

**ESTIMATIVA DA IRRADIAÇÃO SOLAR GLOBAL MENSAL POR MODELOS DE ANGSTROM-PRESCOTT E APRENDIZADO DE MÁQUINA EM BOTUCATU/SP/BRASIL****MAURÍCIO B. P. SILVA<sup>1</sup>, LUCIANO R. ROSSI<sup>2</sup>, CÍCERO. M. DOS SANTOS<sup>3</sup>,  
TAIZA J. ROSSI<sup>4</sup>, JOÃO F. ESCOBEDO<sup>5</sup>**<sup>1</sup>Engenheiro Agrônomo, Doutorando em Irrigação e Drenagem, UNESP-FCA/Botucatu-SP, mauricio.prado19@hotmail.com;<sup>2</sup>Biólogo, Departamento de Engenharia Rural, UNESP-FCA/Botucatu-SP;<sup>3</sup>Físico, Doutor em Irrigação e Drenagem, UNESP-FCA/Botucatu-SP;<sup>4</sup>Tecnóloga em logística e transporte, Mestrando em Energia na Agricultura, UNESP-FCA/Botucatu-SP;<sup>5</sup>Físico, Prof. Dr. Departamento de Engenharia Rural, UNESP-FCA/Botucatu-SP.

Apresentado no

XLVI Congresso Brasileiro de Engenharia Agrícola - CONBEA 2017

30 de julho a 03 de agosto de 2017 - Maceió - AL, Brasil

**RESUMO:** No presente trabalho é descrito o estudo comparativo, entre o métodos estatístico de Angstrom-Prescott (A-P)<sup>m</sup> e duas técnicas de Aprendizado de Máquina (AM) - Support Vector Machine (SVM)<sup>m</sup> e Artificial Neural Network (ANN)<sup>m</sup>, nas estimativa da irradiação solar global ( $H_G$ ) mensal. A base de dados de  $H_G$  usada na modelagem foi medida no período de 1996 a 2011 em Botucatu/SP/Brasil. A equação (A-P)<sup>m</sup> obtida ( $H_G/H_0$ ) = **0,311 + 0,366 \* (n/N)** com  $R^2 = 0,710$  foi validada usando duas bases de dados: típica e atípica. Os indicativos estatísticos  $r$  e  $rRMSE$ , obtidos na comparação entre a estimativa e medida foram:  $r_t=r_a=0,980$ ,  $rRMSE_t=3,9\%$  e  $rRMSE_a=9,8\%$ . As técnicas SVM<sup>m</sup> e ANN<sup>m</sup> foram treinadas em quatro combinações de entradas de variáveis meteorológicas, e validadas nas bases de dados típica e atípica. A comparação dos indicativos estatísticos mostra que: a técnica SVM possui melhor desempenho que o modelo (A-P) e a técnica ANN em estimar  $H_G$ . O modelo (A-P) possui melhor desempenho que a técnica ANN em estimar  $H_G$ .

**PALAVRAS-CHAVE:** Radiação solar, Angstrom-Prescott; inteligência artificial.

**ESTIMATION OF MONTHLY GLOBAL SOLAR IRRADIATION BY MODELS OF ÅNGSTRÖM -PRESCOTT AND MACHINE LEARNING IN BOTUCATU/SP/BRAZIL**

**ABSTRACT:** In this paper we describe the comparative study between the Angstrom-Prescott (AP) model with two machine learning techniques [Support Vector Machine (SVM) and Artificial Neural Network (ANN)], in the estimation of monthly global solar irradiation ( $H_G$ ). The  $H_G$  database used was measured in the period from 1996 to 2011 in Botucatu/SP/Brazil. The (A-P) model obtained ( $H_G/H_0$ ) = **0,311 + 0,366 \* (n/N)**, with  $R^2=0.710$  was validated using two databases: typical (t) and atypical (a). The statistical indicatives  $r$ ,  $rRMSE$ , obtained in the comparison between the estimation and measurement were:  $r_t=r_a=0.980$ ,  $rRMSE_t=3.9\%$  and  $rRMSE_a=9.8\%$ . The SVM and ANN techniques were trained in four combinations of inputs of meteorological variables and validated in the databases: typical and atypical. The comparison shows that: the SVM has better performance than the (A-P) model and the ANN in estimating  $H_G$ .

**KEYWORDS:** Solar radiation, Angstrom-Prescott, artificial intelligence.

**INTRODUÇÃO:** O Brasil, país de dimensão continental, vem aumentando a rede solarimétrica no território com objetivo de conhecer melhor o potencial solar disponível, para estudos nas áreas de conversões da energia solar como térmica, fotovoltaica e biomassa; na agricultura nos modelos de crescimento e produtividade de cultura agrícolas, e estimativas de evapotranspiração (Hsiao et al., 2008; Bosch et al., 2008), entre varias outras. No entanto, em locais onde a radiação global não é monitorada, o uso de modelos de estimativa é a alternativa mais usada nos projetos de pesquisa .Por isso, diversos pesquisadores vêm desenvolvendo e ajustando modelos que permitem estimar a radiação solar global a partir da medida de outra radiação ou de outra variável mais facilmente monitorada nos postos meteorológicos. Em geral, esses modelos são divididos em diferentes classes: estatísticos, físicos (transferência radiativa) e mais recentemente as técnicas de Aprendizado de Máquina (AM) (Oliveira et al., 2002; Soares et al., 2004; Escobedo et al., 2012; Santos et al., 2016).

Os modelos de AM são capazes de solucionar problemas complexos e têm sido aplicados com sucesso para previsão de irradiação solar em aplicações solares. A técnica permite modelar um sistema conhecendo apenas as variáveis de entrada e saída. A maior vantagem do uso de modelos de AM é a sua capacidade de generalização e otimização do tempo (Oliveira et al., 2006). Estimativas de  $H_G$ , por meio dos modelos de AM, têm sido estudadas e comparadas com modelos estatísticos e os resultados têm mostrado que o desempenho dos modelos de AM são similares, em alguns casos, e superiores em outros aos modelos estatísticos (Elizondo et al. 1994; Tymvios et al., 2005; Chen et al., 2013; Piri et al., 2015). Na região Sudeste do Brasil existem poucos estudos com sobre modelagem de  $H_G$  usando AM, portanto, o ajuste e a validação dos modelos de AM na estimativa de  $H_G$  é de grande interesse para o mapeamento e aproveitamento potencial da radiação solar em projetos em território brasileiro. No presente trabalho é apresentado um estudo comparativo entre o modelo estatístico de Angstrom-Prescott (A-P)<sup>m</sup> e modelos AM<sup>m</sup>.

**MATERIAIS E MÉTODOS:** Os dados utilizados neste trabalho foram medidos na Estação de Radiometria Solar na Faculdade de Ciências Agrárias (FCA) da Universidade Estadual Paulista (UNESP), localizada na cidade de Botucatu (22°53'S latitude, 48°26'W longitude e 786m de altitude). Botucatu é um município brasileiro localizado na região centro-oeste do estado de São Paulo, com uma área total de 1482.642 km<sup>2</sup>. A cidade tem elevado gradiente de altitude entre 400 e 500m na região mais baixa e entre 700 e 900 na região serrana. Esta diferença provoca variações na temperatura do ar e nos ventos. Com bioma cerrado e de mata atlântica, segundo a classificação climática de Köppen o clima da região de Botucatu é Cwa, caracterizado pelo clima tropical de altitude, verão quente e úmido com elevada precipitação, inverno seco e temperatura média do mês mais quente superior do que 22°C (Santos e Escobedo, 2016).

Foram utilizados dados de duração do brilho solar, precipitação, temperatura (máxima e a mínima) e umidade relativa do ar do período de 1996 a 2011. A irradiação solar global ( $I_G$ , W m<sup>-2</sup>), foi monitorada por um piranômetro Eppley PSP com erro 4,1% (Reda et al., 2008). Na aquisição dos dados de  $I_G$  foi utilizado um datalogger CR23X da Campbell Scientific operando na frequência de 1 Hz e armazenando médias a cada 5 minutos. Estes dados passaram por um rigoroso controle de qualidade (para eliminação de valores espúrios ou inconsistentes) através de programas desenvolvidos para cálculo da irradiação integrada no dia ( $H_G$ ) (Chaves e Escobedo, 2000). Os dados de brilho solar (n, horas), foram obtidos por um heliógrafo Campbell-Stokes; a precipitação (P, mm) medida através de um pluviógrafo da marca Ota Keiki Seisakusho; temperatura máxima e mínima do ar (T, °C) medida por termômetro de bulbo de mercúrio e álcool, respectivamente; e a umidade relativa do ar (UR,

%) medida através de um higrômetro seguindo as recomendações sugeridas pela World Meteorological Organization-WMO (1981).

#### *Modelo de Angstrom-Prescott (A-P)*

Vários modelos são sugeridos para estimar  $H_G$  utilizando a duração do brilho solar como variável de entrada. O modelo mais conhecido de estimativa da  $H_G$  foi proposto por Angstrom (1924) e posteriormente modificado por Prescott (1940), por meio da Eq. 1: onde ( $H_0$ ) é a irradiação solar no topo da atmosfera, ( $n$ ) é o brilho solar e ( $N$ ) é o fotoperíodo:

$$\frac{H_G}{H_0} = a + b \times \left(\frac{n}{N}\right) \quad (1)$$

O coeficiente “a” pode ser interpretado como a fração de  $H_G$  que atinge a superfície da Terra num dia nublado, sendo dependente do tipo e espessura de nuvens. O coeficiente “b” é um complemento que dá o total de  $H_G$ . A soma (a+b), é a fração potencial de irradiação solar no topo da atmosfera disponível para alcançar a superfície (isto é,  $H_G$  num dia de céu claro).

#### *Support Vector Machine (SVM) com o algoritmo Sequential Minimal Optimization (SMO)*

Máquina de Vetor de Suporte (SVM) é uma técnica de aprendizado supervisionado, baseado na teoria de aprendizado estatístico (Vapnik, 1995). Informações mais detalhadas sobre SVM podem ser encontradas em Vapnik (1998). Devido a sua capacidade de fornecer excelente desempenho de generalização, a SVM tornou-se uma poderosa ferramenta para resolver problemas de reconhecimento de padrões, classificação, previsão e de regressão (Shevade et al., 2000). A solução de problemas de regressão usando a SVM pode ser dada através de um algoritmo iterativo, chamado de Sequential Minimal Optimization (SMO) (Smola e Schölkopf, 1998).

#### *Multilayer Perceptron (MLP) com algoritmo Back Propagation (BP)*

A ANN utilizada foi a Multilayer Perceptron (MLP). A MLP mapeia conjuntos de dados de entrada para um conjunto de dados de saída. Essa técnica é amplamente utilizada em modelagem para resolução de problemas complexos.

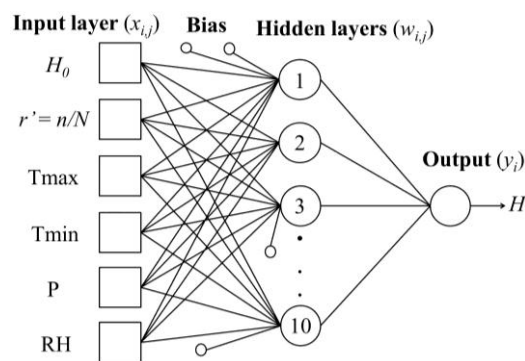


FIGURA 1. Diagrama em blocos demonstrando a arquitetura da Artificial Neural Network (ANN) utilizada, adaptado de Haykin (1998).

Na Figura. 1 é demonstrada a estrutura de uma MLP, onde a primeira camada é a entrada ( $x_{i,j}$ ), a segunda possui uma ou mais camadas ocultas de nós computacionais com pesos de ligações ( $w_{i,j}$ ) e a terceira camada corresponde à saída de nós computacionais ( $y_i$ ), (Lyra et al., 2015). Os sinais de entrada são enviados para a camada oculta. Em seguida, as camadas ocultas e de saída multiplicam os sinais de entrada por um conjunto de pesos.

A MLP típica, com uma camada oculta, pode ser modelada conforme Eq. 2 (Lam et al., 2008b):

$$y_i = \sum_{j=1}^n w_{i,j}x_{i,j} + \theta_i \quad (2)$$

Onde  $\theta_i$  é a bias do neurônio  $i$ . Cada entrada é multiplicada por um peso de ligação. A saída dos neurônios é calculada pela aplicação de uma função de ativação não linear, Eq. 3, que é tipicamente sigmóide padrão (Rehman e Mohandes, 2008).

$$f(x) = \text{sigmoid}(x) = \frac{1}{(1+\exp(-x))} \quad (3)$$

Existem muitos algoritmos de aprendizado específicos para determinados modelos de redes neurais. A MLP foi treinada usando o algoritmo de aprendizagem supervisionado Back Propagation (BP) e o termo momento. Neste algoritmo o valor de cada camada de saída é usado para atualizar o peso da camada anterior. O BP aprende de forma iterativa do processamento conjunto de exemplos de treinamentos de dados. O ajuste de peso na iteração depende da taxa de aprendizado e do momentum. A taxa de aprendizagem durante cada interação controla o tamanho das mudanças de peso e do viés.

### Software utilizado

A caixa de ferramentas do Waikato Environment for Knowledge Analysis (WEKA) foi usada para treinar e validar os dados de  $H_G$  com o algoritmo SMO para SVM e a BP para ANN. WEKA consiste num conjunto de algoritmos de AM, contendo ferramentas para pré-processamento de dados, classificação, regressão, regras de associação e visualização (Witten et al., 2011). O SMO é usado com a função de Kernel RBF para formação dos modelos. Nos modelos de ANN, com o algoritmo BP, foram considerados os valores: taxa de aprendizado = 0,3; momentum = 0,2 e número de interações = 500. As camadas ocultas (*hidden layers*) foram testadas variando de 1 – 10, porém o valor padrão do WEKA foi adotado pelo melhor ajuste encontrado. Na WEKA o padrão das camadas ocultas é definido como “ $\alpha$ ” = [(variáveis de entrada + classes)/2]. A função de ativação sigmoidal é adotada.

### Base de dados de validação dos modelos: seleção dos anos típico e atípico

Na validação dos modelos gerados neste trabalho foi utilizada uma base dados de 2 anos da irradiação solar global ( $H_G$ ), denominadas de ano típico e atípico, separados a partir da base de dados total de 16 anos (1996 a 2011). A seleção dos anos típico e atípico, foi efetuada através de análises estatísticas dos dados onde se comparou para cada mês do ano, o valor da irradiação solar global ( $H_G$ ) média inter-anos com desvio padrão de cada mês e o valor da irradiação média de cada mês por ano. O critério de seleção do ano típico é semelhante ao processo de seleção do ano meteorológico típico divulgado pela *World Meteorological Organization* (WMO) em 1981. Os resultados encontrados para os demais meses estão representados na Tabela 1, a qual mostra a constituição mensal dos anos típico e atípico do banco de dados total de 16 anos.

TABELA 1. Ano típico e atípico obtido do banco de dados de 16 anos.

|       |         | Meses |      |      |      |      |      |      |      |      |      |      |      |
|-------|---------|-------|------|------|------|------|------|------|------|------|------|------|------|
| Meses |         | Jan   | Fev  | Mar  | Abr  | Mai  | Jun  | Jul  | Ago  | Set  | Out  | Nov  | Dez  |
| Ano   | Típico  | 2004  | 2000 | 2008 | 2005 | 2005 | 2005 | 2002 | 2003 | 2000 | 2003 | 2003 | 2005 |
|       | Atípico | 2001  | 2005 | 2011 | 2008 | 2003 | 1997 | 2009 | 1998 | 2009 | 2001 | 1998 | 2011 |

### Indicativos Estatísticos

Existem vários índices estatísticos utilizados para avaliar o desempenho de modelos. Na performance dos modelos de radiação solar com técnicas de aprendizados de maquina normalmente são: Root Mean Square Error (RMSE), Relative Root Mean Square Error (rRMSE), coeficiente de correlação ( $r$ ), Coeficiente de determinação ( $R^2$ ) (Santos et al., 2016). Uma escala classificatória para os diferentes intervalos de rRMSE para avaliar o desempenho dos modelos é utilizada (Jamieson et al., 1991; Li et al., 2013): Excelente se

rRMSE<10%; Bom se  $10\% \leq rRMSE < 20\%$ ; Aceitável se  $20\% \leq rRMSE < 30\%$ ; Pobre se  $rRMSE \geq 30$ .

## RESULTADOS E DISCUSSÃO

*Correlação entre a transmissividade atmosférica da irradiação solar global ( $H_G/H_0$ ) e a razão de insolação ( $n/N$ ): Modelo de Angstrom-Prescott (A-P)*

A Figura. 2 mostra a correlação entre a transmissividade atmosférica da irradiação solar global ( $H_G/H_0$ ) e a razão de insolação ( $n/N$ ) para a base de dados medida no período de 1996-2011 em Botucatu, SP. A correlação é linear em todo intervalo de variação de ( $n/N$ ), entre 0 e 1. O espalhamento da correlação é semelhante à maioria dos trabalhos da equação de Angstrom-Prescott em diversos países (Martinez-Lozano, 1984; Akinoglu, 1990; Bakirci, 2009a).

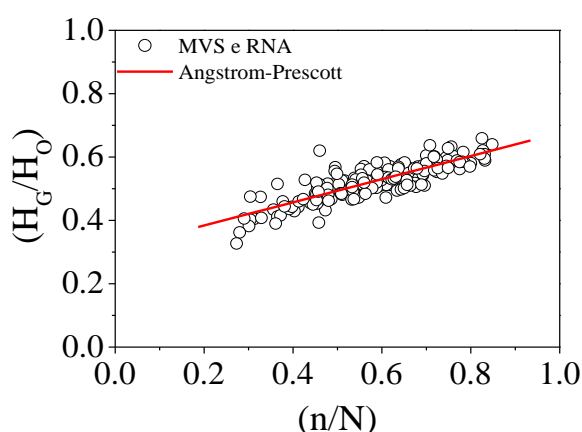


FIGURA 2. Correlação entre a transmissividade atmosférica da irradiação solar global ( $H_G/H_0$ ) e a razão de insolação ( $n/N$ ), e a reta obtida por regressão linear no período de 1996-2011.

A Eq. 4 de (A-P) obtida na correlação da Fig. 2, por regressão linear, foi:

$$(H_G/H_0) = 0,311 + 0,366 \times (n/N) \quad (4)$$

O coeficiente de determinação ( $R^2$ ) = 0,714, bem como o valor da fração de  $H_G$  mínima ( $a = 0,311$ ), do complemento que dá o valor total de  $H_G$  ( $b = 0,366$ ) ou a fração atmosférica máxima ( $a+b = 0,677$ ) diferem dos valores obtidos no intervalo de variação dos coeficientes em diversas localidades do Brasil ( $a = 0,248 \pm 0,081$ ), ( $a+b = 0,689 \pm 0,124$ ) e  $R^2$  cuja média é de 0,793 ( $\pm 0,162$ ) (Tabela 2). Fatores climáticos, como tipo e espessuras de nuvens, concentração de vapor de água e aerossóis na atmosfera são os fatores responsáveis pelas variações de  $a$  e  $a+b$  em cada local.

TABELA 2. Coeficientes de Angstrom-Prescott (A-P) obtidos no território brasileiro por diversos autores.

| Autores                       | Localidades (N°)         | Coeficientes de (A-P) |             |             |             |
|-------------------------------|--------------------------|-----------------------|-------------|-------------|-------------|
|                               |                          | <i>a</i>              | <i>b</i>    | <i>a+b</i>  | $R^2$       |
| Tiba (2001)                   | 34 (NE)                  | 0,22 - 0,35           | 0,31 - 0,58 | 0,58 - 0,84 | 0,71 - 0,98 |
| Ribeiro et al. (1982)         | Manaus (AM)              | 0,26                  | 0,48        | 0,74        | 0,69        |
| Souza et al. (2016)           | 3 (AL)                   | 0,16 - 0,39           | 0,28 - 0,58 | 0,68 - 0,75 | *           |
| Tarifa (1972)                 | Presidente Prudente (SP) | 0,19                  | 0,39        | 0,58        | *           |
| Cervellini et al. (1966)      | 2 (SP)                   | 0,19 - 0,23           | 0,56 - 0,61 | 0,79 - 0,80 | *           |
| <b>Modelo proposto (2015)</b> | <b>Botucatu (SP)</b>     | <b>0,31</b>           | <b>0,37</b> | <b>0,68</b> | <b>0,71</b> |

Siglas: AL = Alagoas, AM = Amazonas, SP = São Paulo, NE = Região Nordeste do Brasil.  
 - Informação não fornecida pelos autores.

A Figura. 3 (a, b) mostra as correlações obtidas na validação entre as estimativas e as medidas de  $H_G$ , para as bases de dados típico e atípico. As distribuições dos valores estimados pela equação (A-P) e as medidas nas duas condições de validação estão em concordância linear com a reta ideal ( $45^\circ$ ) ou  $r = 1$  no ajuste. As retas de regressão linear com os coeficientes de correlação iguais  $r^t = r^a = 0,980$  mostram que a equação (A-P) pode estimar  $H_G$  com elevados coeficientes de determinação, próximos de 1:  $(R^2)^t = 0,960$  (Fig. 3 a) e com  $(R^2)^a = 0,960$  (Fig. 3 b).

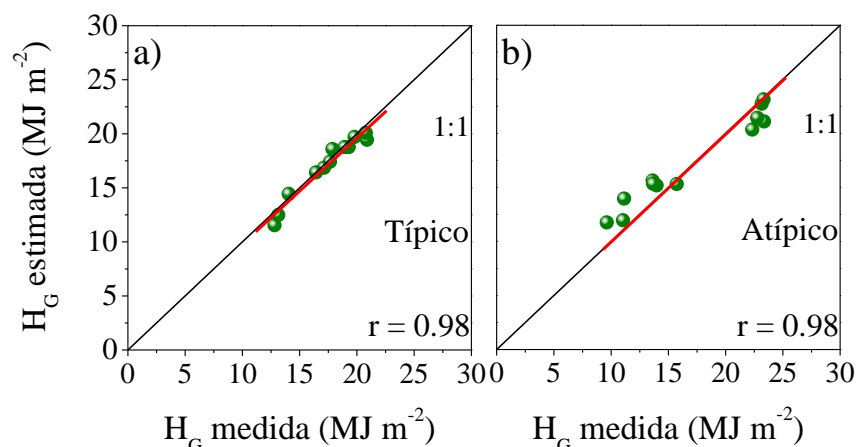


FIGURA 3 (a, b). Comparação entre os valores estimados por meio de Angstrom-Prescott (A-P) e medidas de  $H_G$  para anos típico e atípico.

O indicativo estatístico RMSE obtidos na validação entre as estimativas e as medidas de  $H_G$ , para as bases de dados típico e atípico foram:  $RMSE^t = 3,9\%$  e  $RMSE^a = 9,8\%$ .

#### Treinamento e validação dos modelos das técnicas de SVM e ANN

O mesmo banco de dados da geração do modelo de A-P (Fig. 2) foi usado no treinamento das técnicas SVM e ANN. O Modelo de Angstrom-Prescott e os modelos de SVM1 e ANN1 (combinação1) utilizam as mesmas variáveis de entrada  $H_0$ ,  $N$  (calculados) e  $n$  (medido). Além da combinação1, foram treinadas três novas combinações (SVM2, SVM3 e SVM4; ANN2, ANN3 e ANN4), conservado a mesmas variáveis da combinação 1, e acrescentando na sequência as seguintes variáveis de entrada: temperatura máxima e mínima do ar ( $T$ ,  $^\circ C$ ), precipitação ( $P$ , mm), Umidade Relativa ( $UR$ , %), como mostra a Tabela 3.

TABELA 3. Variáveis de entrada para os modelos de SVM e ANN.

| Modelo       | SVM  | ANN  | Variáveis de entrada                               |
|--------------|------|------|--|
| Combinação 1 | SVM1 | ANN1 | $H_0$ , $n/N$                                      |
| Combinação 2 | SVM2 | ANN2 | $H_0$ , $T_{máx}$ , $T_{mín}$ , $n/N$              |
| Combinação 3 | SVM3 | ANN3 | $H_0$ , $T_{máx}$ , $T_{mín}$ , $P$ , $n/N$        |
| Combinação 4 | SVM4 | ANN4 | $H_0$ , $T_{máx}$ , $T_{mín}$ , $P$ , $UR$ , $n/N$ |

$H_0$  = irradiação solar no topo da atmosfera,  $n/N$  = razão de insolação,  $T_{máx}$  = temperatura máxima do ar,  $T_{mín}$  = temperatura mínima do ar,  $P$  = precipitação pluvial e  $UR$  = umidade relativa do ar.

A Fig. 4 mostra o diagrama de dispersão entre os valores estimados de  $H_G$  pelos modelos SVM e ANN e as medidas, bem como as retas obtidas por regressão linear e coeficientes de correlação ( $r$ ), para os anos típico ( $r^t$ ) e atípico ( $r^a$ ), respectivamente.

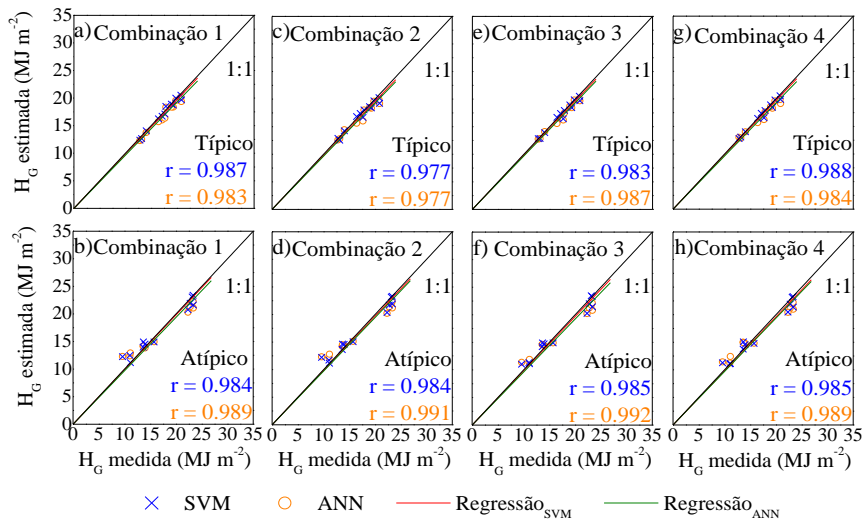


FIGURA 4 (a-h). Correlação entre os valores estimados pelos modelos de AM, e os valores medidos de  $H_G$ .

Os valores estimados por SVM e ANN nas quatro (4) combinações estão em concordância linear com as medidas (Fig. 4 a, b): os valores de coeficientes de correlação ( $r$ ) obtidas nas regressões lineares estão muito próximos da reta ideal ( $45^\circ$ ) ou  $r = 1$  no ajuste.

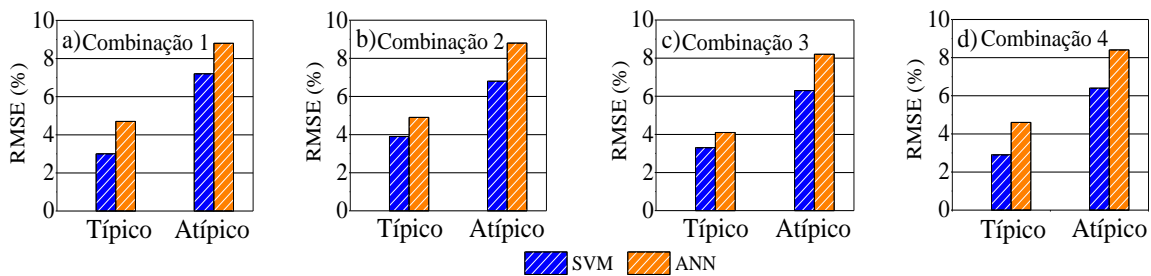


FIGURA 5 (a-d). Indicativo estatístico RMSE (%) para anos típico e atípico.

Para a combinação 1 (SVM1 e ANN1), os valores de  $r^t = 0,987$  e  $r^a = 0,984$ ;  $r^t = 0,983$  e  $r^a = 0,989$ , respectivamente, mostram que a medida e a estimativa de  $H_G$  estão bem correlacionadas para todas as combinações. A técnica SVM1 apresentou maior valor de  $r$  que a rede ANN1, na validação do ano típico e menor no atípico. O indicativo estatístico rRMSE, obtido da comparação dos valores estimados gerados com as técnicas de AM e as medidas de  $H_G$  respectivamente para os anos típico e atípico, são mostrados na Fig. 4:  $rRMSE^t_{SVM1} = 3,0\%$  e  $rRMSE^a_{SVM1} = 7,2\%$ , enquanto para  $rRMSE^t_{ANN} = 4,7\%$  e  $rRMSE^a_{ANN} = 8,8\%$ .

#### *Comparação do desempenho entre os modelos Angstrom-Prescott (A-P) e as redes MVS e RNA mensal em Botucatu e de outros locais*

A Tabela 4 mostra a comparação dos indicativos estatísticos RMSE, RrMSE, em unidades de energia por mês e porcentagem, e, também, os valores de  $r$  e  $R^2$ , obtidos em duas condições validações, típico e atípico, entre medida e estimativa dos modelos estatístico de Angstrom-Prescott e das AM's MVS e RNA (mensal) em Botucatu, bem como para outros locais no globo terrestre. Os valores dos indicativos estatísticos rRMSE, RMSE e  $R^2$  do modelo de A-P e das técnicas de AM (MVS e RNA), obtidas em Botucatu, mostram que os modelo MVS1<sup>m</sup> apresenta melhor resultado que o modelo estatístico (A-P)<sup>m</sup> e o RNA1<sup>m</sup> nas duas condições de validação, respectivamente.

TABELA 4. Comparações entre os modelos estatísticos de A-P com as técnicas de aprendizado de máquina (AM).

| Autores                               | Localidades (N°)         | Modelos                | rRMSE (%)          | RMSE (MJ m <sup>-2</sup> ) | R <sup>2</sup>     |
|---------------------------------------|--------------------------|------------------------|--------------------|----------------------------|--------------------|
| Mohammadi et al. (2015a) <sup>m</sup> | Isfahan (Irã)            | A-P                    | 6,0                | 1,16                       | 0,96               |
| Mohammadi et al. (2015a) <sup>m</sup> | Isfahan (Irã)            | MVS                    | 2,26               | 0,45                       | 0,99               |
| Tymvios et al. (2005) <sup>m</sup>    | Nicósia (Chipre)         | A-P                    | 13,36              | -                          | 0,85               |
| Tymvios et al. (2005) <sup>m</sup>    | Nicósia (Chipre)         | RNA                    | 5,67 a 10,15       | -                          | 0,74 a 0,91        |
| <b>Modelo mensal proposto</b>         | <b>Botucatu (Brasil)</b> | <b>A-P<sup>m</sup></b> | <b>3,94 a 9,76</b> | <b>0,67 a 1,66</b>         | <b>0,96 a 0,96</b> |
| <b>Modelo mensal proposto</b>         | <b>Botucatu (Brasil)</b> | <b>MVS<sup>m</sup></b> | <b>3,0 a 7,2</b>   | <b>0,51 a 1,22</b>         | <b>0,96 a 0,98</b> |
| <b>Modelo mensal proposto</b>         | <b>Botucatu (Brasil)</b> | <b>RNA<sup>m</sup></b> | <b>4,7 a 8,8</b>   | <b>0,82 a 1,50</b>         | <b>0,96 a 0,98</b> |

A comparação entre SVM1 x (A-P) mostra que os valores de  $rRMSE_{SVM1}^t = 3,0\%$  e  $rRMSE_{SVM1}^a = 7,2\%$  são inferiores aos de  $rRMSE_{(A-P)}^t = 3,94\%$  e  $rRMSE_{(A-P)}^a = 9,76\%$  (Fig. 6 b); e os de  $(R^2)_{SVM1}^t = 0,974$  e  $(R^2)_{SVM1}^a = 0,968$  são superiores aos de  $(R^2)_{(A-P)}^t = 0,953$  e  $(R^2)_{(A-P)}^a = 0,960$ . Os resultados obtidos para o indicativo estatístico rRMSE para SVM1 e (A-P): neste trabalho estão de acordo com os resultados dos trabalhos por Mohammadi et al. (2015a) em Isfahan no Irã com  $rRMSE_{SVM} = 2,3\%$  e  $rRMSE_{(A-P)} = 6,0\%$  e valores de  $R^2_{SVM} = 99,0\%$ , e  $R^2_{(A-P)} = 96\%$ .

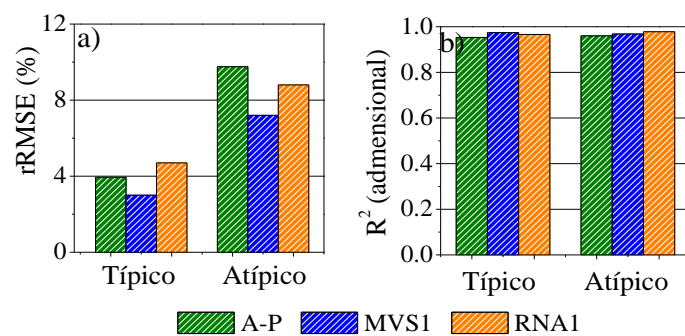


FIGURA 6 (a, b). Indicativos estatísticos rRMSE e R<sup>2</sup> para anos típico e atípico.

A comparação entre (A-P) x ANN1 mostra que: os valores de  $rRMSE_{(A-P)} = 3,94\% - 9,76\%$  são inferiores a  $rRMSE_{ANN4} = 4,7\% - 8,8\%$ , os valores de  $R^2_{(A-P)} = 96,0\% - 96,0\%$  são inferiores a  $R^2_{ANN4} = 96,0\% - 98,0\%$ . Os resultados obtidos neste trabalho para os indicativos estatísticos  $rRMSE_{(A-P)} = 3,94\%$  a  $9,76\%$ ;  $R^2_{(A-P)}$  de  $96\%$  a  $96\%$ ;  $rMBE_{ANN4} = -3,8\% - 0,9\%$ ;  $rRMSE_{ANN4} = 4,7\%$  a  $8,8\%$ ; e  $R^2_{ANN4}$  de  $96\%$  a  $98\%$  são na mesma ordem de grandeza aos de Tymvios et al. (2005) em Nicósia (Chipre), cujos resultados para os indicativos estatísticos  $rRMSE_{(A-P)} = 13,36\%$ ;  $rRMSE_{ANN4} = 5,67\%$  a  $10,15\%$ .

**CONCLUSÃO:** A equação de estimativa (A-P), obtidas por meio de regressão linear em Botucatu, SP, com coeficientes de determinação  $R^2 = 0,715$  é da mesma ordem de grandeza aos valores de  $R^2$  obtidos por outros pesquisadores e outros locais do globo terrestre. Os coeficientes de correlação,  $r = 0,976$  e  $r = 0,980$ , para validação com o ano típico e atípico, mostram que o modelo de A-P pode estimar  $H_G$  com elevados coeficientes de determinação.

O treinamento das técnicas SVM1 e a ANN1 com mesmas variáveis de entrada  $H_0$  e  $n/N$  que o modelo A-P com valores de  $r$  próximos dos 100% mostram que as redes podem estimar  $H_G$  na mesma ordem de grandeza que a equação de Angstrom-Presecott clássica.

Na comparação entre desempenho para os modelos A-P, SVM1 e ANN1 por meio dos valores dos indicativos estatísticos rRMSE, RMSE,  $r$  e  $R^2$  obtidos da validação mostrou que: entre o modelo A-P e a técnica SVM1 em Botucatu, a técnica SVM1 apresentou melhor resultado que o modelo estatístico de A-P; Entre os modelos SVM1 e ANN1 em Botucatu a técnica

SVM1 apresentou melhor resultado do que a rede ANN1; entre o modelo A-P e a rede ANN1 em Botucatu o modelo estatístico A-P, apresentou no geral melhor resultado do que a rede ANN1.

## REFERÊNCIAS

- AKNOGLU, B. G. A review of sunshine-based models used to estimate monthly average global solar radiation. *Renewable Energy*, v. 1, n. 3/4, 479-497, 1990.
- Angstrom, A. Solar and terrestrial radiation. *Quarterly Journal of the Royal Meteorological Society*, v. 50, p. 121-125, 1924.
- BAKIRCI, K. Correlations for estimation of daily global solar radiation with hours of bright sunshine in Turkey. *Energy*, v. 34, 485-501, 2009a.
- BAKIRCI, K. Models of radiation with hours of bright sunshine: A review. *Renewable and Sustainable*, v. 13, p. 2580-2588, 2009b
- BOSCH, J. L.; LÓPEZ, G.; BATLLES, F. J. Daily solar irradiation estimation over a mountainous area using artificial neural networks. *Renewable Energy*, v. 33, 1622-1628, 2008.
- CERVELLINI, A.; SALATI, E. Estimativa da distribuição da energia solar no Estado de São Paulo. *Bragantia*, v. 25, n. 3., 1966.
- CHAVES, M. A.; ESCOBEDO, J. F. A software to process daily solar radiation data. *Renewable Energy*, v. 19, 339-344, 2000.
- CHEN, J-L.; LI, G-S.; WU, S-J. Assessing the potential of support vector machine for estimating daily solar radiation using sunshine duration. *Energy Conversion and Management* v.75, 311–318, 2013.
- ELIZONDO, D.; HOOGENBOOM, G.; MCCLENDON, R. W. Development of a neural network model to predict daily solar radiation. *Agricultural and Forest Meteorology*, v. 71, 115-132, 1994.
- ESCOBEDO, J. F.; TERAMOTO, E. T.; OLIVEIRA, A. P. et al. Equações de estimativa das frações solar direta (kdh) e difusa (kd) em função do índice de claridade (kt) e razão de insolação (n/N), *Avances en y Medio Ambiente*, Vol. 16, 2012.
- HAYKIN, S. *Neural networks: A comprehensive foundation*. 2nd ed. pp. 897. Hamilton: Prentice Hall, 1998.
- HSIAO, T.; HENG, L.; STEDUTO, P. et al. Aqua Crop – The FAO crop model to simulate yield response to water: III. Parameterization and testing for maize. *Agronomy Journal*, v. 101, 448-459, 2008.
- JAMIESON, P. D.; PORTER, J. R.; WILSON, D. R. A test of the computer simulation model ARC - WHEAT1 on wheat crops grown in New Zealand. *Field Crops Research*, v.27, 337-350, 1991.
- LI, M.; TANG, X.; WU, W. et al. General models for estimating daily global solar radiation for different solar radiation zones in mainland China. *Energy Conversion and Management*, v.70, 139-148, 2013.
- LAM, J. C.; WAN, K. K. W.; Yang, L. Solar radiation modeling using ANNs for different climates in China. *Energy Conversion & Management*, v.49, 1080-1090, 2008b.
- MARTINEZ-LOZANO, et al. The historical evolution of the Angstrom Formula and its modifications. *Agricultural and Forest Meteorology*, v. 33, n. 2-3, 109-128, 1984.
- NICÁCIO, R. M.; SOUZA, J. L.; BERNARDO, S. O. Estimation da irradiância solar global para Maceió utilizando o modelo linear de Angstrom-Prescott. In: XII Congresso Brasileiro de Agrometeorologia – Fortaleza, 3 a 6 de 2001.
- OLIVEIRA, A. P.; ESCOBEDO, J. F.; MACHADO et al. Correlation model of diffuse solar-radiation applied to the city of São Paulo, Brazil. *Applied Energy*, v. 71, 59-73, 2002.
- OLIVEIRA, A. P.; SOARES, J.; BOZNAR, M. Z.; MLAKAR, P.; ESCOBEDO, J. F. An application of Neural Network technique to correct the dome temperature effects on pyrgeometer measurements. *Journal of Atmospheric and Oceanic Technology*, v. 23, 80-89, 2006.

MOHAMMADI, K.; SHAMSHIRBAND, S.; ANISI, M. H. et al. Support vector regression based prediction of global solar radiation on a horizontal surface. *Energy Conversion and Management*, v.91, 433-441, 2015a.

PIRI, J.; SHAMSHIRBAND, S.; PETKOVIC, D. et al. Prediction of the solar radiation on the Earth using support vector regression technique. *Infrared Physics & Tecnology*, v. 68, 179-185, 2015.

PRESCOTT, J. A.. Evaporation from a water surface in relation to solar radiation. *Transactions of the Royal Society Science Australian*, v. 64, 114-118, 1940.

REDA, I.M.; MYERS, D.R.; STOFFEL, T.L. Uncertainty estimate for the outdoor calibration of solar pyranometers: NCSLI Measure. *The Journal of Measurement Science*, v. 3, n. 4, 58-66, 2008.

REHMAN, S.; MOHANDDES, M. Artificial neural network estimation of global solar radiation using air temperature and relative humidity. *Energy Policy*, v. 36, 571-576, 2008.

RIBEIRO, M. N. G.; SALATI, E.; VILA NOVA, N. A.; DEMÉTRIO, C. G. B. Radiação solar disponível em Manaus (AM) e sua relação com a duração do brilho solar. *Acta Amazônica*, v. 12, n. 2, p. 339-346, 1982.

SHEVADE, S. K.; KEERTHI, S. S.; BHATTACHARYYA, C. Improvements to the SMO Algorithm for SVM Regression. *IEEE Transactions on Neural Networks*, v. 11, n. 5, 2000.

SMOLA, A. J.; SCHÖLKOPF, B. "A Tutorial on Support Vector Regression". Royal Holloway College, London, U. K., Neuro COLT Tech. Rep. TR 1998-030, 1998.

SOARES, J.; OLIVEIRA, A. P.; BOZNAR, M. Z.; MLAKAR, P.; ESCOBEDO, J. F.; Machado, A. J. Modeling hourly diffuse solar-radiation in the city of São Paulo using a neural-network technique, *Applied Energy*, v.79, 201-214, 2004.

SANTOS, C. M.; ESCOBEDO, J. F. Temporal variability of atmospheric turbidity and DNI attenuation in the sugarcane region, Botucatu/SP. *Atmospheric Research*, v.181, 312–321, 2016.

SANTOS, C. M.; ESCOBEDO, J. F.; TADAO, E. T.; SILVA, S. H. M. G. Assessment of ANN and SVM models for estimating normal directirradiation ( $H_b$ ). *Energy Conversion and Management*, v.126, 826–836, 2016.

SOUZA, J. L.; LYRA, G. B.; SANTOS, C. M. et al. Empirical models of daily and monthly global solar irradiation using sunshine duration for Alagoas State, Northeastern Brazil. *Sustainable Energy Technologies and Assessment*, v. 14, p. 35-35, 2016.

TARIFA, J. R. Estimativa da radiação em função da insolação para Presidente Prudente. *Caderno de Ciências da Terra*, n. 22, p. 1-15, 1972.

TIBA, C. Solar radiation in the Brazilian Northeast. *Renewable Energy*, v. 22, p. 565-578, 2001.

TYMVIOS, F. S.; JACOVIDES, C. P.; MICHAELIDES, S. C. et al. Comparative study of Angstrom and artificial neural networks methodologies in estimating global solar radiation. *Solar Energy*, v. 78, 752-762, 2005.

VAPNIK, V. N. *The nature of Statistical learning theory*. Springer-Verlag, New York, 1995.

VAPNIK V, N. *Statistical learning theory*. New York: Wiley, 1998.

World Meteorological Organization. *Meteorological Aspects of the utilization of Solar Radiation as an Energy Source*. World Meteorological Organization Technical Note No. 172, WMO-No. 557, Geneva, pp. 298., 1981.

WITTEN, I. H.; FRANK, E.; HALL, M. A. *Data Mining: Practical Machine Learning Tools and Techniques*. 3 rd ed. pp. 630, 2011.