Múltipla (RLM) a fim de analisar qual método apresentaria melhores resultados.

Após a introdução, o artigo está organizado nas seguintes seções: Seção 2 descreve os materiais e métodos utilizados na presente pesquisa, a Seção 3 apresenta os resultados e discussões obtidos e a Seção 4 mostra as conclusões do estudo.

2. MATERIAIS E MÉTODOS

O propósito da presente seção é descrever as etapas que foram realizadas no trabalho proposto, desde a obtenção e preparação dos dados aos parâmetros de avaliação do desempenho dos modelos.

2.1 Conjunto de Dados

A área de estudo e investigação da pesquisa compreende os dados obtidos do Instituto Nacional de Meteorologia (INMET). A estação selecionada para o levantamento de dados do presente estudo faz parte da rede de Estações Meteorológicas Automáticas (EMA), a estação Ecologia Agrícola (A601), situada no município de Seropédica (Latitude: -22.75°, Longitude: -43.68°, Altitude: 35 metros) localizada no estado do Rio de Janeiro (RJ).

Os dados foram selecionados no período de 01 de Maio de 2017 a 31 de Janeiro de 2019, no horário compreendido entre 10h e 20h (hora UTC), devido a corresponder a um período significativo de incidência de radiação solar.

As informações meteorológicas da EMA são coletadas por minuto e estes dados são integralizados a cada hora. A partir destes dados, depois da seleção dos horários de interesse para o presente estudo, foram obtidas através da formação de rotinas computacionais as médias diárias de cada uma das variáveis. Optou-se por fazer as médias diárias para evitar a ocorrência de grandes erros, visto que isso ocorreria caso fossem utilizadas todas as medições horárias coletadas.

Após a obtenção das médias diárias os dados foram normalizados a fim de melhorar a eficiência do treinamento da rede. Dessa forma, tanto as variáveis de entrada como as de saída assumem valores entre 0 e 1. É calculado através da equação (Perboni *et.al.*, 2014):

$$x_{norm} = \frac{x_0 - x_{min}}{x_{max} - x_{min}}$$

Onde: x_{norm} é a variável normalizada, x_0 a variável original, x_{max} o valor máximo observado entre as variáveis e x_{min} o valor mínimo observado entre as variáveis.

Na tabela 1 é possível visualizar o resumo dos dados que foram utilizados pelo modelo, e também um exemplo de um dos conjuntos de treinamento e teste utilizados. Realizadas essas etapas, os dados estavam aptos para serem utilizados pelas redes neurais.

Os dados foram divididos randomicamente em 80% para treinamento e 20% para teste. É importante ressaltar que mesmo sendo selecionados randomicamente, os dados dos conjuntos de treinamento e teste possuem amostras igualmente distribuídas, consistindo assim numa divisão homogênea de todo o conjunto de dados.

Tabela 1. Resumo do conjunto de dados e de um dos subconjuntos de treinamento e teste

		Radiação	Temp_Inst.	Temp_Max.	Temp_Min.	Umi_Inst.	Umi_Max.	Umi_Min.	Pto. Orv_Max.	Pto. Orv_Min.	Vel_vento
Dados	Mínimo	146.9	17.36	17.55	16.94	26.27	29.45	24.55	9.491	7.745	0.6364
	1º Quartil	1049.8	23.42	24.02	22.36	52.18	57.55	49.09	16.309	14.691	2.0000
	Mediana	1440.2	26.18	26.87	25.23	61.91	66.73	58.45	18.709	17.218	2.6727
	Média	1480.0	26.26	26.89	25.27	62.00	66.52	58.97	18.438	16.909	2.7022
	3º Quartil	1976.8	28.78	29.59	27.75	71.09	75.27	67.91	20.700	19.282	3.3182
	Máximo	2686.1	34.85	35.77	33.69	93.18	93.73	92.45	24.245	22.800	5.8273
Treinamento	Mínimo	146.9	17.36	17.55	16.94	26.27	29.45	24.55	9.491	7.745	0.6364
	1º Quartil	1049.8	23.40	24.00	22.34	52.18	57.55	49.09	16.309	14.745	2.0000
	Mediana	1438.2	26.13	26.85	25.15	62.00	67.00	58.36	18.682	17.164	2.6909
	Média	1473.8	26.20	26.83	25.20	62.06	66.00	59.03	18.396	16.865	2.7077
	3º Quartil	1976.8	28.77	29.60	27.67	71.18	75.55	68.09	20.700	19.227	3.3182
	Máximo	2686.1	34.85	35.77	33.69	93.18	93.73	92.45	24.245	22.800	5.8273
Teste	Mínimo	328.9	18.37	18.89	18.02	29.00	32.73	27.36	10.71	8.855	0.6545
	1º Quartil	1058.0	23.50	24.26	22.47	52.48	57.91	49.45	16.51	14.618	2.0068
	Mediana	1440.3	26.47	27.20	25.81	61.50	66.73	58.50	19.05	17.550	2.5864
	Média	1505.2	26.48	27.13	25.51	61.77	66.21	58.76	18.61	17.088	2.6806
	3º Quartil	1955.3	28.93	29.49	28.03	70.52	74.43	67.34	20.91	19.579	3.2559
	Máximo	2639.6	34.50	35.08	33.35	91.00	92.00	89.91	23.68	22.391	5.1000

2.2 Seleção dos Grupos de Preditores

A estação Ecologia Agrícola oferece dados como temperatura, umidade, ponto de orvalho, pressão atmosférica, direção e velocidade dos ventos, precipitação e radiação solar. Para as variáveis temperatura, umidade, ponto de orvalho e pressão atmosférica, são disponibilizadas medições instantâneas, máximas e mínimas, ao todo são 16 variáveis que podem ser utilizadas como preditores de radiação solar.

A fim de otimizar o grupo de preditores objetivando selecionar as variáveis que mais contribuíssem para a predição de radiação solar, foram realizados experimentos utilizando Regressão Linear através do software R, com a função glm. Por meio dessa função foram obtidos as estimativas pontuais, os erros padrão das estimativas pontuais e o nível de significância (p). Foi adotado um nível de significância de 5% ($p < 0,05$), ou seja, um intervalo de confiança de 95%. Nessa etapa foram evidenciadas as variáveis que apresentaram um nível de significância superior a 5%.

Logo em seguida iniciaram-se os experimentos com RNA. O primeiro experimento contou com as 16 variáveis disponibilizadas, e a partir daí as variáveis com $p > 0,05$ foram excluídas do modelo uma a uma.

De acordo com Guarneri (2006), para redes que possuem um grande número de entradas, determinadas topologias levam a resultados incoerentes ou até resultados negativos de radiação. Assim para um grupo de preditores com muitas entradas, o autor

sugere a topologia para o treinamento das RNAs a estrutura RNA-(n)-(n)-(0,5n)-1, onde n é o número de neurônios, portanto nesta etapa do trabalho a recomendação do autor foi seguida.

Em todos os experimentos para cada topologia da RNA foram repetidas 50 simulações, concomitantemente para cada conjunto de treino e teste/validação, usou-se RLM para fins de comparação. Nesta etapa a RLM foi utilizada como uma ferramenta norteadora dos ajustes das variáveis preditoras, ficando como critério de seleção das mesmas a análise dos Índices: Índice de Confiança (C), Erro Médio Absoluto (MAE – Mean Absolute Error) e Raiz do Erro Médio Quadrático (RMSE – Root Mean Squared Error).

2.3 Treinamento da RNA

A metodologia adotada foi a de RNA do tipo Perceptron de múltiplas camadas (MLP). Com o intuito de identificar qual topologia seria mais adequada, a partir dessa etapa o presente estudo foi dividido em dois subgrupos de teste. Ambos consistiram em realizar experimentos variando o número de neurônios na camada escondida, assim, no primeiro foram treinadas redes compostas apenas por uma camada interna e no segundo redes compostas por duas camadas internas.

O próximo passo foi definir quais variáveis seriam utilizadas na entrada da rede. Como citado no item anterior as variáveis: temperatura, umidade, ponto de orvalho e pressão atmosférica, possuíam medições instantâneas, máximas e mínimas. Assim foram feitos testes utilizando cada uma dessas medições individualmente e também com todas de uma vez a fim de verificar qual geraria melhor resultado. Com o intuito de facilitar a referência aos experimentos realizados, uma nomenclatura personalizada foi adotada:

- **Modelo completo:** experimento que utiliza todas as variáveis preditoras indicadas como relevantes para predição de radiação solar.
- **Modelo instantâneo:** experimento que utiliza as medições instantâneas das variáveis preditoras indicadas como relevantes para predição de radiação solar.
- **Modelo máximo:** experimento que utiliza as medições máximas das variáveis preditoras indicadas como relevantes para predição de radiação solar.
- **Modelo mínimo:** experimento que utiliza as medições mínimas das variáveis preditoras indicadas como relevantes para predição de radiação solar.

Logo em seguida os dados foram divididos em conjuntos de treinamento e teste. O software R foi programado para dividir randomicamente o conjunto de dados em 80%

para treinamento da rede e os 20% restante foram reservados para a teste/validação do modelo.

Após as simulações de diversas topologias para investigação de qual seria a mais adequada, foi feito um estudo sobre o comportamento do erro MAE a fim de verificar a necessidade de se realizar simulações de outras topologias além das inicialmente propostas. A tabela 2 apresenta um resumo dos parâmetros utilizados pelo modelo proposto.

Tabela 2. Resumo dos parâmetros do modelo proposto

Parâmetros	Valor
Nº de simulações por arquitetura	50
Nº máximo de épocas	100.000 iterações
Função de ativação	Logística
Likelihood	Falsa
Nº de repetições da rede no treinamento	1
Algoritmo de treinamento	<i>Resilient backpropagation</i>
Inicialização dos pesos	Randomico
Threshold (critério de parada)	0,01
Métrica usada erro de treinamento	SSE (Sum of Squared Errors)

Fazendo uso do mesmo conjunto de dados de entrada, foram calculadas regressões lineares múltiplas (RLMs) para fins de comparação com os resultados da RNA.

Por causa da normalização dos dados, os resultados das RNAs do presente estudo pertencem ao intervalo de 0 a 1. Logo, para poder analisar tais resultados foi indispensável realizar o procedimento de desnormalização dos dados, ou seja, tornar os valores para a escala real, conforme a seguinte equação (Coutinho, 2014):

$$V_j = (x_j * x_{max}) - (x_j * x_{min}) + x_{min}$$

Onde V_j é a variável desnormalizada, x_j variável predita pela rede, x_{max} o valor máximo observado entre as variáveis e x_{min} o valor mínimo observado entre as variáveis.

Após as simulações das redes e avaliar as estruturas com diferentes quantidades de neurônios nas camadas ocultas, a arquitetura mais adequada foi determinada. A primeira análise foi avaliar qual dentre os modelos (completo, instantâneo, máximo ou mínimo) geraria melhores resultados. Posteriormente foi feita uma análise entre as redes que possuíam o mesmo número de camadas, selecionando também as que apresentariam

melhores resultados e por fim foram comparadas as melhores redes com uma e duas camadas escondidas.

2.4 Avaliações das Previsões

Para comparar o desempenho dos modelos foram utilizadas diferentes medidas estatísticas, que geralmente são utilizadas na literatura, além de também serem descritas por vários autores (Pezzopane *et al.*, 2012; Guarnieri, 2006; Deshmukh e Ghatol, 2010; Fonseca *et al.*, 2012), são elas:

• Coeficiente de Correlação de Pearson (r):

Por definição o coeficiente de correlação representa uma medida do grau de dependência linear entre duas variáveis, que no presente estudo são os valores estimados pela rede (x_j) e os valores observados pela estação (O_j) é calculado por:

$$r = \frac{\sum_{j=1}^N (x_j - \bar{x}) (O_j - \bar{O})}{\sqrt{\frac{\sum_{j=1}^N (x_j - \bar{x})^2}{n}} \cdot \sqrt{\frac{\sum_{j=1}^N (O_j - \bar{O})^2}{n}}}$$

Onde \bar{x} é a média dos valores estimados e \bar{O} é a média dos valores observados e n representa o número de dados utilizados. O resultado do coeficiente de correlação varia no intervalo de $[-1,1]$, onde 1 representa que existe uma correlação linear positiva perfeita entre x_j e O_j .

• Erro Médio Absoluto (MAE – Mean Absolute Error):

Através do MAE é possível obter a magnitude média dos erros de previsão. O resultado varia no intervalo de $[0, \infty)$ onde 0 representa o valor para a previsão perfeita. É calculado através da seguinte equação:

$$MAE = \frac{\sum_{j=1}^n |O_j - x_j|}{n}$$

• Raiz do Erro Médio Quadrático (RMSE – Root Mean Squared Error):

É o valor que representa a magnitude média dos erros de previsão. O resultado varia no intervalo de $[0, \infty)$ onde 0 representa o valor para a previsão perfeita. É obtido por:

$$RMSE = \sqrt{\frac{\sum_{j=1}^n (O_j - x_j)^2}{n}}$$

• Índice de Concordância (D):

Fornece o grau de exatidão entre as variáveis envolvidas, pois relaciona a diferença entre os valores estimados em relação aos valores observados. O resultado varia no intervalo de [0,1] onde 0 representa nenhuma concordância e 1 significa que houve uma concordância perfeita.

$$D = 1 - \frac{\sum_{j=1}^n (O_j - x_j)^2}{\sum_{j=1}^n \left(\left| x_j - \bar{O} \right| + \left| O_j - \bar{O} \right| \right)^2}$$

• **Coefficiente de Determinação (R²):**

O R² é uma medida que descreve a qualidade do ajuste do modelo. Seu resultado varia de 0 a 1, representado de 0 a 100% da variância total das observações explicadas pelos valores estimados. Sua equação é dada por:

$$R^2 = \frac{\sum_{j=1}^N (x_j - \bar{O})^2}{\sum_{j=1}^N (O_j - \bar{O})^2}$$

• **Índice de Confiança (C):**

Através do índice de confiança, é possível analisar simultaneamente a precisão e a exatidão dos resultados obtidos. Ele é calculado através do produto do coeficiente de correlação (r) pelo índice de concordância:

$$C = (r * D)$$

Seus valores variam no intervalo de [0,1] onde zero representa que não há nenhuma concordância e 1 representa a concordância perfeita (Pezzopane *et. al.*, 2012). A Tabela 3 demonstra os critérios para avaliação de desempenho.

Tabela 3. Critérios para avaliação e análise do desempenho de modelos com base no índice de confiança

Valor de C	Desempenho
> 0,85	Ótimo
0,76 a 0,85	Muito bom
0,66 a 0,75	Bom
0,61 a 0,65	Mediano
0,51 a 0,60	Sofrível
0,41 a 0,50	Mau
≤ 0,40	Péssimo

• **Skill:**

Este índice é utilizado para comparar dois modelos em termos de ganho (ou aperfeiçoamento), é obtido relacionando um modelo a outro adotado como referência. Sua fórmula é dada por:

$$Skill(\acute{I}NDICE, ref) = \frac{\acute{I}NDICE - \acute{I}NDICE_{ref}}{\acute{I}NDICE_{perf} - \acute{I}NDICE_{ref}}$$

onde $\acute{I}NDICE$ pode assumir qualquer um dos índices anteriormente citados, $\acute{I}NDICE_{perf}$ é o valor onde este índice assume a previsão perfeita, ou seja, assume o valor 0 para erros como MAE ou RMSE e assume o valor 1 para r, D ou R^2 , $\acute{I}NDICE_{ref}$ é o valor deste índice calculado para uma outra previsão adotada como referência.

Quando o Skill assume valor igual a 1 representa que a previsão atingida foi perfeita, quando se iguala a 0 significa que não houve ganho se comparado a previsão de referência e por fim, quando se obtém valores negativos quer dizer que a previsão avaliada é inferior a previsão de referência.

3. RESULTADOS E DISCUSSÃO

Nesta seção são exibidos e discutidos os resultados alcançados pelos modelos aplicados no presente estudo.

3.1 Seleção dos Grupos de Preditores

Como mencionado anteriormente, a estação Ecologia Agrícola fornece dados que podem ser utilizados como conjunto de preditores de radiação solar como temperatura (instantânea, máxima, mínima), umidade (instantânea, máxima, mínima), ponto de orvalho (instantâneo, máximo, mínimo), pressão atmosférica (instantânea, máxima, mínima), vento (velocidade, rajada, direção) e chuva.

Inicialmente todas as 16 variáveis preditores foram utilizadas, a partir daí, foram realizados experimentos excluindo os grupos: chuva, pressão atmosférica instantânea, pressão atmosférica máxima, pressão atmosférica mínima, direção dos ventos, rajada dos ventos, ponto de orvalho instantâneo, temperatura máxima e umidade instantânea respectivamente.

Os índices utilizados como critério de seleção nesta etapa, foram utilizados por serem índices que evidenciam a precisão e a exatidão dos resultados além de representam a magnitude média dos erros de previsão. Os erros MAE e RMSE mensuram

o tamanho médio dos módulos dos desvios entre previsões e observações, a diferença entre ambos é que enquanto o MAE é a média aritmética dos módulos dos desvios, o RMSE representa uma média quadrática dos desvios, pode-se pensar nele como uma medida análoga ao desvio padrão.

Dos dados obtidos nas 50 simulações realizadas, foram calculadas as médias tanto para a RNA como para RLM. Foram selecionados também os casos que apresentaram índice de Confiança (C) máximo e mínimo, conforme mostra a tabela 4. Ao analisar a tabela observou-se que em alguns casos, os experimentos mostraram-se bastante similares entre si, houve casos onde a RLM mostrou desempenho superior a RNA.

É importante notar que, embora para alguns dos experimentos a média seja aparentemente boa, analisando os valores máximos e mínimos atingidos pela RNA nas 50 simulações realizadas, observa-se que em todos os casos os valores máximos de C apresentam desempenho ótimo ($C > 0,85$), entretanto ao analisar os valores mínimos obtidos, em alguns dos casos apresentaram desempenho sofrível (0,51 a 0,60), mau (0,41 a 0,50) ou até mesmo péssimo ($C \leq 0,40$).

Após uma análise detalhada das médias obtidas nos experimentos, foi possível verificar e definir o caso onde os melhores resultados foram encontrados, (maior Índice de Confiança e menores MAE e RMSE). Assim o experimento que contava com 9 preditores (temperatura (instantânea, máxima, mínima), umidade (instantânea, máxima, mínima), ponto de orvalho (máximo, mínimo) e velocidade do vento) foi o mais indicado para ser usado no treinamento das RNAs. Foi observado também que em termos do Índice de Confiança, dentre as 50 simulações realizadas, o valor mínimo obtido apresentou desempenho muito bom (0,76 a 0,85), o valor máximo e a média apresentaram ótimo desempenho ($C > 0,85$).

Nesta etapa foi possível notar que excetuando os casos onde a RLM mostrou-se superior a RNA, nos demais experimentos a RNA superou a RLM, houve também uma melhora significativa no modelo com a redução das variáveis de 16 para 9 preditores, constatando que de fato existem variáveis que podem ser excluídas do mesmo. Além disso, verificou-se que as 9 variáveis preditoras selecionadas possuem informação suficiente para gerar uma boa estimativa de radiação solar incidente em Seropédica.

Tabela 3. Índices de avaliação calculados na seleção das variáveis preditora

Experimento		Índices	MAE	RMSE	C
16 preditores (16-16-8-1)	RNA	Máx	0.0712	0.0965	0.8854
		Mín	0.1146	0.2039	0.5775
		Média	0.0928	0.1344	0.7917
	RLM	Máx	0.0639	0.0787	0.9265
		Mín	0.2375	1.7792	0.0084
		Média	0.0956	0.2891	0.7729
15 preditores (15-16-8-1)	RNA	Máx	0.0678	0.0889	0.9053
		Mín	0.1581	0.7681	0.1071
		Média	0.0896	0.1465	0.7850
	RLM	Máx	0.0723	0.0922	0.9086
		Mín	0.2743	2.2341	0.0044
		Média	0.1098	0.4431	0.6878
14 preditores (14-14-7-1)	RNA	Máx	0.0674	0.0848	0.9054
		Mín	0.0966	0.3030	0.4272
		Média	0.0813	0.1141	0.8451
	RLM	Máx	0.0728	0.0906	0.9117
		Mín	0.0833	0.1091	0.8361
		Média	0.0784	0.1000	0.8757
13 preditores (13-14-7-1)	RNA	Máx	0.0671	0.0849	0.9155
		Mín	0.1264	0.5638	0.1613
		Média	0.0828	0.1219	0.8296
	RLM	Máx	0.0709	0.0881	0.9044
		Mín	0.0866	0.1135	0.8158
		Média	0.0787	0.0998	0.8730
12 preditores (12-12-6-1)	RNA	Máx	0.0656	0.0855	0.9161
		Mín	0.1022	0.3046	0.4488
		Média	0.0778	0.1090	0.8565
	RLM	Máx	0.0692	0.0868	0.9082
		Mín	0.0882	0.1142	0.8242
		Média	0.0779	0.0996	0.8752
11 preditores (11-12-6-1)	RNA	Máx	0.0656	0.0844	0.9158
		Mín	0.0817	0.1454	0.7763
		Média	0.0752	0.1012	0.8734
	RLM	Máx	0.0666	0.0849	0.9080
		Mín	0.0883	0.1182	0.8172
		Média	0.0796	0.1018	0.8687
10 preditores (10-10-5-1)	RNA	Máx	0.0646	0.0823	0.9142
		Mín	0.0816	0.1118	0.8329
		Média	0.0751	0.1011	0.8729
	RLM	Máx	0.0699	0.0894	0.8972
		Mín	0.0983	0.1264	0.8119
		Média	0.0828	0.1055	0.8589
9 preditores (9-10-5-1)	RNA	Máx	0.0591	0.0747	0.9317
		Mín	0.0901	0.1297	0.8145
		Média	0.0748	0.1004	0.8741
	RLM	Máx	0.0737	0.0901	0.8996
		Mín	0.0955	0.1189	0.8223
		Média	0.0820	0.1050	0.8599
8 preditores (8-8-4-1)	RNA	Máx	0.0681	0.0867	0.9155
		Mín	0.0861	0.1185	0.8180
		Média	0.0751	0.1018	0.8707
	RLM	Máx	0.0730	0.0917	0.8956
		Mín	0.0891	0.1210	0.8092
		Média	0.0820	0.1047	0.8618
7 preditores (7-8-4-1)	RNA	Máx	0.0646	0.0907	0.9014
		Mín	0.0850	0.1195	0.8139
		Média	0.0760	0.1047	0.8625
	RLM	Máx	0.0782	0.0997	0.8894
		Mín	0.0929	0.1227	0.8043
		Média	0.0844	0.1086	0.8490

3.2 Seleção Entre os Modelos: Completo, Instantâneo, Máximo e Mínimo

Como especificado na seção anterior, foi adotada uma nomenclatura personalizada para os experimentos nesta fase do trabalho, que consistiram em modelo completo (utiliza todas as variáveis preditoras selecionadas na etapa anterior de uma única vez), modelo instantâneo, modelo máximo e modelo mínimo que utiliza as variáveis preditoras individualmente.

Observou-se assim que para cada topologia entre os modelos instantâneo, máximo e mínimo não havia grande variação entre as médias dos índices, apresentando valores aproximados entre si. No que se refere a comparação entre os modelos instantâneo, máximo e mínimo com o modelo completo, notou-se que este supera os demais substancialmente, pois os maiores valores médios do Índice de Confiança atingidos pelos modelos individuais não superaram os valores de C atingidos pelo modelo completo, além disso o modelo completo apresenta menores valores de MAE e RMSE.

Ainda analisando o Índice de Confiança foi possível observar que o desempenho dos modelos, tanto para uma quanto para duas camadas escondidas, variou bastante. A tabela 4 apresenta um resumo dos índices observados dos experimentos. A fim de abreviar a tabela, foi adotado o uso de siglas para representar o desempenho do modelo:

- OT: Ótimo ($> 0,85$)
- MB: Muito bom (0,76 a 0,85)
- BO: Bom (0,66 a 0,75)
- ME: Mediano (0,61 a 0,65)
- SO: Sofrível (0,51 a 0,60)
- MA: Mau (0,41 a 0,50)
- PE: Péssimo ($C \leq 0,40$)

Tabela 4. Resumo do Índice de Confiança para topologias com uma e duas camadas escondidas

Modelo \ nº de camadas	Valores Médios		Valores Máximos		Valores Mínimos	
	Uma	Duas	Uma	Duas	Uma	Duas
Instantâneo	MB	BO e MB	MB e OT	MB e OT	BO, ME e MA	BO, ME, SO e PE
Máximo	MB	BO e MB	MB e OT	MB e OT	BO, MB e PE	BO, ME, SO e PE
Mínimo	MB	BO e MB	MB e OT	MB e OT	BO, SO, MA	BO, ME, SO, PE e MA
Completo	OT	OT	OT	OT	MB	MB

Fazendo a análise dos modelos em termos de tempo médio computacional necessário para a conclusão do treinamento da RNA, obteve-se a tabela 5. Antes mesmo das simulações serem realizadas, esperava-se que o modelo completo levasse mais tempo para convergência que os demais modelos, visto que a quantidade de dados na camada de entrada era maior. Essa suspeita foi corroborada após a mensuração do

tempo médio computacional, contudo, o modelo completo mostrou ser o mais indicado, pois leva a menores erros e a maior Índice de Confiança. Assim o modelo completo demonstrou ser o mais adequado para estimar a radiação solar e, portanto foi adotado no presente trabalho.

Tabela 5. Tempo médio computacional entre os modelos

RNA \ Modelo	Completo	Instantâneo	Máximo	Mínimo
1 Camada	99s	14s	21s	10s
2 Camadas	55s	23s	24s	13s

3.3 Comportamento do Erro MAE

Fazendo uso das mesmas topologias utilizadas na subseção anterior e considerando a média do modelo completo, realizou-se então um estudo do comportamento do erro MAE, com o propósito de verificar se haveria a necessidade de acrescentar outras topologias ao estudo.

Assim, foi observado o comportamento do erro MAE para as RNAs de uma e duas camadas escondidas com quantidades diferentes de neurônios em cada uma. Primeiramente foram analisadas as RNAs com uma camada escondida (figura 2).

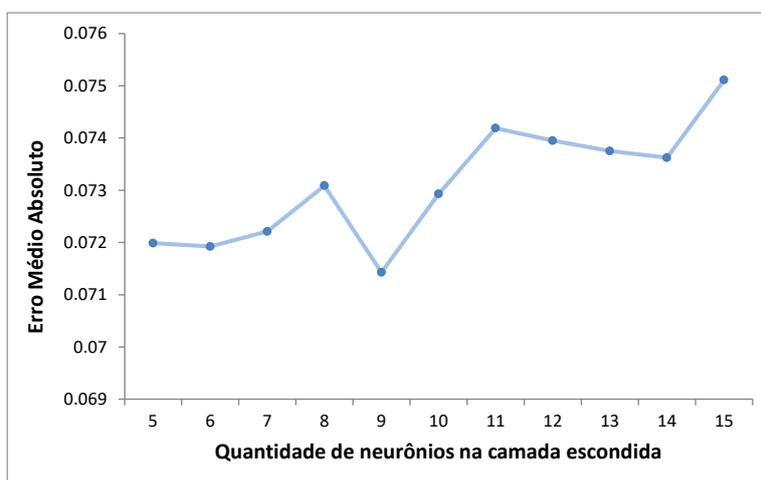


Figura 2. Comportamento do MAE para diferentes quantidades de neurônios na camada escondida

Notou-se que quanto maior a quantidade de neurônios na camada escondida o erro MAE cresceu consideravelmente. E, portanto verificou-se que não haveria a necessidade de treinar outras redes além dessas visto a tendência do erro.

Para redes com duas camadas escondidas, a um primeiro momento, foram treinadas todas as redes apresentadas na figura 3 com exceção da topologia (RNA-9-10-5-1) e até

esta etapa foi possível analisar a um primeiro momento que o erro MAE diminui, logo em seguida torna a crescer, porém a partir da topologia (RNA-9-8-7-1) não foi possível verificar um comportamento expressivo do erro, impossibilitando afirmar se o mesmo iria aumentar ou diminuir. Por esse motivo foi realizado um teste adicional utilizando a topologia (RNA-9-10-5-1) para investigar qual seria a tendência do erro MAE, a partir daí percebeu-se claramente que o erro voltou a crescer.

Portanto, foi verificado que para as RNAs tanto para uma quanto para duas camadas escondidas, na medida em que se aumenta a quantidade de neurônios nas camadas o erro aumenta. Constatando assim que não há a necessidade de se fazer uso de RNAs com um número muito elevado de neurônios nas camadas internas, visto que o desempenho das redes é aproximado entre si e, além disso, erro e o custo computacional no treinamento aumentam.

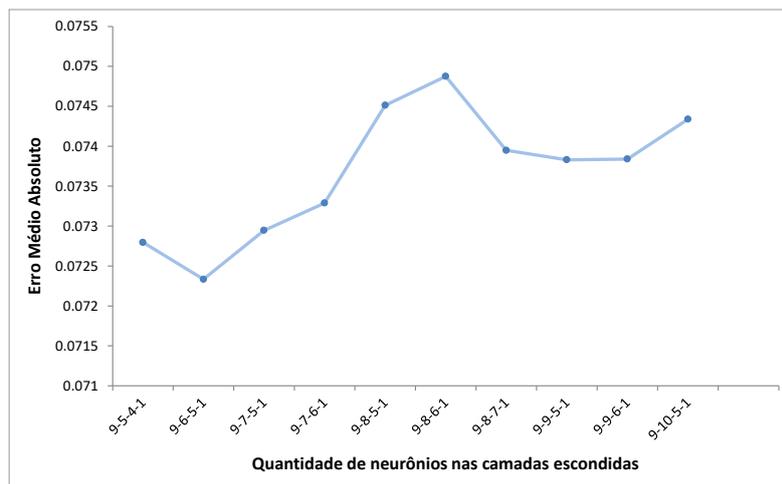


Figura 3. Comportamento do MAE para diferentes quantidades de neurônios nas camadas escondidas

3.4 Resultados de Irradiação Solar

Após a seleção do modelo completo, buscou-se fazer uma análise mais detalhada do mesmo a fim de definir qual seria a topologia mais adequada para estimar a radiação solar, para isso foram examinadas as médias das 50 simulações para cada topologia, tanto para uma quanto para duas camadas escondidas. Observou-se que em nenhum caso a média da RLM superou a RNA, tanto para uma quanto para duas camadas. Notou-se também que os Índices de Confiança não apresentaram valores muito distantes entre si, quando analisado os casos de valor máximo e mínimo para cada topologia.

Dessa forma as redes que demonstraram melhores desempenhos foram RNA-9-9-1 e RNA-9-6-1 (tabela 7) para uma camada escondida e RNA-9-7-5-1 e RNA-9-6-5-1

(tabela 8) para duas camadas escondidas. Dentre essas topologias, a que apresentou maiores Índice de Confiança (C) e Coeficiente de Determinação (R^2), além de menores erros MAE, MSE e RMSE foi a RNA-9-9-1. Quando comparada ao modelo de RLM, a RNA mostrou uma redução em termos de Skill de aproximadamente de 12% do MAE, 14% do MSE e 7% do RMSE, em termos de R^2 e C apresentou um ganho de 14% e 15% respectivamente, conforme a tabela 6.

Tabela 6. Comparação entre RNA e RLM para previsão de radiação solar

	MAE	MSE	RMSE	R^2	C
RNA-9-9-1	0.071	0.009	0.095	0.853	0.888
RLM	0.081	0.011	0.102	0.830	0.868
SKILL	0.12	0.14	0.07	0.14	0.15

Tabela 7. Diferentes topologias para uma camada escondida

Topologia	Índices	épocas	MAE	MSE	RMSE	R^2	C
(9-5-1)	Máx	68254	0.0641	0.0070	0.0834	0.8953	0.9233
	Min	40783	0.0841	0.0137	0.1172	0.7789	0.8352
	Média		0.0720	0.0094	0.0968	0.8456	0.8827
(9-6-1)	Máx	75693	0.0613	0.0068	0.0823	0.8995	0.9257
	Min	45576	0.0819	0.0133	0.1151	0.7774	0.8330
	Média		0.0719	0.0093	0.0961	0.8494	0.8843
(9-7-1)	Máx	53688	0.0628	0.0067	0.0817	0.8968	0.9208
	Min	36257	0.0856	0.0144	0.1201	0.7666	0.8215
	Média		0.0722	0.0094	0.0966	0.8452	0.8818
(9-8-1)	Máx	55727	0.0633	0.0069	0.0833	0.8893	0.9142
	Min	42704	0.0834	0.0137	0.1169	0.7627	0.8320
	Média		0.0731	0.0097	0.0980	0.8409	0.8786
(9-9-1)	Máx	24040	0.0628	0.0065	0.0804	0.8998	0.9229
	Min	49744	0.0811	0.0148	0.1217	0.7533	0.8235
	Média		0.0714	0.0091	0.0947	0.8533	0.8879
(9-10-1)	Máx	96165	0.0615	0.0067	0.0821	0.8932	0.9182
	Min	48777	0.0914	0.0152	0.1233	0.7582	0.8208
	Média		0.0729	0.0095	0.0972	0.8425	0.8800
(9-11-1)	Máx	54232	0.0683	0.0079	0.0890	0.8774	0.9063
	Min	93757	0.0822	0.0136	0.1167	0.7571	0.8171
	Média		0.0742	0.0102	0.1004	0.8325	0.8723
(9-12-1)	Máx	46060	0.0668	0.0073	0.0855	0.8793	0.9098
	Min	70725	0.0860	0.0134	0.1157	0.7588	0.8169
	Média		0.0739	0.0099	0.0994	0.8373	0.8758
(9-13-1)	Máx	77180	0.0639	0.0072	0.0847	0.8910	0.9166
	Min	32861	0.0871	0.0135	0.1162	0.7801	0.8327
	Média		0.0737	0.0099	0.0990	0.8390	0.8774
(9-14-1)	Máx	88951	0.0633	0.0071	0.0842	0.8837	0.9111
	Min	85266	0.0858	0.0140	0.1183	0.7696	0.8230
	Média		0.0736	0.0099	0.0991	0.8376	0.8760
(9-15-1)	Máx	63620	0.0601	0.0067	0.0818	0.8879	0.9165
	Min	87117	0.0891	0.0194	0.1391	0.6811	0.7677
	Média		0.0751	0.0102	0.1007	0.8329	0.8725

Tabela 8. Diferente topologias para duas camadas escondidas

Índices Topologia		n.epochs	MAE	MSE	RMSE	R ²	C
(9-5-4-1)	Máx	53442	0.0591	0.0063	0.0792	0.9014	0.9246
	Mín	38530	0.0821	0.0127	0.1125	0.7905	0.8400
	Média		0.0728	0.0097	0.0979	0.8423	0.8796
(9-6-5-1)	Máx	43931	0.0598	0.0061	0.0782	0.9069	0.9294
	Mín	45830	0.0813	0.0138	0.1175	0.7589	0.8249
	Média		0.0723	0.0097	0.0982	0.8424	0.8800
(9-7-5-1)	Máx	47310	0.0621	0.0063	0.0795	0.9076	0.9290
	Mín	52997	0.0841	0.0132	0.1147	0.7662	0.8290
	Média		0.0729	0.0097	0.0982	0.8426	0.8807
(9-7-6-1)	Máx	73357	0.0633	0.0065	0.0808	0.8832	0.9100
	Mín	36628	0.0899	0.0147	0.1212	0.7746	0.8301
	Média		0.0733	0.0098	0.0987	0.8405	0.8787
(9-8-5-1)	Máx	39366	0.0636	0.0077	0.0877	0.8873	0.9137
	Mín	44741	0.0905	0.0164	0.1279	0.7346	0.8009
	Média		0.0745	0.0102	0.1005	0.8331	0.8736
(9-8-6-1)	Máx	50477	0.0661	0.0070	0.0838	0.8809	0.9085
	Mín	59942	0.0860	0.0156	0.1249	0.7303	0.8061
	Média		0.0749	0.0104	0.1015	0.8292	0.8708
(9-8-7-1)	Máx	97294	0.0673	0.0070	0.0838	0.8940	0.9218
	Mín	36781	0.0810	0.0150	0.1225	0.7333	0.7982
	Média		0.0739	0.0100	0.0998	0.8366	0.8767
(9-9-5-1)	Máx	43900	0.0596	0.0063	0.0792	0.8989	0.9225
	Mín	63025	0.0856	0.0150	0.1226	0.7463	0.8098
	Média		0.0738	0.0100	0.0995	0.8364	0.8765
(9-9-6-1)	Máx	81172	0.0632	0.0070	0.0834	0.8935	0.9187
	Mín	72834	0.0784	0.0280	0.1672	0.5478	0.7060
	Média		0.0738	0.0103	0.1008	0.8314	0.8734
(9-10-5-1)	Máx	30924	0.0601	0.0065	0.0806	0.8986	0.9222
	Mín	25855	0.0843	0.0155	0.1243	0.7516	0.8276
	Média		0.0743	0.0101	0.1000	0.8342	0.8746

A figura 4 apresenta a dispersão dos valores estimados (X) pelo método e os valores observados (O) pela estação Ecologia Agrícola. As linhas cinza possuem coeficientes angulares unitários e passam pelas origens dos gráficos, elas denotam os casos onde as previsões seriam perfeitas, ou seja, os valores estimados se igualam aos valores observados. (X=O). Na figura 4 encontram-se também as retas que representam o comportamento geral das previsões (pretas), os dados estimados pela RNA em vermelho e pela RLM em azul.

Observa-se nos diagramas de dispersão da figura 4 que a linha de tendência do modelo de RNA aproxima-se consideravelmente da linha de previsão perfeita, ao passo que a linha de tendência do modelo de RLM encontra-se abaixo da linha de previsão perfeita, o que indica que este modelo estima intensidades de radiação menores do que as observadas pela estação, ou seja, os dados estimados por RLM subestimam a radiação solar.

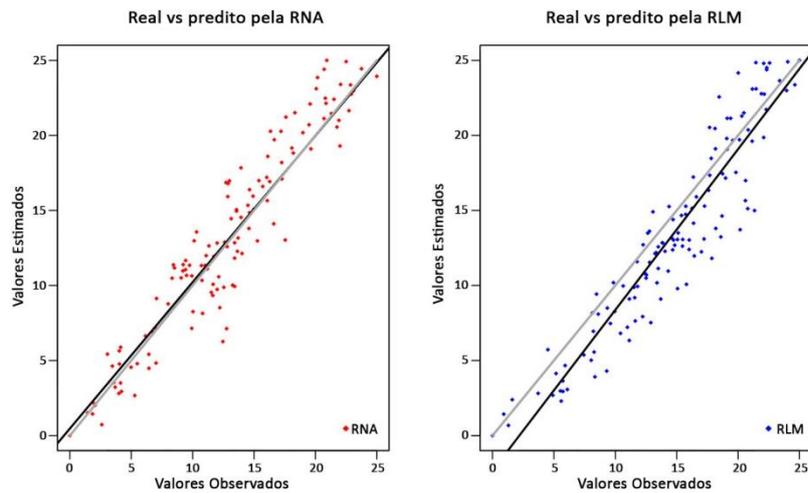


Figura 4. Diagrama de dispersão dos valores estimados de radiação solar do modelo RNA-9-9-1 e de RLM

É apresentada na figura 5 a arquitetura da RNA que demonstrou melhor resultado no modelo proposto. Interessante observar que a rede que demonstrou melhor desempenho consta com apenas uma camada interna de neurônios ocultos. Resultado semelhante foi encontrado por Fiorin *et.al.* (2011), que em seu estudo utilizou o algoritmo de Levenberg-Marquardt, a fim de analisar a incidência de radiação solar em São Martinho da Serra no Rio Grande do Sul. Os dados de irradiação solar foram coletados na estação de superfície pertencentes à rede SONDA. Como resultado obteve que as redes neurais podem fornecer estimativas confiáveis para avaliação da disponibilidade de energia solar, apresentando também melhor desempenho em comparação a outras técnicas estatísticas utilizadas e inclusive concluiu também que uma camada oculta é suficiente para gerar uma boa estimativa.

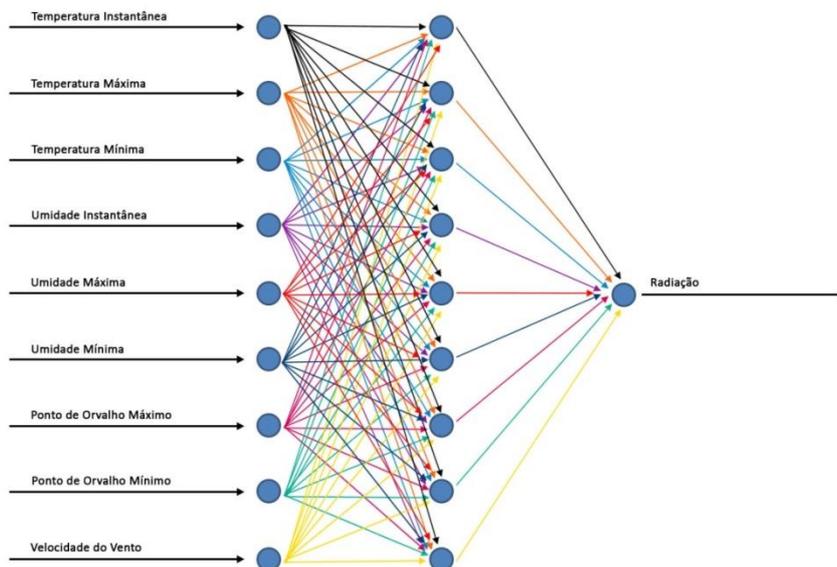


Figura 5. Arquitetura da rede MLP proposta

5. CONSIDERAÇÕES FINAIS

O objetivo do presente estudo foi alcançado com êxito, que foi avaliar a aplicabilidade de Redes Neurais Artificiais (RNAs) do tipo Múltiplas Camadas (MLP) para estimar a incidência de radiação solar em Seropédica, concluindo que este método computacional é apropriado para tal fim, pois é um recurso que gera boas estimativas com pequenos erros. Além disso, foi possível comprovar também que o software R com o uso do algoritmo resilient backpropagation é uma boa ferramenta para a proposta do presente estudo.

No que diz respeito à comparação do desempenho entre as RNAs e os modelos de Regressão Linear Múltipla (RLM), notou-se que o método de Redes Neurais Artificiais superou a RLM, contudo não apresentaram ganhos significativos frente a estas. Acredita-se que a similaridade entre os métodos se dá em virtude do período relativamente curto de coleta de dados.

Foi possível observar também que para estimar a radiação solar não se faz necessário o uso de todas as variáveis disponibilizadas pelo Instituto Nacional de Meteorologia através da estação Ecologia Agrícola. Assim, as variáveis consideradas foram temperatura (instantânea, máxima, mínima), umidade (instantânea, máxima, mínima), ponto de orvalho (máximo, mínimo), velocidade do vento.

Na análise entre os modelos: completo, instantâneo, máximo e mínimo, comprovou-se que, embora o tempo computacional gasto pelo modelo completo seja maior do que os outros modelos, ele leva a melhores resultados.

Para cada topologia de RNA utilizada no trabalho foram repetidas 50 simulações e ao analisar as médias dos experimentos realizados das RNAs com uma e duas camadas escondidas, obteve-se como resultado que uma camada escondida é suficiente para gerar boas estimativas de radiação solar, destacando-se a topologia RNA-9-9-1 que apresentou um ótimo desempenho de Índice de Confiança (C) médio de 88% e coeficiente de determinação médio (R²) de 85%. Dentre as 50 simulações da topologia RNA-9-9-1 o maior Índice de Confiança (C) alcançado foi de 92% apresentando também um ótimo desempenho. Concluindo que não é viável investir em redes muito grandes tendo em vista que os resultados são muito próximos à redes com menos camadas e neurônios ocultos, além disso, notou-se que em redes com muitas camadas e neurônios ocultos o erro e o custo computacional no treinamento aumentam.

Finalmente, tendo em vista a tendência atual por formas de geração de energia que menos agridam ao meio ambiente, e que nos últimos anos a energia solar fotovoltaica tem se mostrado uma opção promissora, o Brasil vive um momento de oportunidades, pois possui um território propício para o aproveitamento desse tipo de energia. O município de Seropédica localizado no estado do Rio de Janeiro possui uma vasta área onde poderia ser aproveitada para instalação de painéis solares, a própria Universidade Federal Rural do Rio de Janeiro, situada no município, dispõe de uma ampla área que também poderia ser utilizada. Dessa forma, caso o Brasil continue desenvolvendo sua capacidade de geração fotovoltaica fica em evidência, neste trabalho, o elevado potencial da região de Seropédica para investimentos desta natureza.

AGRADECIMENTO

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, e assim, deixo o meu agradecimento pelo auxílio financeiro da bolsa de mestrado (Processo 23038.013648/2018-51).

REFERÊNCIAS

- ABDALA, P.J.P. **Energia Solar e Eólica**. Ponta Grossa (PR): Atena Editora; v. 1. 2019.
- BENTLEY, R.W., **Energy Policy** 30, 189, 2002.
- CASTRO, J.O; TEIXEIRA, R.O; OLIVEIRA, A.D E CARVALHO, A. G. **Comportamento de Coleópteros Degradadores de Madeira em Plantio de *Mimosa caesalpiniaefolia* (Sabiá) Seropédica, RJ**. XIII Encontro Latino Americano de Iniciação Científica e IX Encontro Latino Americano de Pós-Graduação – Universidade do Vale do Paraíba, 2009.
- COSTA, R. C. **Modelos Preditivos de Velocidade de Vento para Sistemas Eólicos Baseados em Redes Neurais Artificiais**. Monografia, UFPB, João Pessoa, 2016.
- COUTINHO, E. R. **Aplicação de Técnicas de Inteligência Computacional no preenchimento de Falhas de Séries Temporais Climatológicas**. Dissertação. UFRRJ, Seropédica, 2014.
- _____, E. R. SILVA, R. M.; DELGADO, A. R. S. **Utilização de Técnicas de Inteligência Computacional na Predição de Dados Meteorológicos**. **Revista Brasileira de Meteorologia**, v. 31, n. 1, p. 24-36, 2016.
- DESHMUKH, R. P.; GHATOL, A. A. **Comparative Study of Temporal Neural Networks for Short Term Flood Forecasting**. **International Journal of Computer Applications**, v. 5, n. 12, p. 24-28, 2010.

FIORIN, D.V., MARTINS, F.R, SCHUCH, N.J., PEREIRA, E.B. **Aplicações de Redes Neurais e Previsões de disponibilidade de Recursos Energéticos Solares.** Revista Brasileira de Ensino de Física, v. 33, n. 1, 1309, 2011.

FONSECA, J. S.; MARTINS, G. A.; TOLEDO, G. L. **Estatística Aplicada.** Ed. 2, p. 267. 2012.

GUARNIERI, R.A., **Emprego de Redes Neurais Artificiais e Regressão Linear Múltipla no Refinamento das Previsões de Radiação Solar do Modelo Eta.** Dissertação de Mestrado, INPE, São José dos Campos, 2006.

HEINEMANN, D., LORENZ, E. and GIRODO, M., in: Proc. **International Workshop on Solar Resource from the Local Level to Global Scale in Support to the Resource Management of Renewable Electricity Generation.** Institute for Environment and Sustainability, Joint Research Center, Ispra. 2004.

IBGE – **Instituto Brasileiro de Geografia e Estatística.** v4.3.16.1 Brasília: IBGE. 2017. Disponível em <<https://www.ibge.gov.br/>> acesso em 31/01/2019.

IPEA – Instituto de Pesquisa Econômica Aplicada. **Caderno ODS: Assegurar o Acesso Confiável, Sustentável, Moderno e a Preço Acessível à Energia para Todos.** 2019. Disponível em <http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=34715&Itemid=444> Acesso em 10/05/2019.

INMET, Instituto Nacional de Meteorologia. **BDMEP - Banco de Dados Meteorológicos para Ensino e Pesquisa.** Disponível em: <<http://www.inmet.gov.br/portal/>> Acesso em 01/02/2019

MENDES, I.A., REZENDE, R.A.D., NASCIMENTO, T.H.F., SILVA, J.S.F. e SILVA, O. F. A **Utilização das Redes Neurais Artificiais na Previsão de Radiação Solar Global.** Congresso Técnico Científico da Engenharia e da Agronomia. Belém – PA 2017.

MESINGER F., in: **Proceedings of ICTP Regional Weather Predictability and Modeling Workshop,** CPTEC/INPE, São José dos Campos 2005. Disponível em <ftp://ftp1.cptec.inpe.br/etamd/Download/model/EtaGuide.2008.8.pdf>. Acesso em 21/02/2018

PERBONI, A; FRIZZONE, J.A.; CAMARGO, A. P. **Artificial Neural Network Based Equation to Estimate Head Loss Along Dripp Irrigation Laterals.** Revista Brasileira de Agricultura Irrigada v.8, nº. 2, p. 77 - 85, 2014

PEREIRA, E.B., MARTINS, F.R., GONÇALVES, A.R; COSTA, R.S.; LIMA, F.J.L.; RUTHER, R.; ABREU, S.L.; TIEPOLO, G.M.; PEREIRA, S.V.; SOUZA, J.G. **Atlas Brasileiro de Energia Solar:** Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP. 2ªed. 88p. 2017.

PEZZOPANE, J.E. M.; CASTRO, F.S.; PEZZOPANE, J.R.M.; CECÍLIO, R.A. **AGROMETEOROLOGIA: Aplicações para o Espírito Santo.** Alegre, ES: CAUFES. 2012.

PINHEIRO, E.; LOVATO, A.; RÜTHER, R. Aplicabilidade de Redes Neurais Artificiais para Análise de Geração de Energia de um Sistema Fotovoltaico Conectado a Rede Elétrica. Revista Brasileira de Energias Renováveis. v.6, n.5, p. 875-892, 2017.

PMS - Prefeitura Municipal de Seropédica. Dados Gerais. 2019. Disponível em <<http://www.seropedica.rj.gov.br/>> Acesso em 31/01/2019.