

# ESTIMATIVA DA IRRADIAÇÃO SOLAR GLOBAL PELOS MODELOS DE HARGREAVES-SAMANI E APRENDIZADO DE MÁQUINA SVM E ANN EM BOTUCATU/SP/BRASIL

Angélica Castilho Paes - gecacp@hotmail.com

Maurício Bruno Prado da Silva - mauricio.prado19@hotmail.com

Eduardo Nardini Gomes - engomes@registro.unesp.br

João Francisco Escobedo - escobedo@fca.unesp.br

Universidade Estadual Paulista, FCA-Botucatu, Departamento de Engenharia Rural

Cícero Manoel dos Santos - ciceromanoel2007@gmail.com

Faculdade de Engenharia Agrônômica - UFPA/Altamira/PA/Brasil

**Resumo.** Neste trabalho é apresentado um estudo comparativo entre dois métodos de estimativas da radiação solar global, modelo estatístico de Hargreaves-Samani (H-S) e os modelos de aprendizado de máquinas Artificial Neural Network (ANN) e Support Vector Machine (SVM), para três combinações de entradas de variáveis meteorológicas. Por meio dos valores dos indicadores estatísticos  $r$ ,  $rMBE$ ,  $rRMSE$  e  $d$  de Wilmott, obtidos na validação dos modelos, foi conclusivo que: a técnica SVM3 mostrou melhor desempenho que a rede ANN3 e do modelo estatístico (H-S) em estimar  $H_G$ , enquanto que a técnica ANN3 possui desempenho superior ao modelo estatístico (H-S) em estimar  $H_G$ .

**Palavras - chave:** Radiação solar, modelo Hargreaves-Samani (H-S), modelagem estatística; inteligência artificial, variáveis meteorológicas.

## 1. INTRODUÇÃO

A radiação solar é um importante parâmetro para várias atividades do setor produtivo e de pesquisa, principalmente em processos de conversão de energia térmica, fotovoltaico e biomassa. No Brasil, país de dimensões continentais, as medidas de radiação solar ainda são restritas a centros de pesquisas e universidades. Uma opção para locais onde não são efetuadas medidas da radiação global, é o uso de modelos de estimativas, desenvolvidos para se estimar a radiação a partir de outros parâmetros meteorológicos medidos com maior frequência em postos meteorológicos (Santos *et al.*, 2014; Souza *et al.*, 2016; Silva *et al.*, 2017). Dentre os modelos mais importantes estão os estatísticos, transferência radiativa e técnicas de aprendizados de máquinas.

Atualmente, existem muitos modelos estatísticos, sendo que o mais importante é o modelo de Angstrom- Prescott (A-P), que estima a radiação solar global a partir da medida do brilho solar. O modelo A-P obtido por regressão linear é simples, mas atualmente, tem a desvantagem de usar o heliógrafo na medida do brilho solar, tendo em vista que o instrumento ainda é importado a custo elevado. O heliógrafo possui ainda várias limitações de natureza experimental, o qual não é tão preciso principalmente em regiões tropicais onde a umidade relativa do ar é elevada. Outro modelo importante é o desenvolvido por Hargreaves & Samani (1982), o qual estima a radiação solar global a partir de medidas da temperatura máxima e mínima, e da radiação extraterrestre. A medida da temperatura é mais simples e possui baixo custo, além de que, os termômetros são instalados até em estações meteorológicas ordinárias de muitas cidades, aumentando o campo espacial de estudos. A grande desvantagem desses modelos de regressão linear é que são precisos apenas para locais específicos de origem das medidas. Outro modelo que usa a amplitude térmica da temperatura do ar como variável de entrada é o de Bristow & Campbell (1984). Há também outros modelos modificados em outros locais que associam a temperatura do ar a outras variáveis de entrada (Hunt *et al.*, 1998; Bechini *et al.*, 2000; Almorox *et al.*, 2011).

Os modelos de transferência radiativa são mais complexos que os modelos estatísticos e exigem entrada mista de parâmetros meteorológicos medidos na superfície terrestre e na atmosfera por satélites (Madkour *et al.*, 2006). A desvantagem destes modelos é a necessidade de informações detalhadas dos constituintes atmosféricos como da concentração de ozônio, aerossóis, gases e vapor d'água na atmosfera que são de difícil acesso (Gueymard, 2003). Os modelos de transferência radiativa necessitam ainda de extenso trabalho computacional nos complexos processos de transferência atmosférica devido à variação temporal e espacial dos aerossóis e vapor d'água (Dai e Fang, 2014; Zhang *et al.*, 2014).

Os modelos de AM são capazes de solucionar problemas que os demais modelos não solucionam e têm sido aplicados com sucesso para previsão de irradiação solar e aplicações solares, principalmente nas relações em que são baixas as correlações entre as radiações (Rehman e Mohandes, 2008). 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). As principais técnicas de AM que vêm sendo utilizadas

são: Adaptive Neuro-Fuzzy Inference System (ANFIS), Artificial Neural Network (ANN), Genetic Programming (GP) e Support Vector Machine (SVM). Estudos recentes, concluíram que a técnica SVM é superior à ANN, ANFIS, na estimativa da radiação solar. Especificamente, as estimativas de radiação global, por meio dos modelos de SVM e ANN têm sido estudadas e comparadas com modelos estatísticos, e os resultados têm mostrado que o desempenho dos modelos de AM (SVM e ANN) 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; Quej *et al.*, 2017; Silva *et al.*, 2017). No Brasil existem vários estudos com modelagem da radiação solar usando AM, como mostram os trabalhos (Silva *et al.*, 2017; Santos *et al.*, 2016; Lima *et al.*, 2016).

Diante da necessidade de se conhecer melhor a eficiência dos modelos estatísticos e de AM, este trabalho teve como objetivo realizar um estudo comparativo entre o modelo de Hargreaves-Samani (H-S), e as técnicas Artificial Neural Network (ANN) e Support Vector Machine (SVM) de desempenho nas estimativas da radiação global.

Uma base de dados de temperaturas máxima e mínima, e de irradiação solar global diária de 16 anos, período 1996 a 2011, foi utilizada para gerar e validar os modelos de H-S e das técnicas ANN e SVM. Para as técnicas ANN e SVM foram estudadas três diferentes combinações combinando variáveis de entrada no modelo: a combinação ANN1 e SVM1 têm como variáveis de entrada a amplitude de temperatura e a irradiação solar no topo da atmosfera igual ao do modelo H-S. Nas combinações seguintes 2 e 3 foram acrescentadas outras variáveis, uma a uma, como razão de insolação e umidade relativa do ar, respectivamente. Os resultados obtidos da validação, comparação entre medidas e estimativas dos modelos, por meio de correlação ( $r$ ) e indicadores estatísticos ( $r$ MBE,  $r$ RMSE e  $d$  de Willmott), permitiram classificar o desempenho dos modelos nas estimativas de radiação global.

## 2. MATERIAL E MÉTODOS

Os dados utilizados neste trabalho foram: irradiância solar global, temperatura (máxima e mínima) e umidade relativa, coletados em Botucatu, no período de 1996-2011. As medidas de irradiâncias solar global ( $I$ ,  $W\ m^{-2}$ ) foram realizadas por um piranômetro Eppley, a temperatura máxima e mínima do ar ( $T$ ,  $^{\circ}C$ ) foi medida por termômetro de bulbo de mercúrio e álcool, respectivamente; e a umidade relativa do ar ( $UR$ , %) medida através de um higrômetro.

### 2.1 Modelo Hargreaves-Samani

O modelo estatístico desenvolvido por Hargreaves & Samani (1982) usado neste trabalho foi desenvolvido para estimar  $H_G$  através de dados de temperatura mínima e máxima e da radiação solar extraterrestre, como mostra a Eq. (1), sendo  $H_0$  a radiação solar extraterrestre,  $T_{max}$  e  $T_{min}$  a temperatura máxima e a temperatura mínima, respectivamente.

$$\frac{H_G}{H_0} = a \cdot (T_{max} - T_{min})^{0,5} \quad (1)$$

### 2.2 Software utilizado

O software utilizado para realizar o aprendizado de máquinas foi o Waikato Environment for Knowledge Analysis (WEKA). Para os modelos de ANN foi utilizado o algoritmo BP e a função de ativação sigmoideal, com os seguintes valores: taxa de aprendizado = 0,3; momentum = 0,2 e número de interações = 500. Para o número de camadas ocultas foi utilizado o padrão do Weka, definido como: “ $\alpha$ ” = [(variáveis de entrada + classes)/2]. Para a SVM foi utilizado o algoritmo SMO com a função de Kernel RBF.

### 2.3 Multilayer Perceptron

Rede neural artificial é uma técnica de aprendizado de máquinas baseado no funcionamento do cérebro. A rede possui três grupos de neurônios: os neurônios de entrada na camada de entrada, os neurônios ocultos, e os neurônios de saída, na camada de saída da rede, sendo que pode existir mais de uma camada de neurônios ocultos na rede (Ramedani *et al.*, 2012). A rede utilizada neste trabalho é a Multilayer Perceptron (MLP), um tipo de rede de aprendizado supervisionado.

Os neurônios multiplicam seu valor de entrada ( $x_i$ ) pelo valor do peso da conexão ( $w_{ji}$ ), e então são somados os produtos realizados com cada neurônio e o resultado é colocado em uma função de transferência, que determinará a resposta de saída ( $y_i$ ) (Citakoglu, 2015). A Eq. (2) apresenta um exemplo deste processo utilizando-se uma função sigmoideal:

$$y_i = f(\sum w_{ji} \cdot x_i) = \frac{1}{1 + e^{-\sum w_{ji} x_i}} \quad (2)$$

O algoritmo utilizado na rede foi o Back Propagation, que realiza a retropropagação dos resultados para o ajuste dos pesos a cada iteração da rede, até que o resultado esteja o mais próximo do esperado.

## 2.4 Support Vector Machine (SVM)

A máquina de vetor de suporte é uma técnica de aprendizado de máquinas desenvolvida inicialmente para resolver problemas de classificação de dados, mas que pode também pode realizar regressões. A máquina mapeia os dados de forma não linear em um espaço multidimensional, gerando uma superfície de decisão linear (Cortes e Vapnik, 1995).

O uso das funções de Kernel para o treino da SVM apresenta facilidades para se trabalhar com a SVM, porque elas convertem os dados de um espaço dimensional de poucas dimensões, dos dados de origem, em um espaço multidimensional de forma implícita.

Uma das funções de Kernel é a Função de Base Radial (RBF), que foi utilizada neste trabalho. A vantagem do seu uso é que a função RBF faz uso de soluções de um conjunto de equações lineares ao contrário de outros casos que utilizam soluções quadráticas (Mohammadi *et al.*, 2015).

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

Na validação do modelo foi utilizada uma base dados de 1 ano da irradiação solar global diária ( $H_G$ ), denominadas de ano típico e atípico, selecionados a partir da base de dados total de 16 anos (1996-2011). 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.

Tabela 1- Ano típico a atípico.

		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

O processo de seleção do ano típico é similar ao processo de seleção do Ano Meteorológico Típico (WMO, 1981). Em cada mês do ano típico, os valores das irradiações media-inter-anos praticamente coincidem com os valores das irradiações médias do ano. Ao contrário, no ano atípico, os valores das irradiações médias do ano estão mais distantes dos valores das irradiações médias inter-anos, externos ao intervalo de variação do desvio padrão inter-anos. Os resultados encontrados para os meses dos anos típico e atípico estão representados na Tab. 1 (Silva *et al.*, 2017).

## 2.6 Indicativos estatísticos

Os indicativos estatísticos rMBE (Relative Mean Bias Error), rRMSE (Root Mean Square Error) e  $d$  (Índice de Concordância de Willmott) empregados na avaliação da performance das equações de estimativa diária da irradiação global infravermelha foram:

$$rMBE (\%) = 100 \times \frac{\left( \frac{\sum_{i=1}^N (y_i - x_i)}{N} \right)}{\bar{x}} \quad (1)$$

$$rRMSE (\%) = 100 \times \left[ \frac{\sum_{i=1}^N (y_i - x_i)^2}{N} \right]^{\frac{1}{2}} \quad (2)$$

$$d = 1 - \frac{\sum_{i=1}^N (P_i - O_i)^2}{\sum_{i=1}^N (|P_i| + |O_i|)^2} \quad (3)$$

onde:  $P_i$  representa os valores estimados,  $O_i$  os valores medidos,  $N$  o número de observações,  $|P_i|$  o valor absoluto da diferença  $P_i - \bar{O}_i$ , e  $|O_i|$  representa o valor absoluto da diferença  $O_i - \bar{O}_i$ .

## 3. RESULTADOS E DISCUSSÃO

### 3.1 Correlação entre a transmissividade atmosférica da irradiação solar global ( $H/H_0$ ) e a amplitude térmica $(T_{\max} - T_{\min})^{1/2}$ : Modelo de H-S

A Fig. (1) mostra a correlação entre a transmissividade atmosférica da irradiação solar global ( $H/H_0$ ) e a amplitude térmica  $(T_{\max} - T_{\min})^{1/2}$  conforme o modelo H-S para a base de dados medida no período de 1996-2011 em Botucatu, SP. A Fig. 1 mostra ainda a reta de correlação obtida por regressão entre a  $H/H_0$  e a  $(T_{\max} - T_{\min})^{1/2}$ . O espalhamento na correlação é elevado, onde para cada valor de  $(T_{\max} - T_{\min})^{1/2}$  há um largo intervalo de variação nos valores de ( $H/H_0$ ). O efeito é devido a grande variabilidade do saldo de irradiação diário local, os quais dependem das variações das concentrações de nuvens (quando combina o tipo, número e espessura da camada), vapor d'água e aerossóis na

atmosfera, que absorvem e espalham a radiação global, interferindo no balanço de radiação de ondas curtas, diferentemente nos diversos tipos de cobertura de céu para os mesmos valores de  $(T_{\max} - T_{\min})^{1/2}$ .

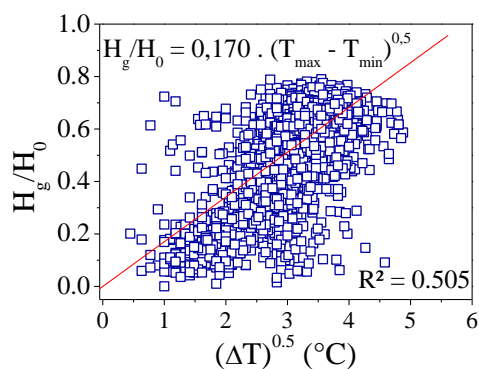


Figura 1 - Relação entre os valores da transmissividade atmosférica da radiação solar global e a amplitude térmica para o período de 1996 - 2011, e a curva de regressão.

A Eq.(4) obtida apresenta o coeficiente de determinação  $R^2 = 0,505$  e o valor do coeficiente de calibração de  $a = 0,170$ :

$$\frac{H_g}{H_0} = 0,170 \cdot (T_{\max} - T_{\min})^{0,5} \quad (4)$$

O valor do coeficiente de calibração  $a = 0,170$  encontrado para Botucatu está dentro do intervalo de variação dos valores divulgados na literatura, e o valor do coeficiente de determinação mostrou  $R^2$  abaixo aos obtidos para as localidades, como mostrado na Tab. 2.

Tabela 2 - Valores do coeficiente de calibração e o coeficiente de determinação de equação de H-S para diferentes localidades do Brasil.

Autores	Localidades	Latitude (°)	a	R <sup>2</sup>
Santos <i>et al.</i> (2014)	3 (Alagoas)	-9,2 a -9,7	0,168 - 0,187	0,704 - 0,790
Borges <i>et al.</i> (2010)	Cruz das Almas (Bahia)	-12,7	0,16 - 0,19 (predeterminado)	0,688
Jerszurki <i>et al.</i> (2013)	Telêmaco Borba (Paraná)	-24,2	0,11 - 0,12	0,740
Silva <i>et al.</i> (2012)	Minas Gerais	-18,8° a -20,8	0,161	0,643

Para a análise sazonal da Eq. (4) a base de dados total foi separada nas quatro estações, e o coeficiente de calibração  $a$  foi calculado para cada estação seguindo o mesmo procedimento de obtenção da Eq. (4). A Fig. 2 apresenta a regressão entre a transmissividade atmosférica e a raiz quadrada da amplitude térmica para cada estação. As equações (5, 6, 7 e 8) mostram as equações sazonais, com seus respectivos valores de  $a$ .

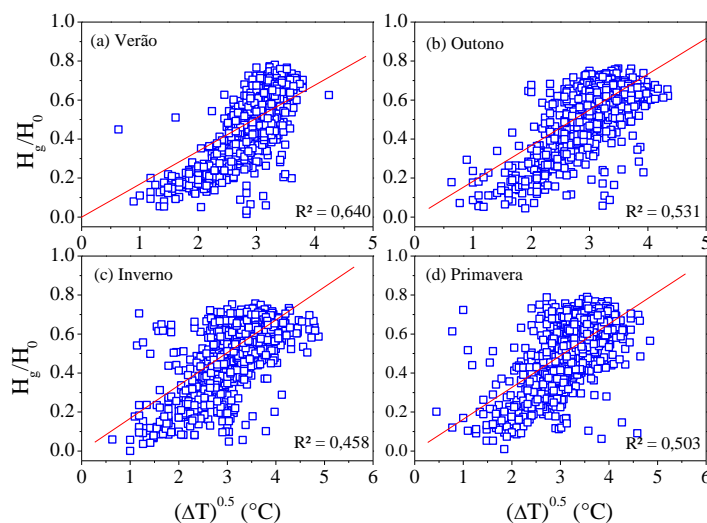


Figura 2 - Regressão linear entre os valores da transmissividade atmosférica e a raiz quadrada da amplitude térmica, sazonais: (a) Verão, (b) Outono, (c) Inverno e (d) Primavera.

Como mostrado das Eq. (5 - 8), a equação da estação verão foi a que apresentou maior coeficiente de correlação,  $R^2 = 0,640$ , e a equação no inverno o menor coeficiente de correlação,  $R^2 = 0,458$ .

$$\text{Verão} \quad \frac{H_G}{H_0} = 0,169 \cdot (T_{\max} - T_{\min})^{0,5} \quad (5)$$

$$\text{Outono} \quad \frac{H_G}{H_0} = 0,183 \cdot (T_{\max} - T_{\min})^{0,5} \quad (6)$$

$$\text{Inverno} \quad \frac{H_G}{H_0} = 0,168 \cdot (T_{\max} - T_{\min})^{0,5} \quad (7)$$

$$\text{Primavera} \quad \frac{H_G}{H_0} = 0,163 \cdot (T_{\max} - T_{\min})^{0,5} \quad (8)$$

Numericamente, a sazonalidade da Eq. (4) foi determinada por meio do cálculo do desvio relativo  $D(\%)$  por meio da Eq. (9) entre os coeficientes angulares da equação anual ( $a_a$ ) e das equações sazonais no verão, outono, inverno e primavera ( $a_s$ ) como mostrado na Fig. (3):

$$D(\%) = 100 \cdot (a_a - a_s)/a_a \quad (9)$$

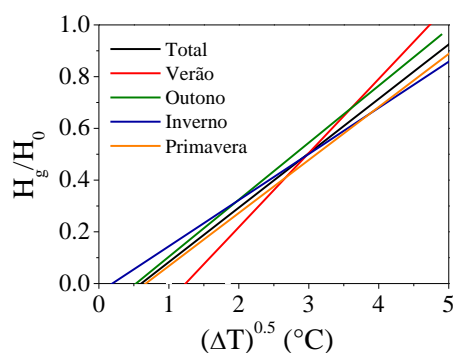


Figura 3 - Retas de regressão linear entre a transmissividade atmosférica e a raiz quadrada da amplitude térmica, anual e sazonal.

O desvio percentual  $D(\%)$  expressa à capacidade de  $H_G$  anual de superestimar ou subestimar  $H_G$  sazonal.

Tabela 3 - Desvio linear entre o valor do coeficiente de calibração da equação anual e os valores dos coeficientes de calibração sazonais.

Estações do ano	Coefficiente $a$	Desvio (%)
<b>Verão</b>	0,169	0,588
<b>Outono</b>	0,183	-7,647
<b>Inverno</b>	0,168	1,176
<b>Primavera</b>	0,163	4,117

Os resultados apresentados na Tab. 3 mostra que para a estimativa de  $H_G$ , o valor do desvio  $D(\%)$  mostra que o valor  $H_G$ , anual superestima aos valores das  $H_G$ , sazonal no verão e inverno e primavera em 0,6%, 1,2% e 1,1% respectivamente, e subestima no outono em 7,7% respectivamente. Os desvios elevados no outono são devidos, a precipitação do tipo frontal originada do encontro das massas frias e secas vindas da região sul com as massas quentes e úmidas, da região sudeste. A distribuição da chuva frontal gera aumento na nebulosidade com ocorrência de chuvas generalizada na região. A intensidade da precipitação é fraca a moderada, dependendo do tipo de frente frio, e a duração é de média a longa (horas ou dias), dependendo da velocidade de deslocamento da frente (Satyamurty *et al.*, 1998).

### 3.2 Validação do Modelo de H-S

A Fig. 4 (a, b) mostra a correlação entre os valores medidos e estimados da radiação solar global para os anos típico (a) e atípico (b). Os coeficientes de correlação encontrados foram  $r^t = 0,823$  para o ano típico e  $r^a = 0,831$  para o ano atípico, sendo o coeficiente de determinação  $(r^2)^t = 0,677$  e  $(r^2)^a = 0,690$ , para os anos típico e atípico, respectivamente. Os indicativos estatísticos obtidos foram:  $rMBE^t = -3,1\%$ ,  $rRMSE^t = 21,9\%$  e  $(d)$  de Wilmott  $d^t = 0,87$  para o ano típico, e  $rMBE^a = 3,2\%$ ,  $rRMSE^a = 24,7\%$  e  $(d)$  de Wilmott  $d^a = 0,88$  para o ano atípico. Os resultados obtidos por meio destes indicativos estatísticos são comparados com os encontrados por outros autores (Tab. 4).

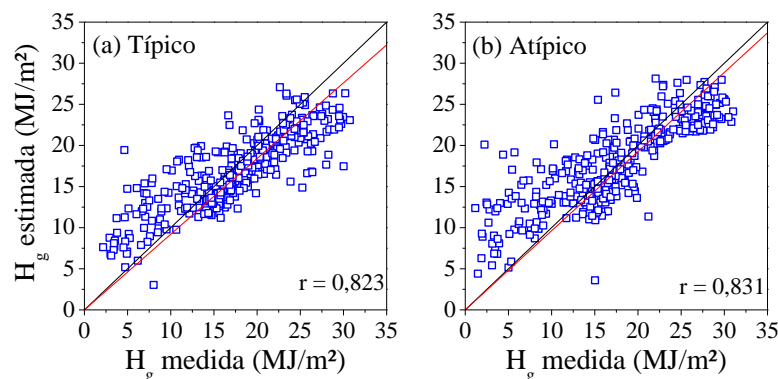


Figura 4 – Comparação entre os valores medidos e estimados para a radiação solar global através do modelo estatístico de Hargreaves-Samani (H-S) para o ano típico (a) e o atípico (b)

Os valores de rMBE, rRMSE e  $d$  de Wilmott, obtidos na validação para os anos típico e atípico são próximos, e mostram que as estimativas de  $H_G$  por meio da Eq. anual (4) pode ser feita com erro inferior ao erro do piranômetro Eppley usado nas medidas da radiação global, que é de 4,0% (Reda *et al.*, 2008), com espalhamento inferior a 25% e concordância de ajuste superior a 87,0%. Os resultados são da mesma ordem de grandeza de validações com outros modelos estatísticos para outras radiações como a difusa (Oliveira *et al.*, 2002; Soares *et al.*, 2004; Dal Pai *et al.*, 2014).

Tabela 4 – Indicativos estatísticos obtidos para diversas localidades.

Autores	Localidades	coeficientes de calibração (a)	RMSE	rRMSE (%)	r
			(MJ m <sup>-2</sup> d <sup>-1</sup> )		
Mehdizadeh <i>et al.</i> (2016)	Kerman (Irã)	0,15	3,23		
Jimenez <i>et al.</i> (2016)	Tucumán (Argentina)		3,69		0,87
Rahimikhoob (2010)	Ahwaz (Irã)	0,16	3,64		
Landeras <i>et al.</i> (2012)	4 (País Basco)		3,94		
Sanusi <i>et al.</i> (2013)	Sokoto (Nigéria)	0,11	0,12		
Sharifi <i>et al.</i> (2016)	2 (Irã)	0,16	3,58 e 4,20	0,004 - 0,018	
//	4 (EUA)	0,14 - 0,15	3,25 - 3,92	0,002 - 0,004	
Grillone <i>et al.</i> (2012)	Sicília (Itália)	0,15 - 0,21	2,58 - 3,79		
Estévez <i>et al.</i> (2012)	65 (Espanha)	0,16 - 19	3,33 - 3,31	18,60 - 19,80	
Daut <i>et al.</i> (2012)	2 (Malásia)			0,17 - 0,21	
Liu <i>et al.</i> (2009)	15 (China)	0,14 - 0,17			
Almorox <i>et al.</i> (2013)	Córdoba (Argentina)	0,15	2,07		
Allen (1997)	9 (EUA)	0,15 - 0,20			
Almorox <i>et al.</i> (2011)	Madrid (Espanha)	0,15 - 0,16	2,79 - 3,06		
Gois <i>et al.</i> (2016)	Seropédica (Brasil)	0,16 - 0,19 (predeterminado)	14,60 - 13,93		
Lêdo <i>et al.</i> (2012)	Barbalha	0,16 (predeterminado)			0,56
Nogueira <i>et al.</i> (2012)	Jaguaruana	0,16 (predeterminado)			0,59
Almorox (2009)	Aranjuez (Espanha)	0,14	3,00		
Ugwu & Ugwuanyi (2011)	Abuja (Nigéria)		0,41 - 2,11		
<b>Presente estudo<sup>t</sup></b>	<b>Botucatu</b>	<b>0,17</b>	<b>3,81</b>	<b>21,87</b>	<b>0,82</b>
<b>Presente estudo<sup>a</sup></b>	<b>Botucatu</b>	<b>0,17</b>	<b>4,23</b>	<b>24,69</b>	<b>0,83</b>

t – ano típico; a – ano atípico.

### 3.3 Treinamento e validação dos modelos gerados por AM

A mesma base de dados da geração do modelo de H-S (Fig.1) foi usada no treinamento das técnicas SVM e ANN. O modelo de H-S e os modelos de SVM1 e ANN1 (combinação1) utilizaram as mesmas variáveis de entrada  $H_0$ ,  $H_G$  e  $(\Delta T)^{0,5}$ . Além da combinação1, foram treinadas duas novas combinações (ANN2 e SVM2 - ANN3 e SVM3), conservando a mesmas variáveis da combinação1, e acrescentando na sequência as seguintes variáveis de entrada: Umidade Relativa (UR, %) e razão de insolação (RI), como mostra a Tab. 5.

Tabela 5 - Parâmetros de entrada da RNA e da SVM para cada combinação,

sendo  $H_o$  – radiação solar extraterrestre;  $T_{\max}$  – Temperatura máxima;  
 $T_{\min}$  – temperatura mínima,  $H_G$  – Radiação solar global; RI – razão de insolação;  
 UR – umidade relativa do ar.

Combinações	Parâmetros de entrada
ANN1	$H_o, (\Delta T)^{0.5}$ e $H_G$
SVM1	$H_o, (\Delta T)^{0.5}$ e $H_G$
ANN2	$H_o, (\Delta T)^{0.5}$ , RI e $H_G$
SVM2	$H_o, (\Delta T)^{0.5}$ , RI e $H_G$
ANN3	$H_o, (\Delta T)^{0.5}$ , RI, UR e $H_G$
SVM3	$H_o, (\Delta T)^{0.5}$ , RI, UR e $H_G$

A Fig. 5 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. Os valores estimados por ANN e SVM nas combinações 2 e 3 estão em concordância linear com as medidas Fig. 5(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.

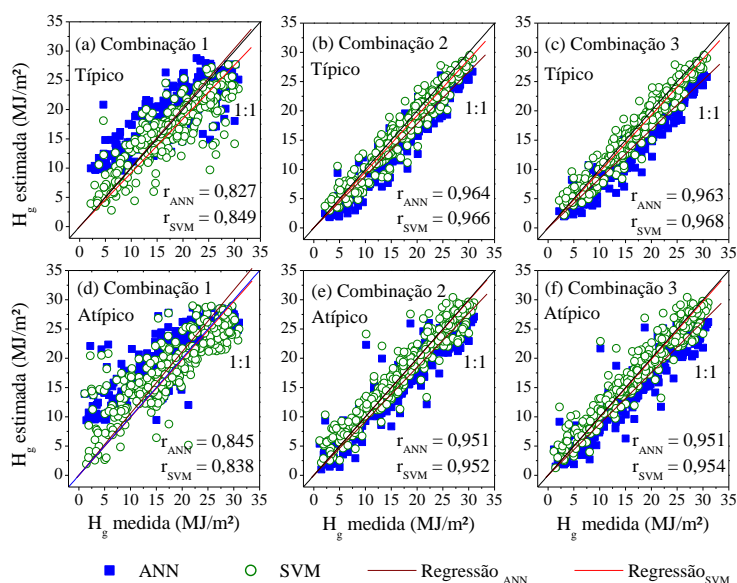


Figura 5 - Gráfico de dispersão entre os valores medidos e estimados para a radiação solar global por meio de AM, para as três combinações de parâmetro de entrada, para os anos típico (a, b e c), e atípico (d, e, f).

Para a combinação1 (ANN1 e SVM1), os valores de  $r^t = 0,849$  e  $r^a = 0,838$ ;  $r^t = 0,827$  e  $r^a = 0,845$ , respectivamente, mostram que a medida e a estimativa de  $H_G$  estão correlacionadas e superiores ao modelo estatístico. A técnica SVM1 apresentou maior valor de  $r$  que a rede ANN1, nas duas condições de validação. O indicativo estatístico RMSE, obtido da comparação dos valores estimados com as técnicas de AM (ANN1 e SVM1) e as medidas de  $H_G$  respectivamente para os anos típico e atípico foram:  $rRMSE^t_{SVM1} = 3,0\%$  e  $rRMSE^a_{SVM1} = 7,2\%$ , enquanto para  $rRMSE^t_{ANN1} = 4,7\%$  e  $rRMSE^a_{ANN1} = 8,8\%$ . A comparação dos indicativos estatísticos  $r$ , e  $rRMSE$  obtidos com as técnicas de SVM1 e ANN1 com a do modelo de H-S com as mesmas variáveis de entrada mostra que a combinação SVM1 e ANN1 tem desempenho superior ao modelo de H-S nas duas condições de validações.

O desempenho das técnicas aumentou com as entradas das outras duas variáveis meteorológicas, e para fins de comparação com o modelo estatístico foi selecionado a combinação3 (ANN3 e SVM3). Os indicativos estatísticos foram validados para a combinação 3 com os anos típico e atípico obtendo os seguintes resultados para  $r$ ,  $rMBE$ ,  $rRMSE$  e  $d$  de Willmott: ( $r^t_{SVM3} = 0,97$ ,  $r^a_{SVM3} = 0,95$ ) e ( $r^t_{ANN3} = 0,96$ ,  $r^a_{ANN3} = 0,95$ ); ( $MBE^t_{SVM3} = -3,3\%$  e  $MBE^a_{SVM3} = 1,0\%$ ) e ( $MBE^t_{ANN3} = -16,2\%$  e  $MBE^a_{ANN3} = -11,6\%$ ); ( $RMSE^t_{SVM3} = 9,9\%$  e  $RMSE^a_{SVM3} = 12,9\%$ ), e ( $RMSE^t_{ANN3} = 19,2\%$  e  $RMSE^a_{ANN3} = 17,9\%$ ) e ( $d$ ) de Willmott ( $d^t_{SVM3} = 0,98$  e  $d^a_{SVM3} = 0,97$ ) e ( $d^t_{ANN3} = 0,93$  e  $d^a_{ANN3} = 0,95$ ).

Os resultados mostram que a técnica SVM possui melhor desempenho, ou maior valor de  $r$  e  $d$  de Willmott, e menor valor de  $rMBE$  e  $rRMSE$  comparado aos valores da técnica ANN e aos do modelo estatístico H-S, nas duas condições de validação típico e atípico. A técnica ANN apresentou melhor resultado do que o modelo estatístico H-S.

Fatores como arquitetura das técnicas, treinamento e processo de validação das técnicas podem ter influenciado na variabilidade dos indicativos estatísticos, principalmente no valor de  $rRMSE$ , não só de SVM3 e ANN3 como também nos valores de SVM e ANN obtidos em outros locais. Ainda assim, os valores de  $r$ ,  $rMBE$ ,  $rRMSE$  e  $d$  de Willmott obtidos com as técnicas SVM3 e ANN3 são comparáveis aos resultados de SVM e ANN obtidos em outros locais (Tab. 6).

Tabela 6 - Índices estatísticos obtidos neste trabalho e comparados aos obtidos por outros autores.

Autores	Localidades	Técnica	RMSE (MJ m <sup>-2</sup> dia <sup>-1</sup> )	rRMSE (%)	r
Urraca <i>et al.</i> (2015)	2 (Espanha)	SVM	2,18 – 2,78		
Chen <i>et al.</i> (2013)	3 (China)	SVM	1,99 – 2,38		
Mohammadi <i>et al.</i> (2015)	Bandar Abbas (Irã)	SVM	1,42 – 2,69		
Ramedani <i>et al.</i> (2014)	Tehran (Irã)	SVM	3,30 – 3,40		
Mohammadi <i>et al.</i> (2015)	Isfahan (Irã)	SVM	2,00 – 2,19		
Quej <i>et al.</i> (2017)	Yucatán (México)	SVM	2,14 – 3,32		
Silva <i>et al.</i> (2017)	Botucatu	SVM	1,64 – 2,14	9,40 – 12,50	0,96 – 0,97
<b>Modelo proposto</b>	<b>Botucatu</b>	<b>SVM<sup>d</sup></b>	<b>1,71</b>	<b>9,87</b>	<b>0,97</b>
<b>Modelo proposto</b>	<b>Botucatu</b>	<b>SVM<sup>b</sup></b>	<b>2,21</b>	<b>12,94</b>	<b>0,95</b>
Urraca <i>et al.</i> (2015)	4 (Espanha)	ANN	2,11 – 2,90		
Mohammadi <i>et al.</i> (2015)	Bandar Abbas (Irã)	ANN	1,84 – 2,85		
Ramedani <i>et al.</i> (2014)	Tehran (Irã)	ANN	3,70		
Fortin <i>et al.</i> (2008)	Montreal (Canadá)	ANN	3,83 – 5,45		
Mejdoul <i>et al.</i> (2013)	8 (Marrocos)	ANN	1,23 – 1,50		
Elizondo <i>et al.</i> (1994)	7 (EUA)	ANN	2,92 – 3,64		
Quej <i>et al.</i> (2017)	Yucatán (México)	ANN	2,48 – 3,61		
Silva <i>et al.</i> (2017)	Botucatu	ANN	2,68 – 2,89	15,60 – 16,60	0,95 – 0,96
<b>Presente estudo</b>	<b>Botucatu</b>	<b>ANN<sup>d</sup></b>	<b>3,34</b>	<b>19,18</b>	<b>0,96</b>
<b>Presente estudo</b>	<b>Botucatu</b>	<b>ANN<sup>b</sup></b>	<b>3,07</b>	<b>17,93</b>	<b>0,95</b>

#### 4. CONCLUSÕES:

Do que foi apresentado e discutido no trabalho podemos concluir que: a equação estatística de H-S obtida neste trabalho possui baixa correlação mostrando-se inadequado em estimar  $H_G$ . A análise da sazonalidade efetuada por meio dos valores de desvio linear entre a equação anual e as equações sazonais, não apontam vantagem em se utilizar as equações sazonais no lugar da equação total. As técnicas de aprendizado de máquinas apresentaram melhores desempenhos do que o modelo estatístico, sendo que o melhor desempenho seguiu a ordem das técnicas SVM, ANN, e modelo H-S em estimar  $H_G$ . As técnicas de aprendizado de máquinas SVM, ANN podem estimar  $H_G$  com precisão e exatidão a partir das variáveis meteorológicas radiação solar extraterrestre; Temperatura máxima; temperatura mínima, razão de insolação e umidade relativa do ar.

#### Agradecimentos

Os Autores agradecem o CNPq e à FAPESP pelo apoio financeiro.

#### REFERÊNCIAS

- Allen, R. G., 1997. Self- Calibrated Method For Estimating Solar Radiation From Air Temperature. *Journal of Hydrologic Engineering*, v. 2(2), p. 56 – 67.
- Almorox, J.; Hontoria, C.; Benito, M., 2011. Models for obtaining daily global solar radiation with measured air temperature data in Madrid (Spain). *Applied Energy*, v. 88, p.1703-1709.
- Almorox, J.; Bocco, M.; Willington, E., 2013. Estimation of daily global solar radiation from measured temperatures at Cañada de Luque, Córdoba, Argentina. *Renewable Energy*, v. 60, p. 382 – 387.
- Bechini, L.; Ducco, G.; Donatelli, M.; Stein, A., 2000. Modeling, interpolation and stochastic simulation in space and time of global solar radiation. *Agriculture Ecosystems & Environment*, v. 81, p. 29-42.
- Borges, V. P.; Oliveira, A. S.; Filho, M. A. C.; Silva, T. S. M.; Pamponet, B. M., 2010. Avaliação de modelos de estimativa da radiação solar incidente em Cruz das Almas, Bahia. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.14, n.1, p.74-80.
- Bristow, K. L.; Campbell, G. S., 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agricultural and Forest Meteorology*, v.31, n. 2, p.159-166.
- Citakoglu, H., 2015. Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation. *Computers and Electronics in Agriculture*, v. 118, p. 28-37.
- Chen, J-L; Li, G-S, Wu, S-J., 2013. Assessing the potential of support vector machine for estimating daily solar radiation using sunshine duration. *Energy Conversion and Management*, v.75, p. 311-318.
- Cortes, C.; Vapnik, V., 1995. Support – Vector Networks. *Machine Learning*, v. 20, p. 273-297.
- Dai, Q.; Fang, X., 2014. A simple model to predict solar radiation under clear sky conditions. *Advances in Space Research*, v.53, 1239-1245.
- Dal Pai, A.; Escobedo, J. F.; Dal Pai, J. F. *et al.*, 2014. Estimation of Hourly, Daily and Monthly Mean Diffuse Radiation Based on MEO Shadowring Correction. *Energy Procedia*, v. 57, p. 1150-1159, 2014.



- Daut, I.; Irwanto, M.; Irwan, Y. M.; Gomesh, N.; Ahmad, N. S., 2011. Combination of Hargreaves method and linear regression as a new method to estimate solar radiation in Perlis, Northern Malaysia. *Solar Energy*, v. 85, p. 2871 – 2880.
- Elizondo, D.; Hoogenboom, G.; McClendon, R. W., 1994. Development of a neural network model to predict daily solar radiation. *Agricultural and Forest Meteorology*, v. 71, p. 115-132.
- Estévez, J.; Padilla, F. L. M.; Gavilán, P., 2012. Evaluation and Regional Calibration of Solar Radiation Prediction Models in Southern Spain. *Journal of Irrigation and Drainage Engineering*, v. 138(10), p. 868 – 879.
- Fortin, J.; Anctil, F.; Parent, L. E.; Bolinder, M. A., 2008. Comparison of empirical daily surface incoming solar radiation models. *Agricultural and Forest Meteorology*, v. 148, p. 1332 – 1340.
- Gois, G.; Souza, J. C. S.; Junior, J. F. O.; Delgado, R. C.; Lyra, G. B.; Teodoro, P. E., 2016. Models to estimate incidente solar radiation on Seropédica, Rio de Janeiro. *Bioscience Journal*, v. 32, n. 2, p. 505 – 513.
- Grillone, G.; Agnese, C.; D’Asaro, F., 2012. Estimation of Daily Solar Radiation from Measured Air Temperature Extremes in the Mid-Mediterranean Area. *Journal of Irrigation and Drainage Engineering*, v. 138(10), p. 939 – 947.
- Gueymard, C. A., 2003. Direct solar transmittance and irradiance predictions with broadband models. Part I: detailed theoretical performance assessment. *Solar Energy*, v. 74, p. 355-379.
- Hargreaves, G. H.; Samani, Z. A., 1982. Estimating potential evapotranspiration. *Journal of Irrigation and Drainage Engineering*, v. 108, p. 225-230.
- Hunt, L.; Kuchar, L.; Swanton, C. J., 1998. Estimation of solar radiation for use crop modeling. *Agricultural Forest Meteorology*, v. 91, p. 293-300.
- Jerszurki, D.; Souza, J. L. M., 2013. Parametrização das equações de Hargreaves & Samani e Angström-Prezcott para estimativa da radiação solar na região de Telêmaco Borba, Estado do Paraná. *Ciência Rural*, v.43, n.3, p.383-389.
- Jimenez, V. A.; Barrionuevo, A.; Will, A.; Rodríguez, S., 2016. Neural Network for Estimating Daily Global Solar Radiation Using Temperature, Humidity and Pressure as Unique Climatic Input Variables. *Smart Grid and Renewable Energy*, v. 7, p. 95 – 104.
- Landeras, G.; López, J. J.; Kisi, O.; Shiri, J., 2012. Comparison for Gene Expression Programming with neuro-fuzzy and neural network computing techniques in estimating daily incoming solar radiation in the Basque Country (Northern Spain). *Energy Conversion and Management*, v. 62, p. 1 – 13.
- Ledo, E. R. F.; Silva, M. G.; Nogueira, D. H.; Arraes, F. D. D., 2012. Avaliação de modelo de estimativa da radiação solar global (Rs) com base na amplitude térmica. *Conexões – Ciência e Tecnologia*, v. 6, n. 1, p. 15 – 26.
- Lima, F. J. L.; Martins, F. R.; Pereira, E. B. *et al.*, 2016. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. *Renewable Energy*, v. 87, p.807-818.
- Madkour, M. A.; El-metwally, M.; Harmed, A. B., 2006. Comparative study on different models for estimation of direct normal irradiance (DNI) over Egypt atmosphere. *Renewable Energy*, v. 31, p. 361-382.
- Mehdizadeh, S.; Behmanesh, J.; Khalili, K., 2016. Comparison of artificial intelligence methods and empirical equations to estimate daily solar radiation. *Journal of Atmospheric and Solar-Terrestrial Physics*, v. 146, p. 215 – 227.
- Mejdoul, R.; Taqi, M.; Blouaggadia, N., 2013. Artificial neural network based prediction modelo f daily global solar radiation in Morocco. *Journal of Renewable and Sustainable Energy*, v. 5.
- Mohammadi, K.; Shamshirband, S.; Anisi, M. H.; Alam, K. A.; Petkovic, D., 2015. Support vector regression based prediction of global solar radiation on a horizontal surface. *Energy Conversion and Management*, v.91, p.433-441.
- Mohammadi, K.; Shamshirband, S.; Tong, C. W.; Arif, M.; Petkovic, D.; Ch, S, 2015. A new hybrid support vector machine-wavelet transform approach for estimation of horizontal global solar radiation. *Energy Conversion and Management*, v. 92, p. 162 – 171.
- Nogueira, D. H.; Silva, M. G.; Ledo, E. R. F.; Araújo, E. M., 2012. Estimativa da radiação solar global (RS) através da amplitude térmica diária. IV Workshop Internacional de Inovação Tecnológica da Irrigação, Fortaleza – CE, Brasil.
- Oliveira, A. P.; Escobedo, J. F.; Machado, A. J.; Soares, J., 2002. Correlation model of diffuse solar-radiation applied to the city of São Paulo, Brazil. *Applied Energy*, v. 71, p. 59-73.
- Oliveira, L. L. A., 2006. Estimation of software project effort with Support Vector Regression. *Neurocomputing*, v. 69, p. 1749-1753.
- Piri, J.; Kisi, O., 2015. Modelling solar radiation reached to the Earth using ANFIS, NN-ARX, and empirical models (Case studies: Zahedan and Bojnurd stations). *Journal of Atmospheric and Solar-Terrestrial Physics*, v.123, p. 39 – 47.
- Piri, J.; Shamshirband, S.; Petkovic, D.; Tong, C. W.; Rehman, M. H., 2015. Predictio of the solar radiation on the Earth using support vector regression technique. *Infrared Physics and Technology*, v. 68, p. 179-185.
- Quej, V. H.; Almorox, J.; Arnaldo, J. A.; Saito, L., 2017. ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. *Journal of atmospheric and solar-terrestrial physics*, v.155, p.62-70.
- Rahimikhoob, A., 2010. Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment. *Renewable Energy*, v. 35, p. 2131 – 2135.
- Ramedani, Z.; Omid, M.; Keyhani, A., 2012. A method on neural networks for generating solar radiation map. *International Journal of Energy and Environment*, v. 3, p. 775-786.

- Ramedani, Z.; Omid, M.; Keyhani, A.; Shamshirband, S.; Khoshnevisan, B., 2014. Potential of radial basis function based support vector regression for global solar radiation prediction. *Renewable and Sustainable Energy Reviews*, v. 39, p. 1005 – 1011.
- Reda, I.M.; Myers, D.R.; Stoffel, T.L., 2008. Uncertainty estimate for the outdoor calibration of solar pyranometers: NCSLI Measure. *The Journal of Measurement Science*, v.3, n.4, p.58-66.
- Rehman, S.; Mohandes, M., 2008. Artificial neural network estimation of global solar radiation using air temperature and relative humidity. *Energy Policy*, v. 36, p. 571-576.
- Santos, C. M.; Souza, J. L.; Junior, R. A. F. *et al.*, 2014. M. On modeling global solar irradiation using air temperature for Alagoas State, Northeastern Brazil. *Energy*, v.71, p. 388-398.
- Santos, C. M.; Escobedo, J. F.; Teramoto, É. T.; Modenese, S. H., 2016. Assessment of ANN and SVM models for estimating normal direct irradiation. *Energy Conversion and Management*, v.126, p.826-836.
- Sanusi, Y. K.; Abisoye, S. G.; Abiodun, A. O., 2013. Application of Artificial Neural Networks to Predict Daily Solar Radiation in Sokoto. *International Journal of Current Engineering and Technology*, v. 3, n. 2, p. 647 – 652.
- Satyamurty, P.; Nobre, C. A.; Silva Dias, P. L. *Topics: South America. Meteorological Monographs*, v.27, n.49, p.119-139, 1998.
- Sayed Saber Sharifia\*, Vahid Rezaverdinejada, Vahid Nouranib Sharifi, S. S.; Resaverdineja, V.; Nourani, V., 2016. Estimation of daily global solar radiation using wavelet regression, ANN, GEP and empirical models: A comparative study of selected temperature based approaches. *Journal of Atmospheric and Solar-Terrestrial Physics*, v. 149, p. 131 – 145.
- Silva, C. R.; Silva, V. J.; Júnior, J. A.; Carvalho, H. P., 2012. Radiação solar estimada com base na temperatura do ar para três regiões de Minas Gerais. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.16, p. 281-288.
- Silva, M. B. P.; Escobedo, J. F.; Rossi, T. J.; *et al.*, 2017. Performance of the Angstrom-Prescott model and SVM and ANN techniques to estimate daily global solar radiation in Botucatu/SP/Brazil. *Journal of Atmospheric and Solar-Terrestrial Physics*, v.160, p. 11-23.
- Souza, J. L.; Lyra, G. B.; Santos, C. M. *et al.*, 2016. Empirical models of daily and monthly global solar irradiation using sunshine duration for Alagoas State, Northeastern Brazil. *Sustainable Energy Technologies and Assessments*, v.14, p. 35-45.
- Tymvios, F. S.; Jacovides, C. P.; Michaelides, S. C.; Scouteli, C., 2005. Comparative study of Angstrom's and artificial neural networks' methodologies in estimating global solar radiation. *Solar Energy*, v. 78, p. 752-762.
- Urraca, R.; Antonanzas, J.; Pisón, F. J. M.; Torres, F. A., 2015. Estimation of solar global irradiation in remote areas. *Journal of Renewable and Sustainable Energy*, v. 7.
- Ugwu, A. I.; Ugwuanyi, J. U., 2011. Performance assessment of Hargreaves model in estimating solar radiation in Abuja using minimum climatological data. *Journal of the Physical Sciences*, v. 6(31), p. 7285 – 7290.
- Yacef, R.; Mellit, A.; Belaid, Sen, Z., 2014. New combined models for estimating daily global solar radiation from measured air temperature in semi-arid climates: Application in Ghardaia, Algeria. *Energy Conversion and Management*, v. 79, p. 606 – 615.
- Zhang, F.; Zhou, X.; Zhang, H.; Peng, X.; Wang, Z., 2014. On the relationship between direct and anisotropic diffuse radiation. *Infrared Physics & Technology*, v.65, p.5-8.

#### **ESTIMATION OF GLOBAL SOLAR IRRADIATION BY HARGREAVES-SAMANI MODELS AND LEARNING OF SVM AND ANN MACHINE IN BOTUCATU/SP/BRAZIL**

**Abstract.** *In this paper the comparative study between two methods of global solar radiation estimation Hargreaves-Samani statistical model (HS) and the machine learning models Artificial Neural Network (ANN) and Support Vector Machine (SVM), for three combinations of inputs of meteorological variables. The values of the  $r$ ,  $rMBE$ ,  $rRMSE$  and Wilmott statistical indicative values obtained in the validation of the models were conclusive: the SVM3 technique showed better performance than the ANN3 network and the statistical model (HS) in estimating  $H_G$ , ANN3 technique has superior performance to the statistical model (HS) in estimating  $H_G$ .*

**Key words:** *Solar radiation, Hargreaves-Samani model, statistical modeling; artificial intelligence, meteorological variables.*