



## Estimativa da radiação global diária a partir de variáveis meteorológicas comuns em ambiente subtropical úmido

Érico Tadao TERAMOTO<sup>1\*</sup>, Darllan Collins da SILVA e CUNHA<sup>1</sup>, Cícero Manoel dos SANTOS<sup>2</sup>,  
Levi Pompermayer MACHADO<sup>1</sup>, João Francisco ESCOBEDO<sup>3</sup>

<sup>1</sup>Campus Experimental de Registro, Universidade Estadual Paulista, Registro, SP, Brasil.

<sup>2</sup>Universidade Federal do Pará, Altamira, PA, Brasil.

<sup>3</sup>Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista, Botucatu, SP, Brasil.

\*E-mail: [erico.teramoto@unesp.br](mailto:erico.teramoto@unesp.br)

Recebido em fevereiro/2019; Aceito em junho/2019.

**RESUMO:** O conhecimento do potencial solar é considerado uma informação essencial no dimensionamento e simulação do desempenho de sistemas que convertem radiação solar em energia elétrica ou calor disponível em fluidos, bem como em estudos agrícolas voltados ao manejo de irrigação e avaliação da produtividade potencial. Nem sempre as informações sobre a radiação solar incidente estão disponíveis e, por isso, torna-se importante o desenvolvimento de métodos capazes de estimar com acurácia a radiação solar. O presente estudo avaliou o desempenho de diferentes métodos de estimativa da radiação solar global ( $R_G$ ) em ambiente subtropical úmido brasileiro, a fim de determinar qual deles apresenta maior acurácia. As bases de dados meteorológicas utilizadas foram registradas pela estação meteorológica automática pertencente ao Instituto Nacional de Meteorologia (INMET) e instalada em Iguape, município localizado na porção paulista do Vale do Ribeira. Modelos empíricos foram calibrados e tomados como referência na avaliação do desempenho da Máquina de Vetores de Suporte (SVM) utilizando diferentes conjuntos de variáveis de entrada. Dentre os métodos testados, a SVM utilizando o conjunto de variáveis de entrada  $\{R_0, T_{max}, T_{min}, e_{s,max}, e_{s,min}\}$  apresentou o melhor desempenho.

**Palavras-chave:** modelos empíricos; Vale do Ribeira; SVM.

## Estimating daily global radiation from common meteorological variables in a humid subtropical environment

**ABSTRACT:** Knowledge of solar potential is essential in the design and simulation of systems that convert solar radiation in electric power or available heat in fluids, as well as in agricultural studies focused on irrigation management and potential productivity assessment. Oftentimes, information about solar radiation are not available and, therefore, it becomes important the development of methods for obtain reliable data of solar radiation. The present study evaluated the performance of different methods of estimating global solar radiation (RG) in a Brazilian humid subtropical environment in order to determine which is more accurate. The meteorological databases used were registered by the automatic meteorological station belonging to the National Institute of Meteorology (INMET) and installed in Iguape, a municipality located in the São Paulo portion in the Ribeira Valley. Empirical models were calibrated and taken as reference in the performance evaluation of the Support Vector Machine (SVM) using different sets of input variables. Among the tested methods, the SVM using the set of input variables  $\{R_0, T_{max}, T_{min}, e_{s,max}, e_{s,min}\}$  presented the best performance.

**Keywords:** empirical models; Ribeira Valley; SVM.

### 1. INTRODUÇÃO

Levando em consideração o uso de fontes renováveis de energia para aumentar a sustentabilidade de sistemas de produção agrícolas e aquícolas (VALENTI et al., 2018), diante também da preocupação em relação às mudanças climáticas, a energia solar se destaca por ser limpa (sem a emissão de gases do efeito estufa durante sua utilização), além de ter ampla disponibilidade em locais de clima tropical, como é o caso de grande parte do território brasileiro (ALVARES et al., 2014).

Para realizar o aproveitamento da energia solar, o levantamento do potencial solar local é considerado um dos primeiros passos, que é feito a partir de séries temporais de dados de radiação solar global ( $R_G$ , espectro 300-2800 nm). Com base nas informações sobre o potencial solar local é

possível dimensionar e simular o desempenho de sistemas que convertem a radiação solar em energia elétrica ou calor disponível em fluidos como água ou ar. Séries temporais de  $R_G$  também são consideradas úteis em estudos agrícolas voltados ao manejo de irrigação e avaliação da produtividade potencial, pois a radiação solar é um dos principais fatores condicionantes das temperaturas do ar e do solo, fotossíntese e da evapotranspiração. No caso da evapotranspiração de referência, o método de estimativa de Penman-Monteith, que é amplamente utilizado por ser recomendado pela FAO (Organização das Nações Unidas para a Alimentação e a Agricultura) e pela Organização Mundial de Meteorologia (WMO), tem a acurácia das estimativas muito dependente da qualidade dos dados de  $R_G$  (CITAKOGLU et al., 2014). A

qualidade dos dados de  $R_G$  também é importante em avaliações dos recursos solares, onde pequenas diferenças são decisivas para considerar um projeto viável economicamente ou não (MENDELSON et al., 2012).

Dados de  $R_G$  podem ser obtidos através do monitoramento por estações meteorológicas terrestres ou por sensoriamento remoto. Porém, quando o monitoramento não é realizado ou para completar séries temporais, dados de  $R_G$  podem ser obtidos através de estimativa utilizando modelos empíricos ou Técnicas de Aprendizado de Máquinas (TAM) (MEHDIZADEH et al., 2016; MEENAL; SELVAKUMAR, 2018). Os modelos empíricos são obtidos por meio da correlação entre a  $R_G$  e outras variáveis medidas rotineiramente em estações meteorológicas, e podem ser classificados em quatro categorias baseadas no (BESHARAT et al., 2013): (i) brilho solar, (ii) nebulosidade, (iii) temperatura do ar e (iv) outros parâmetros meteorológicos. Em geral, os modelos empíricos baseados no brilho solar e os baseados na temperatura do ar são considerados os mais e menos precisos na estimativa da  $R_G$ , respectivamente (BESHARAT et al., 2013; QUEJ et al., 2016; MEENAL; SELVAKUMAR, 2018). Contudo, a temperatura do ar é considerada uma variável meteorológica cujos dados são de fácil obtenção e, por isso, os modelos empíricos da categoria (iii) são normalmente os mais utilizados (SELVAKUMAR, 2018).

Paralelas aos modelos empíricos, as TAM foram foco de estudos nos últimos vinte anos e tornaram-se métodos alternativos importantes, a partir da popularização dos microcomputadores, disponibilidade de compilações gratuitas e de fácil utilização e devido à boa capacidade de generalização mesmo na presença de dados inconsistentes (MOHAMMADI et al., 2015). Atualmente, as redes neurais artificiais (RNA) e a Máquina de Vetores de Suporte (SVM) são as TAM mais conhecidas e utilizadas (KUMAR et al., 2015; MEENAL; SELVAKUMAR, 2018). E dentre as duas, a SVM é considerada o método de maior acurácia na estimativa da  $R_G$ , sendo capaz de gerar estimativas com maior acurácia utilizando um número menor de variáveis de entrada (QUEJ et al., 2017; SILVA et al., 2017; SELVAKUMAR, 2018).

Contudo, é importante ressaltar que o desempenho da SVM na estimativa da  $R_G$  é dependente das variáveis de entrada utilizadas, que por sua vez são selecionadas de acordo com o clima do local de estudo (BELAID; MELLIT, 2016; BAKHASHWAIN, 2016; QUEJ et al., 2017; SELVAKUMAR, 2018). Grande parte dos estudos avaliando o uso da SVM na estimativa da  $R_G$  foram realizados em ambiente árido ou semiárido, poucos em ambiente subtropical úmido e tendo foco no uso de variáveis meteorológicas consideradas comuns (temperatura do ar, precipitação e umidade relativa do ar).

Com base nas informações acima, no presente estudo a acurácia de diferentes métodos de estimativa da  $R_G$  foi avaliada em ambiente subtropical úmido brasileiro. O objetivo específico do estudo foi identificar, dentre todos os métodos testados, qual é o mais simples e preciso para estimar a  $R_G$  diária nesse tipo de clima que é predominante na região do Vale do Ribeira. Seis modelos empíricos foram calibrados utilizando séries temporais de variáveis meteorológicas monitoradas em Iguape, São Paulo, Brasil. Tendo os modelos empíricos como referência, foram testados oito conjuntos de variáveis de entrada na SVM, todos compostos por variáveis meteorológicas consideradas comuns.

## 2. MATERIAL E MÉTODOS

### 2.1. Local de estudo e obtenção dos dados

O Vale do Ribeira, que abrange 31 municípios distribuídos entre os Estados de São Paulo e Paraná, é uma região importante do ponto de vista ambiental, devido à sua rica biodiversidade presente nos ecossistemas aquáticos e terrestres (IBGE, 2016). Ele contém a maior área remanescente da Mata Atlântica do mundo, que é dividida em várias Áreas de Preservação Permanente (APP). Em grande parte dessas APP, verifica-se a presença de comunidades tradicionais, como grupos indígenas, quilombolas e caiçaras, muitos dos quais organizados em associações que realizam atividades agropecuárias e de subsistência (COELHO; FAVARETO, 2008).

A economia regional é baseada nas áreas de serviços, agricultura, bubalicultura, aquicultura e pesca, tendo como grande desafio o desenvolvimento sustentável destas atividades, que passa por etapas como a utilização de fontes renováveis de energia. Em relação ao clima, predomina na região o subtropical úmido com verão quente, Cfa pelo critério de classificação de Köppen (ALVARES et al., 2014).

Dentre os municípios que compõe o Vale do Ribeira, Iguape é o maior em área territorial, com um total de 1.978 Km<sup>2</sup> (IBGE, 2016), e um dos poucos que monitoram continuamente a  $R_G$ . Em Iguape, a densidade demográfica é baixa, com uma população de 30,4 mil habitantes, pois aproximadamente 70% do território do município é área natural protegida, que inclui a Estação Ecológica dos Chauás e cerca de 85% da Estação Ecológica Juréia-Itatins, além de estar parcialmente em Área de Proteção Ambiental (APA – Cananéia-Iguape-Peruíbe).

Os dados diários de  $R_G$ , temperatura máxima do ar ( $T_{max}$ ), temperatura mínima do ar ( $T_{min}$ ), umidade relativa do ar (UR) e precipitação pluvial acumulada (P) utilizados no presente estudo foram medidos pela estação meteorológica automática pertencente ao Instituto Nacional de Meteorologia (INMET) e instalada em Iguape, São Paulo, Brasil (Figura 1).

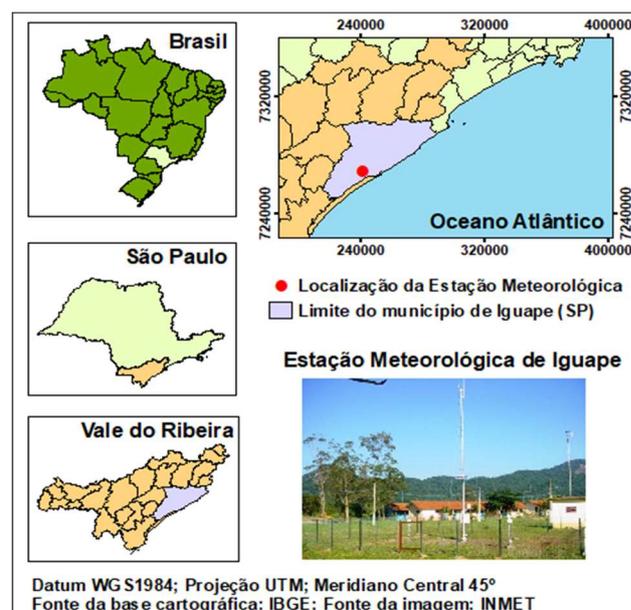


Figura 1. Localização da estação meteorológica do INMET instalada em Iguape/SP/Brazil.

Figure 1. Location of the INMET weather station in Iguape/SP/Brazil.

O monitoramento da  $R_G$  foi realizado utilizando um piranômetro modelo CM21 da Kipp & Zonen e das variáveis  $T_{max}$ ,  $T_{min}$  e UR foi realizada utilizando uma sonda modelo HMP45C da Vaisala HUMICAP. O período de medição foi de janeiro de 2008 a dezembro de 2011.

Os dados de  $R_0$  (expressos em  $MJm^{-2}dia^{-1}$ ) foram estimados a partir da Equação 1 abaixo (IQBAL, 1983):

$$R_0 = \frac{24 \times 3.600}{\pi} I_{CS} E_0 \left( \cos \phi \times \cos \delta \times \sin \omega_s + \frac{\pi}{180} \omega_s \times \sin \phi \times \sin \delta \right) \quad (1)$$

em que:  $I_{CS}$  = a constante solar,  $1.367 Wm^{-2}$ ;  $E_0$  = fator de correção da excentricidade terrestre, adimensional; expressos em graus,  $\phi$  = latitude local,  $\delta$  = declinação solar e  $\omega_s$  = ângulo horário. As variáveis  $E_0$ ,  $\delta$ ,  $\omega_s$  e  $N$  foram estimadas a partir das Equações 2, 3, 4 e 5, respectivamente:

$$E_0 = 1,0 + 0,034 \cos \Gamma + 0,001 \sin \Gamma + 0,0007 \cos 2\Gamma + 0,00008 \sin 2\Gamma \quad (2)$$

$$\delta = 23,45 \sin \left[ \frac{360}{365} (DJ - 80) \right] \quad (3)$$

$$\omega_s = \ar \cos \left[ - \tan(\phi) \tan(\delta) \right] \quad (4)$$

$$N = \frac{2\omega_s}{15} \quad (5)$$

em que: DJ = dia juliano e  $\Gamma$  é calculado a partir da Equação  $\Gamma = 0,986 (DJ - 1)$ .

As bases de dados obtidas foram analisadas em relação a inconsistências, seguindo os critérios recomendados por WMO (2008), sendo eliminados os dados considerados inconsistentes. Em uma etapa seguinte, as bases de dados foram sincronizadas, sendo considerados dias com registros uteis apenas os que apresentaram a medição simultânea das quatro variáveis meteorológicas. Por último, foram adicionados os dados diários de radiação solar no topo da atmosfera ( $R_0$ ), fotoperíodo ( $N$ ) e de índice de transmissividade atmosférica  $K_T$  (adimensional, é razão entre  $R_G$  e  $R_0$ ).

A utilização delas envolve duas etapas: (i) treinamento, que ocorre através da apresentação de exemplos, análogo ao processo de calibração dos modelos empíricos; (ii) estimativa, após o treinamento e aquisição da capacidade de generalização dos exemplos apresentados. Porém, após a realização do treinamento da TAM, não é possível identificar uma equação de estimativa como no caso da geração de modelos empíricos. Por isso, tal processo de modelagem é chamado de “caixa preta” e apresenta como vantagem em relação aos modelos empíricos a possibilidade de utilizar um número grande de variáveis de entrada sem gerar com isso uma equação complexa algebricamente (HALL et al., 2009).

Com o propósito de ajustar os modelos empíricos, treinar a SVM e validar os dois métodos de estimativa, as bases totais de dados foram divididas em duas partes: base de dados de treinamento (75% da base total de dados) e base de dados de validação (25% da base total de dados). Na validação dos métodos de estimativa foram utilizados os parâmetros

estatísticos MBE (Erro Médio Absoluto) e RMSE (Raiz Quadrada do Erro Médio) e “d” de Wilmott. O MBE permite avaliar quanto o modelo subestima ( $MBE < 0$ ) ou superestima ( $MBE > 0$ ) os valores de  $R_G$ , sendo expresso em percentual. A extensão do erro na estimativa foi avaliada a partir do RMSE, expresso em percentual. O “d” de Wilmott, adimensional, varia de 0 a 1 e permite avaliar a acurácia do modelo. Assim, os menores valores de RMSE e o MBE mais próximo de zero e valores de “d” próximos de 1 indicam melhor acurácia do método de estimativa. Os parâmetros estatísticos MBE, RMSE e índice d foram obtidos utilizando as seguintes equações (WILMOTT, 1981; WILMOTT; MATSUURA, 2005):

$$MBE = 100 * \frac{n^{-1} \sum_{i=1}^n (X_{E_i} - X_{M_i})}{\bar{X}} \quad (6)$$

$$RMSE = 100 * \frac{\sqrt{n^{-1} \sum_{i=1}^n (X_{E_i} - X_{M_i})^2}}{\bar{X}} \quad (7)$$

$$d = 1 - \frac{\sum_{i=1}^n (X_{E_i} - X_{M_i})^2}{\sum_{i=1}^n (|X_{E_i}| + |X_{M_i}|)^2} \quad (8)$$

em que:  $X_{E_i}$  = valor estimado;  $X_{M_i}$  = valor medido;  $\bar{X}$  =  $(\sum_{i=1}^n X_{M_i})/n$  é a média dos valores medidos;  $n$  = número de observações.

## 2.2. Modelos empíricos

A seleção dos modelos empíricos a serem calibrados para estimativa da  $R_G$  diária foi realizada em função dos dados de variáveis meteorológicas disponíveis em Iguape e que são medidas como rotina nas estações meteorológicas de superfície. Foram calibrados seis modelos empíricos considerados clássicos, dos quais três baseados em temperatura do ar e os demais em outras variáveis meteorológicas. Os modelos empíricos calibrados aqui serviram como referência na avaliação do desempenho da SVM operando sob as mesmas condições (estimativa utilizando as mesmas variáveis de entrada).

Na Tabela 1, RT é a precipitação pluvial transformada, (se  $P = 0$ ,  $RT = 0$ ; se  $P > 0$ ,  $RT = 1$ );  $T_{med} = (T_{max} - T_{min})/2$ , expresso em  $^{\circ}C$ ;  $e_{s,max}$  e  $e_{s,min}$  são as pressões de saturação de vapor de água nas condições das temperaturas máxima e mínima do ar, respectivamente, expressas em kPa; e “a”, “b”, “c” e “d” são coeficientes calibrados. Na Tabela 2 são apresentados os coeficientes dos modelos empíricos calibrados para a cidade de Iguape.

## 2.3. Máquinas de vetores de suporte

No presente trabalho, a estimativa da  $R_G$  com a SVM foi realizada utilizando o software WEKA (HALL et al., 2009). Baseado nos modelos empíricos os modelos SVM testados utilizaram como variáveis de entrada a  $R_0$ ,  $T_{max}$ ,  $T_{min}$ ,  $P$ , RT, UR,  $N$  e mês do ano. Nesse caso, foram avaliados oito modelos SVM (Tabela 3), que serão simbolizados como SVM(x), onde o x indica o agrupamento de variáveis de entrada utilizado.

Com a utilização dos agrupamentos de variáveis de entrada SVM(1), SVM(2), SVM(3) e SVM(4) foi possível avaliar o desempenho da SVM comparando-a com os modelos empíricos operando sob as mesmas condições. Os modelos

SVM(5), SVM(6), SVM(7) e SVM(8) possuem agrupamento de variáveis de entrada similar aos apresentados pelos modelos SVM(1), SVM (2), SVM (3) e SVM(4), respectivamente, sendo o uso das variáveis N e mês o diferencial apresentado por elas. A adição das variáveis N e mês não implica em

aumento no número de variáveis medidas utilizadas e tem como objetivo permitir que a SVM diferencie os exemplos em função da sazonalidade (MEENAL; SELVAKUMAR, 2018). Devido à eficiência computacional, no presente estudo foi utilizada a função de Kernel de Base Radial (RBF).

Tabela 1. Modelos empíricos calibrados para estimativa da R<sub>G</sub> diária.  
Table 1. Empirical models calibrated for estimating daily R<sub>G</sub>.

Modelo	Ident.	Variáveis de entrada	Equação
Hargreaves e Samani (1982)	HS	R <sub>0</sub> , T <sub>max</sub> e T <sub>min</sub>	$R_G = R_0 a \sqrt{T_{max} - T_{min}}$
Hunt et al. (1998)	HM	R <sub>0</sub> , T <sub>max</sub> e T <sub>min</sub>	$R_G = R_0 a \sqrt{T_{max} - T_{min}} + b$
Almox et al. (2011)	AL	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> , e <sub>s max</sub> e e <sub>s min</sub>	$R_G = R_0 a (T_{max} - T_{min})^b [1 - \exp(-c(e_{s,max}/e_{s,min})^d)]$
DeJong e Stewart (1993)	DS	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> e P	$R_G = R_0 a (T_{max} - T_{min})^b (1 + cP + dP^2)$
Wu et al. (2007)	WM	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> e RT	$R_G = R_0 [a + b(T_{max} - T_{min})^{0.5} + cT_{med} + dRT]$
Quej et al. (2016)	QJ	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> , RT e UR	$R_G = R_0 [(T_{max} - T_{min})^a + b](1 + cRH) + dRT$

Tabela 2. Coeficientes dos modelos empíricos HS, HM, AL, DS, WM e QJ calibrados para Iguape.

Table 2. Coefficients of the empirical models HS, HM, AL, DS, WM and QJ calibrated for Iguape.

Modelo empírico	Coeficiente calibrado			
	a	b	c	d
HS	0,172			
HM	0,209	-0,095		
AL	0,2001	1,086	0,562	-1,185
DS	0,156	0,564	-0,011	-0,00006
WM	-0,168	0,188	0,008	-0,071
QJ	0,226	-0,938	-0,002	-0,844

Tabela 3. Variáveis de entrada utilizadas pelos modelos SVM.  
Table 3. Input variables used by SVM models.

Modelo	Variáveis de entrada
SVM(1)	R <sub>0</sub> , T <sub>max</sub> e T <sub>min</sub>
SVM(2)	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> , e <sub>s max</sub> e e <sub>s min</sub>
SVM(3)	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> e P
SVM(4)	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> , P e UR
SVM(5)	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> , N e mês
SVM(6)	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> , e <sub>s max</sub> , e <sub>s min</sub> , N e mês
SVM(7)	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> , P, N e mês
SVM(8)	R <sub>0</sub> , T <sub>max</sub> , T <sub>min</sub> , P, UR, N e mês

### 3. RESULTADOS

A Figura 2 abaixo apresenta a evolução dos valores diários de R<sub>G</sub> em Iguape, dados medidos durante o período de 2008 a 2011 e utilizados na calibração, treinamento e validação dos métodos de estimativa testados. No mesmo gráfico, a R<sub>0</sub> diária é apresentada para comparação.

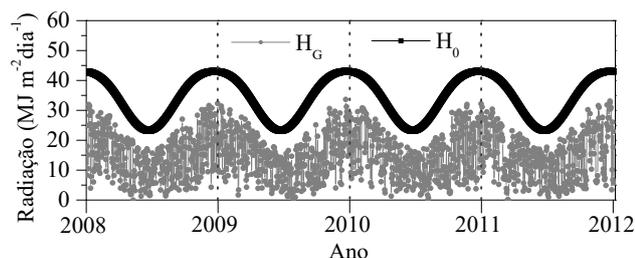


Figura 2. Evolução anual dos valores diários de R<sub>G</sub> e R<sub>0</sub> registrados em Iguape no período de janeiro de 2008 a dezembro de 2011.

Figure 2. Annual evolution of daily values of R<sub>G</sub> and R<sub>0</sub> recorded in Iguape in the period from January 2008 to December 2011.

A evolução temporal da R<sub>G</sub> durante o período considerado é periódica anualmente e acompanha as variações da R<sub>0</sub>, sendo a sazonalidade decorrente de fatores astronômicos (declinação solar) e climáticos (nebulosidade, umidade do ar e turbidez atmosférica). O espalhamento observado nos valores de R<sub>G</sub>, com grande amplitude de variação em relação a R<sub>0</sub>, é resultante principalmente da influência de nuvens e vapor de água modificando a transmissividade atmosférica.

Na Figura 3 abaixo são apresentadas as evoluções anuais dos valores médios de precipitação pluvial acumulada (P), temperatura do ar (T<sub>ar</sub>), umidade relativa do ar (UR) e frequência das condições de céu, todos obtidos a partir dos dados do período de janeiro de 2008 a dezembro de 2011. A classificação das condições de céu foi realizada a partir do índice de transmissividade atmosférica K<sub>T</sub>, seguindo metodologia proposta por Escobedo et al. (2009), onde são consideradas as seguintes classes: céu nebuloso (K<sub>T</sub> ≤ 0,35); céu parcialmente nebuloso com dominância para o difuso – PNDD (0,35 > K<sub>T</sub> ≤ 0,55); céu parcialmente nebuloso com dominância para o claro – PNDC (0,55 > K<sub>T</sub> ≤ 0,65); e céu claro (K<sub>T</sub> > 0,65).

Iguape possui um clima caracterizado pela elevada nebulosidade, umidade relativa do ar e precipitação acumulada. Nos meses mais quentes do ano (verão austral), a temperatura média mensal diária varia de 23,55°C (dezembro) a 25,75°C (fevereiro), e no mês mais frio é de 17,12°C (junho). A umidade relativa do ar é maior nos meses de outono e inverno, sendo os meses de fevereiro e julho os mais e menos úmidos, respectivamente. A precipitação pluvial acumulada é maior nos meses do verão austral e em julho, sendo em todos os meses do ano superior a 90mm e o total acumulado anual igual a 2.310,35mm. Os meses de janeiro e maio são os mais e menos chuvosos, respectivamente.

Em relação à nebulosidade, ao longo do ano há a predominância de dias com céu nebuloso ou PNDD. Nas duas condições de céu, a concentração e a espessura das nuvens são maiores em relação ao céu PNDC ou claro e reduzem o K<sub>T</sub> até valores inferiores a 0,55, e fazem a radiação global ser composta com predominância pela radiação difusa (ESCOBEDO et al., 2009). Em setembro, a frequência das duas condições de céu juntas é de 23 dias enquanto em dezembro é de 17 dias e, por isso, tais meses podem ser considerados o mais e menos nebuloso, respectivamente. A frequência de dias com céu claro em todos os meses do ano é inferior a 6, sendo igual a 2 em setembro.

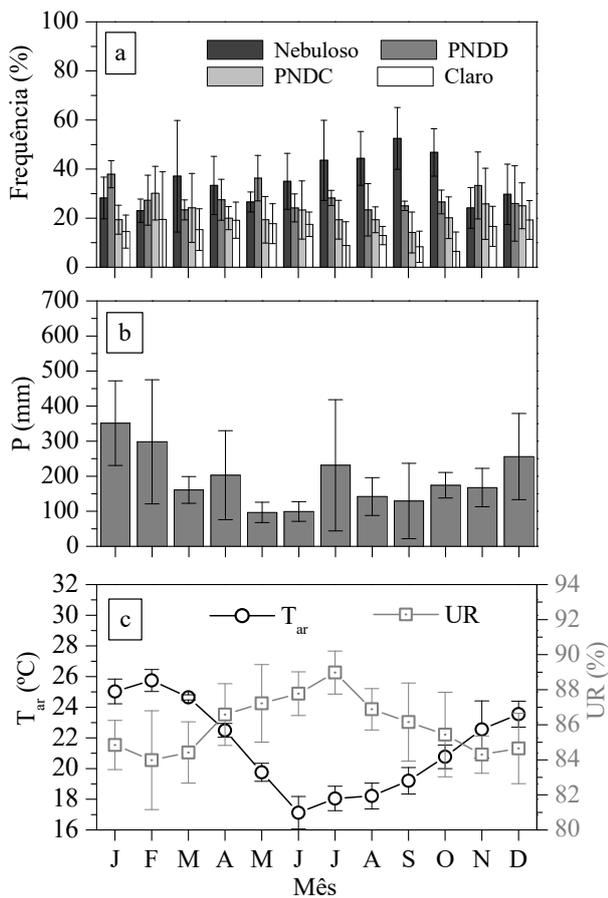


Figura 3. Médias mensais diárias: (a) Frequência das condições de céu nebuloso, parcialmente nebuloso com dominância para o difuso (PNDD), parcialmente nebuloso com dominância para o claro (PNDC) e claro; (b) Precipitação pluvial acumulada; (c) Temperatura do ar e umidade relativa do ar.

Figure 3. Monthly means of daily values: (a) Frequency of sky conditions cloudy, partially cloudy with dominance of diffuse (PNDD), partially cloudy with dominance of clear (PNDC) and clear; (b) Accumulated pluvial precipitation; (c) Temperature and moisture of air.

As variações sazonais da  $R_G$  podem ser verificadas com mais detalhe na Figura 4 abaixo. Os valores apresentados na Figura 4 são valores médios mensais interanuais, que foram compostos pelos valores médios mensais diários de cada um dos 4 anos (período de 2008 a 2011). Os valores de desvios padrões (linhas verticais) indicam a variabilidade interanual. Valores elevados de desvio padrão indicam elevada variabilidade interanual.

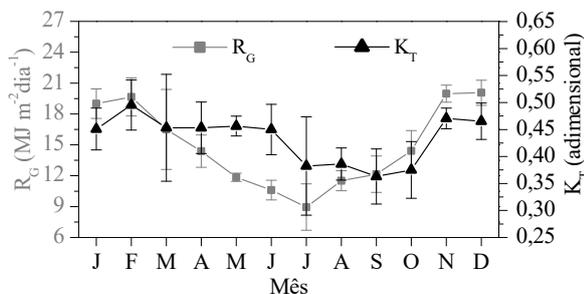


Figura 4. Evolução anual das médias mensais diárias da  $R_G$  e do  $K_T$ , valores médios interanuais obtidos a partir da base de dados medidas no período de janeiro de 2008 a dezembro de 2011.

Figure 4. Annual evolution of monthly means of daily values of  $R_G$  and  $K_T$ , interannual values obtained from the database measured in the period from January 2008 to December 2011.

Os valores médios de  $R_G$  foram maiores nos meses do verão austral, quando a declinação solar apresenta valores próximos ao da latitude local e a atmosfera apresenta as menores concentrações de nuvens e vapor de água do ano. Por outro lado, os meses do inverno austral são os mais úmidos e nebulosos e, por isso, apresentam os menores valores de  $R_G$ . Os meses de março e julho apresentam maior amplitude de variação interanual da  $R_G$ . Por outro lado, os meses de novembro e janeiro e também maio apresentaram as menores variações interanuais da  $R_G$ . Em relação à transmissividade atmosférica, verifica-se que ela é aproximadamente constante ( $K_T$  apresentando valores próximos de 0,45) ao longo do ano, com exceção dos meses de fevereiro, julho, agosto, setembro e outubro.

A calibração dos modelos empíricos, treinamento da SVM e validação dos dois métodos de estimativa ocorreram nas condições climáticas da  $R_G$  citadas acima. Em específico, a validação foi realizada a partir dos dados meteorológicos registrados no ano de 2010. Na Tabela 4 abaixo são apresentados os valores dos parâmetros estatísticos MBE, RMSE e índice d obtidos na validação dos modelos empíricos. Quanto mais próximo de zero são os valores de MBE e RMSE e mais próximo de 1 o valor do índice d, melhor é o desempenho do modelo analisado. Os valores de  $R_G$  diários estimados são apresentados na Figura 5, gráfico de dispersão (estimado x medido) para comparação visual. As regressões lineares entre os valores medidos e os estimados estão também apresentados na Figura 5 de forma algébrica e geométrica (linha contínua vermelha); e a linha contínua azul representa a reta de referência 1:1.

Tabela 4. Valores de MBE, RMSE e índice d verificados na validação dos modelos empíricos.

Table 4. MBE, RMSE and d index values verified in the validation of the empirical models.

Método	MBE	RMSE	Índice d
HS	-0,15	30,45	0,87
HM	-0,38	29,24	0,89
AL	-0,91	27,66	0,91
DS	-3,76	29,87	0,89
WM	1,33	28,29	0,91
QJ	4,93	27,87	0,91

Dos seis modelos empíricos testados, três pertencem a categoria baseada na temperatura do ar e três são baseadas em outras variáveis meteorológicas. Dentre os três modelos empíricos que dependem apenas da mediação da temperatura do ar, o HS e o AL apresentaram o pior e o melhor desempenho, respectivamente. Em ambos os modelos empíricos foi observada uma tendência de subestimativa inferior a 1%, porém a extensão do erro foi maior e o índice d menor na validação do HS. Em relação aos modelos empíricos que utilizam outras variáveis meteorológicas medidas além da temperatura do ar, QJ e DJ apresentaram o melhor e o pior desempenho, respectivamente. Quando os seis modelos empíricos são comparados entre si, levando em consideração os valores de MBE, RMSE e índice d, verifica-se uma maior acurácia nas estimativas do AL.

A seguir são apresentados os resultados obtidos na validação dos modelos SVM. A tabela 5 contém os valores dos parâmetros estatísticos MBE, RMSE e índice d obtidos enquanto a Figura 6 apresenta os valores de  $R_G$  diários estimados, gráfico de dispersão (estimado x medido) para

comparação visual. Nos gráficos, a linha vermelha corresponde a regressão linear entre os valores medidos e os estimados, a linha contínua azul representa a reta de referência 1:1.

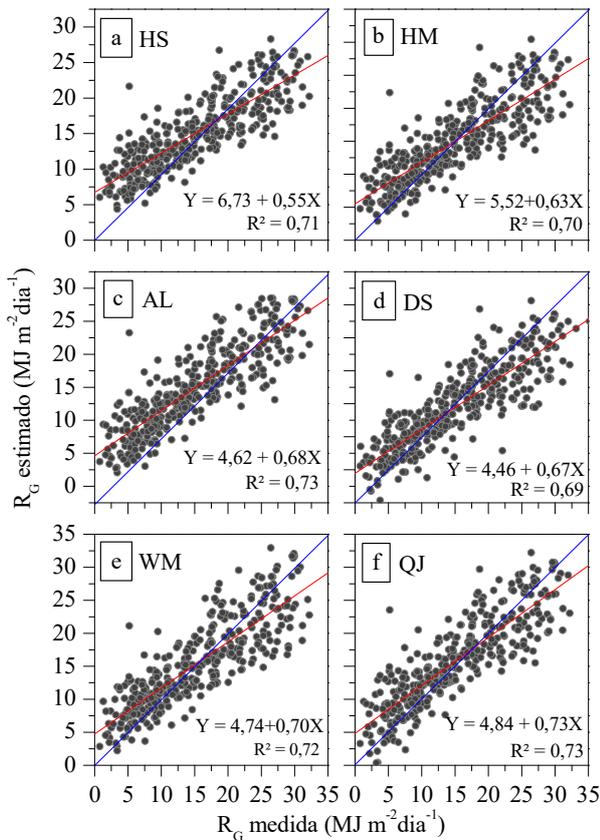


Figura 5. Gráficos de dispersão mostrando a comparação entre os valores diários de  $R_G$  medidos e os estimados utilizando os modelos empíricos.

Tabela 5. Valores de MBE, RMSE e índice d calculado verificados na validação dos modelos SVM.

Table 5. MBE, RMSE and d index values verified in the validation of the SVM models.

Método	MBE (%)	RMSE (%)	Índice d (adimensional)
SVM(1)	16,76	44,24	0,74
SVM(2)	-0,51	23,63	0,94
SVM(3)	-2,37	26,42	0,93
SVM(4)	-1,12	24,46	0,93
SVM(5)	-3,13	28,24	0,91
SVM(6)	1,03	26,04	0,93
SVM(7)	-2,64	26,34	0,93
SVM(8)	-1,16	25,14	0,92

Considerando agora o desempenho da SVM na estimativa da  $R_G$  diária, utilizando conjuntos de variáveis de entrada equivalentes aos dos modelos empíricos, ela apresentou maior acurácia na maioria dos casos. A única exceção foi a SVM(1), que apresentou o pior desempenho dentre todos os métodos avaliados no presente estudo, com uma tendência de superestimar a  $R_G$  diária em 16,76% e extensão do erro de 44,24%. Outro detalhe verificado na maioria dos casos avaliados foi a tendências de subestimativa da radiação solar global diária (MBE < 0) pela SVM.

Em geral, a acurácia nas estimativas aumentou de forma gradativa com a adição em sequência da P e da UR no conjunto de variáveis de entrada. O RMSE foi reduzido para 26,42%

com a adição da P e depois para 24,46% com a adição da UR, sendo o índice d igual a 0,93 em ambos os casos. A exceção nesse caso foi a SVM(2), que apresentou desempenho superior a SVM(3) e SVM(4) sem utilizar a P e a UR como variáveis de entrada e tendo como diferencial o usos dos parâmetros  $e_s$  max e  $e_s$  min.

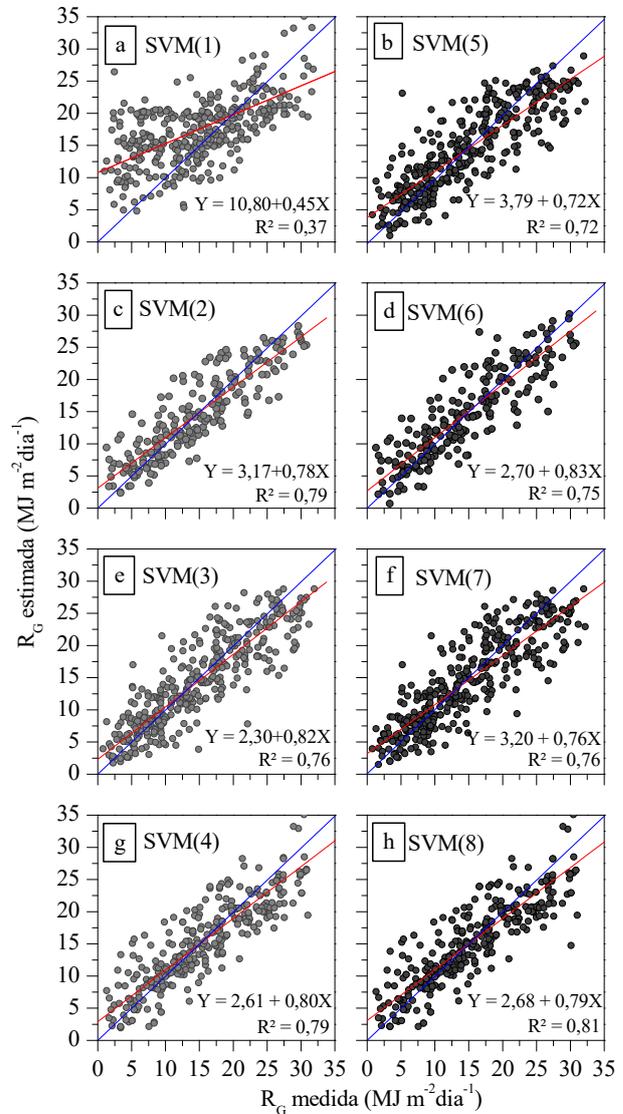


Figura 6. Gráficos de dispersão mostrando a comparação entre os valores diários de  $R_G$  medidos e os estimados utilizando os modelos SVM.

Figure 6. Scatter graph showing the comparison between measured and estimated values of daily  $R_G$  in the case of SVM models.

Os modelos SVM(5), SVM(6), SVM(7) e SVM(8) foram testados a fim de verificar se o uso de variáveis astronômicas melhoram a acurácia das estimativas da  $R_G$  pela SVM, ao permitir a capacidade de diferenciar os exemplos apresentados levando em consideração as variações sazonais. Tais modelos SVM apresentam as variáveis mês e N como diferencial em relação aos demais. Conforme Tabela 5, a adição das duas variáveis aumentou a acurácia nas estimativas apenas no caso da SVM(5), ou seja, quando são utilizadas junto das variáveis  $R_0$ ,  $T_{max}$  e  $T_{min}$ . Nesse caso, a extensão do erro nas estimativas foi reduzida de 44,42% para 28,24%. No caso da SVM(8), houve redução no nível de acurácia, que é verificado em comparação ao SVM(4).

#### 4. DISCUSSÃO

Numerosos modelos empíricos foram propostos na literatura para estimar a  $R_G$  diária utilizando diferentes variáveis meteorológicas, geográficas e astronômicas, que são resultantes de uma tarefa rigorosa para selecionar as variáveis de entrada mais adequadas para se obter o maior nível de acurácia possível (QUEJ et al., 2016; CHUKWUJINDU, 2017; MEENAL; SELVAKUMAR, 2018; HASSAN et al., 2018). Em geral, verifica-se a partir destes estudos que: (i) o aumento no número de variáveis de entrada utilizadas é limitado e implica em modelos empíricos complexos; (ii) modelos baseados no brilho solar apresentam maior acurácia em relação aos demais, porém, os baseados na temperatura do ar são os mais utilizados; (iii) a disponibilidade de dados das variáveis de entrada medidas é considerada um fator limitante no desenvolvimento e seleção dos modelos empíricos; (iii) as características climáticas e localização geográfica definem quais variáveis meteorológicas de entrada são as mais adequadas. O desempenho dos modelos empíricos também pode ser influenciado pela extensão da base de dados utilizada nas calibrações e/ou avaliações de desempenho, uma vez que bases de dados curtas podem eventualmente não abranger períodos com ocorrência de anomalias climáticas (SOUZA et al., 2017). Tais anomalias tendem a gerar maior amplitude de variação interanual na temperatura do ar, precipitação e/ou nebulosidade.

Portanto, na seleção de modelos empíricos clássicos que sejam adequados para um determinado local ou região, os resultados podem variar de um estudo para outro devido às diferenças climáticas e de extensão da base de dados utilizada. São considerados adequados os modelos empíricos com maior acurácia, que utilizam o menor número possível de variáveis de entrada medidas e que sejam simples na forma algébrica.

Dos modelos empíricos avaliados, a maioria com base na temperatura do ar apresentou tendência de subestimar os valores diários de  $R_G$  enquanto os demais apresentaram tendência de superestimativa. A extensão dos erros na estimativa, expressa pelo RMSE, variou de 27,66% a 30,45%, sendo em geral maior nos modelos baseados apenas na temperatura do ar. Os valores observados do índice  $d$  permaneceram dentro do intervalo entre 0,87 e 0,91. Os níveis de acurácia verificados foram próximos dos divulgados por outros autores na literatura (SILVA et al., 2012; SOUSA et al., 2017; QUEJ et al., 2016).

Dentre os seis modelos empíricos testados, o AL apresentou o maior nível de acurácia na estimativa da  $R_G$  diária. Tal resultado se difere dos observados na maioria dos estudos realizados em outras localidades de clima úmido. Quej et al. (2016) na Península de Yucatán, México, região de clima tropical úmido (Aw pelo critério de classificação de Köppen), verificaram maior acurácia no modelo QJ quando comparado aos modelos HS, AL e WM. Silva et al. (2012) em quatro cidade da região noroeste do Estado de Minas Gerais, Brasil, região de clima tropical úmido (Aw), verificam maior acurácia no modelo DS quando comparado aos modelos HS e HM. Porém, Souza et al. (2017) na região de bioma Amazônico e transições no Estado do Mato Grosso, Brasil, com clima tropical úmido (Aw), também verificam maior acurácia do modelo AL quando comparado aos modelos HS, HM e DS. O modelo empírico AL foi proposto por Almorox et al. (2011) e, quando validado utilizando dados medidos na cidade de Madri, Espanha, com clima Csa pelo critério de classificação de Köppen, apresentou desempenho superior ao de outros sete

modelos empíricos clássicos que também dependem apenas da medição da temperatura do ar.

Segundo Quej et al. (2016), quanto maior o número de variáveis de entrada utilizados, que apresentam correlação com a  $R_G$ , maior é a acurácia dos modelos empíricos. Porém, cada variável meteorológica utilizada apresenta em sua medida não só a imprecisão característica do sensor utilizado, mas também a imprecisão decorrente da má calibração ou manutenção deles. Por isso, nem sempre os modelos que utilizam um número maior de variáveis de entrada medidas irão apresentar o melhor desempenho, dada a possibilidade de ocorrerem erros experimentais na medição delas. Nesse contexto, a SVM apresenta como vantagem, em relação aos modelos empíricos, a elevada capacidade de generalização dos exemplos independente da presença de dados inconsistentes. Outra vantagem apresentada pela SVM é capacidade de atingir maior acurácia utilizando um número menor de variáveis de entrada (BELAID; MELLIT, 2016).

Sobre o uso da SVM, conforme esperado, houve uma tendência de aumento na acurácia das estimativas da  $R_G$  diária em função do aumento no número de variáveis de entrada. O pior desempenho da SVM foi observado com a utilização do conjunto de variáveis de entrada  $\{R_0, T_{max}, T_{min}\}$  sendo a acurácia observada muito inferior à dos modelos empíricos HS e HM.

A acurácia nas estimativas aumentou de forma gradativa primeiro com a adição da P e depois com a adição da UR no conjunto de variáveis de entrada. Tanto a P quanto a UR apresentam elevada correlação com a transmissividade atmosférica da radiação solar, especialmente em regiões úmidas e próximas do litoral. Os resultados verificados aqui são similares aos observados por Meenal; Selvakumar (2018) em 8 cidades na Índia, que observaram aumento da acurácia nas estimativas da  $R_G$  diária pela SVM ao adicionar a UR à base de dados composta pela  $R_0, T_{max}$  e  $T_{min}$ . Quej et al. (2017), na Península de Yucatán, obtiveram o melhor desempenho da SVM na estimativa da  $R_G$  diária utilizando o conjunto de variáveis de entrada  $\{R_0, T_{max}, T_{min}, P\}$ . Contudo, ressalta-se que o maior nível de acurácia da SVM em Iguape foi atingido ao utilizar o conjunto de variáveis de entrada  $\{R_0, T_{max}, T_{min}, e_{s_{max}}, e_{s_{min}}\}$ . Ressalta-se também que somente no caso do conjunto de variáveis de entrada  $\{R_0, T_{max}, T_{min}\}$  a adição das variáveis mês e N aumentou de forma expressiva a acurácia nas estimativas da  $R_G$  diária.

Os resultados obtidos aqui servem de referência não só na seleção do método de estimativa da  $R_G$  mais adequado para regiões de clima subtropical úmido no Brasil. Os modelos empíricos calibrados, que foram apresentados no presente estudo, também podem ser utilizados na estimativa da  $R_G$  diária em outras localidades com clima similar ao de Iguape, para a realização de estudos nas áreas de agricultura, aquicultura e meio ambiente. No caso específico do Vale do Ribeira, embora a região abrigue algumas universidades e núcleos de pesquisa com interesse em dados de radiação solar, a quantidade de informações e o acesso a elas ainda são restritos.

#### 5. CONCLUSÕES

Sobre os diferentes métodos de estimativa da radiação solar global diária testados em ambiente subtropical úmido do Vale do Ribeira, verifica-se:

Dentre todos os modelos empíricos avaliados, o proposto por Almorox et al. (2011) é considerado o mais adequado,

dado o maior nível de acurácia e por depender apenas da medição da temperatura do ar.

De acordo com o esperado, a SVM apresentou melhor desempenho na estimativa da radiação solar global diária se comparada aos modelos empíricos. Dentre os conjuntos de variáveis de entrada testados na SVM, o  $\{R_0, T_{\max}, T_{\min}, e_{s\max}, e_{s\min}\}$  permite maior acurácia na estimativa da radiação solar global, sendo necessária a disponibilidade apenas de dados medidos de temperatura do ar.

O uso das variáveis mês e fotoperíodo aumentam a acurácia nas estimativas da radiação solar global diária pela SVM apenas quando utilizadas junto das variáveis  $R_0, T_{\max}$  e  $T_{\min}$ .

## 6. AGRADECIMENTOS

Ao INMET (Instituto Nacional de Meteorologia) pelo fornecimento das séries históricas de dados meteorológicos.

## 7. REFERÊNCIAS

- ALMOROX, J.; HONTORIA, C.; BENITO, M. Models for obtaining daily global solar radiation with measured air temperature data in Madrid (Spain). *Applied Energy*, London, v. 88, p. 1703-1712, 2011. DOI: <https://dx.doi.org/10.1016/j.apenergy.2010.11.003>
- ALVARES, C. A.; STAPE, J. C.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, G. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, Berlin, v. 22, n. 6, p. 711-728, 2014. DOI: <https://dx.doi.org/10.1127/0941-2948/2013/0507>
- BAKHASHWAIN, J. M. Prediction of global solar radiation using support vector machines. *International Journal of Green Energy*, Abingdon, v. 13, n. 14, p. 1467-1472, 2016. DOI: <https://dx.doi.org/10.1080/15435075.2014.896256>
- BELAID, S.; MELLIT, A. Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate. *Energy Conversion and Management*, Oxford, v. 118, p. 105-118, 2016. DOI: <https://dx.doi.org/10.1016/j.enconman.2016.03.082>
- BESHARAT, F.; DEHGHAN, A. A.; FAGHIH, A. R. Empirical models for estimating global solar radiation: a review and case study. *Renewable and Sustainable Energy Review*, v. 21, p. 798-821, 2013. DOI: <https://dx.doi.org/10.1016/j.rser.2012.12.043>
- CITAKOGLU, H.; COBANER, M.; HAKTANIR, T.; KISI, O. Estimation of monthly mean reference evapotranspiration in Turkey. *Water Resources Management*, Cham, v. 28, n. 1, p. 99-113, 2014. DOI: <https://dx.doi.org/10.1007/s11269-013-0474-1>
- COELHO, V. S. P.; FAVARETO, A. S. Dilemas da participação e desenvolvimento territorial. *Revista de Desenvolvimento Econômico*, Salvador, v. 10, n. 18, p. 15-26, 2008.
- DE JONG, R.; STEWART, D. W. Estimating global solar radiation from common meteorological observations in western Canada. *Canadian Journal of Plant Science*, v. 73, n. 2, p. 509-518, 1993. DOI: <https://dx.doi.org/10.4141/cjps93-068>
- ESCOBEDO, J. F.; GOMES, E. N.; OLIVEIRA, A. P.; SOARES, J. Modeling hourly and daily fractions of UV, PAR and NIR to global solar radiation under various sky conditions at Botucatu, Brazil. *Applied Energy*, London, v. 86, p. 299-309, 2009. DOI: <https://dx.doi.org/10.1016/j.apenergy.2008.04.013>
- HALL, M.; FRANK, E.; HOLMES, G.; REUTEMANN, B. P.; WITTEN, I. H. The WEKA Data Mining Software: An Update, *SIGKDD Explorations*, New York, v. 11, n. 1, p. 10-18, 2009. DOI: <https://dx.doi.org/10.1145/1656274.1656278>
- HARGREAVES, G. H.; SAMANI, Z. A. Estimating potential evapotranspiration. *Journal of Irrigation and Drainage Engineering*, v. 108, n. 3, p. 225-230, 1982.
- IBGE\_INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Perfil dos Estados e dos Municípios Brasileiros 2014**. Disponível em: <[http://www.ibge.gov.br/home/estatistica/economia/perfil\\_munic/2014/](http://www.ibge.gov.br/home/estatistica/economia/perfil_munic/2014/)>. Acesso em: 05 mar 2016.
- IQBAL, M. **An introduction to solar radiation**. London: Academic Press, 1983. 390 p.
- KUMAR, R.; AGGARWAL, R. K.; SHARMA, J. D. Comparison of regression and artificial neural network models for estimation of global solar radiations. *Renewable and Sustainable Energy Reviews*, v. 52, p. 1294-1299, 2015. DOI: <https://dx.doi.org/10.1016/j.rser.2015.08.021>
- MEENAL, R.; SELVAKUMAR, I. Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters. *Renewable Energy*, Oxford, v.121, p.324-343, 2018. DOI: <https://dx.doi.org/10.1016/j.renene.2017.12.005>
- MEHDIZADEH, B.; BEHMANSH, J. N.; KHALILI, K. Comparison of artificial intelligence methods and empirical equations to estimate daily solar radiation. *Journal of Atmospheric and Solar-Terrestrial Physics*, Oxford, v. 146, p. 215-227, 2016. DOI: <https://dx.doi.org/10.1016/j.jastp.2016.06.006>
- MENDELSON, M.; LOWDER, T.; CANAVAN, B. **Rating Utility-scale photovoltaic, concentrating photovoltaic, and concentrating solar power**. Colorado: National Renewable Energy Laboratory, 2012. 55 p.
- MOHAMMADI, K.; SHAMSHIRBAND, S.; ANISI, M. H.; ALAM, K. A.; PETKOVIC, D. Support vector regression based prediction of global solar radiation on a horizontal surface. *Energy Conversion and Management*, Oxford, v. 91, p.433-441, 2015. DOI: <https://dx.doi.org/10.1016/j.enconman.2014.12.015>
- QUEJ, V. H.; ALMOROX, J.; ARNALDO, J. A.; SAITO, L. 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*, Oxford, v. 155, p. 62-70, 2017. DOI: <https://dx.doi.org/10.1016/j.jastp.2017.02.002>
- QUEJ, V. H.; ALMOROX, J.; IBRAKHIMOV, M.; SAITO, L. Empirical models for estimating daily global solar radiation in Yucatán. *Energy Conversion and Management*, Oxford, v. 110, p. 448-456, 2016. DOI: <https://doi.org/10.1016/j.enconman.2015.12.050>
- SILVA, C. R. da; SILVA, V. J.; ALVES JÚNIOR, J.; CARVALHO, H. de P. Radiação solar estimada com base na temperatura do ar para três regiões de Minas Gerais. *Revista Brasileira de Engenharia Agrícola e Ambiental*, Campina Grande, v.16, n. 3, p.281-288, 2012.
- SILVA, M. B. P.; ESCOBEDO, J. F.; SANTOS, C. M.; ROSSI, T. J.; GORLA-SILVA, S. H. M. Performance of the Angstrom-Prescott Model (A-P) and SVM and ANN

- Techniques to estimate the daily Global Solar Irradiation in Botucatu/SP/Brazil. **Journal of Atmospheric and Solar-Terrestrial Physics**, Oxford, v. 155, p. 62-70, 2017. DOI: <https://dx.doi.org/10.1016/j.jastp.2017.04.001>
- SOUZA, A. P.; SILVA, A. C.; TANAKA, A. A.; ULIANA, A. M.; ALMEIDA, F. T.; KLAR, A. E.; GOMES, A. W. A. Global radiation by simplified models for the state of Mato Grosso, Brazil. **Pesquisa Agropecuária Brasileira**, Brasília, v. 52, n. 4, p. 215-227, 2017. DOI: <http://dx.doi.org/10.1590/S0100-204X2017000400001>
- VALENTI, W. C.; KIMPARA, J. M.; PRETO, B. L.; MORAES-VALENTI, P. Indicators of sustainability to assess aquaculture systems. **Ecological indicators**, v. 88, p. 402-413, 2018. DOI: <https://dx.doi.org/10.1016/j.ecolind.2017.12.068>
- WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. **Climatic Research**, v.30, p.79-3, 2005. DOI: <https://dx.doi.org/10.3354/cr030079>
- WILLMOTT, C. J. On the validation of models. **Physical Geography**, v. 2, p. 184-194, 1981. DOI: <https://dx.doi.org/10.1080/02723646.1981.10642213>
- WMO\_WORLD METEOROLOGICAL ORGANIZATION. **Guide of Meteorological Instruments and Methods of Observation**. 2008. 681 p.