

REDE PERCEPTON MULTICAMADAS (MLP) NA ESTIMATIVA DA FRAÇÃO DA IRRADIAÇÃO SOLAR DIFUSA NA PARTIÇÃO DIÁRIA.

E. L. Bassetto¹, L. R. Z. Ortega², J. F. Escobedo³ A. Dal Pai⁴.

¹Departamento de Engenharia Elétrica – UTFPR – Cornélio Procópio/PR/Brasil

²Departamento de Engenharia Rural – FCA/UNESP – Botucatu/SP/Brasil

³Departamento de Engenharia Rural – FCA/UNESP – Botucatu/SP/Brasil

⁴Departamento de Bioprocessos e Biotecnologia – FCA/UNESP – Botucatu/SP/Brasil

Tel. + 554399976-1739 e-mail: bassetto@utfpr.edu.br

Recibido 08/08/18, aceptado 19/09/18

RESUMO: Este trabalho propõe a utilização de Redes Neurais Artificiais (RNA/MLP) na estimativa da fração difusa (K_d) da irradiação solar a partir de um conjunto de combinações de variáveis astronômicas e meteorológicas medidas e calculadas. Os dados são obtidos no período de 2000 a 2006 na Estação Meteorológica da Universidade Estadual Paulista, campus Botucatu/SP. Para verificar o desempenho das técnicas é elaborado um modelo estatístico (ME) utilizando somente a fração transmitida da irradiação solar global (K_T) como variável de referência. As variáveis de entrada para treinamento das RNA/MLP compreendem valores de oito variáveis astronômicas e meteorológicas medidas e calculadas. Para validação a base total de dados é reorganizada formando duas bases denominadas de Ano Típico (AT) e Ano Atípico (AAT). O desempenho é avaliado pelo coeficiente de correlação (r) e o índice de precisão (RMSE). O modelo estatístico (ME) elaborado e ajustado por regressão polinomial de 4ª ordem apresenta um coeficiente de determinação (R^2) de 0,91626 o que mostra que K_d está bem correlacionado com K_T estatisticamente. O índice de desempenho do modelo estatístico (ME) na validação com as duas bases (AT e AAT) foi de $RMSE_{AT} = 19,95\%$ e $RMSE_{AAT} = 18,66\%$ o que mostra uma boa aproximação dos valores estimados aos medidos e de correlação de $r_{AT}=0,9533$ e $r_{AAT}=0,9490$ o que mostra a importância das variáveis independente para variável dependente. Com as RNA/MLP os indicadores de desempenho (RMSE) apresentaram uma redução em relação ao (ME) em ambas as bases de validação em aproximadamente 20%, o que mostra que a inserção de variáveis contribui nas estimativas de (K_d) com as condições astronômicas, meteorológicas e geográficas utilizadas.

Palavras-chave: Irradiação difusa; Redes Neurais Artificiais; Perceptron Multicamadas.

INTRODUÇÃO

A radiação solar é considerada uma energia limpa, sustentável e abundante na superfície terrestre. Países com grande incidência dessa energia como o Brasil, são beneficiados com o uso desta fonte em diversas áreas científicas e tecnológicas (Bakirci, 2015; Teke et al., 2015). No entanto para um bom aproveitamento é importante quantificar ou estimar este potencial através de medidas (séries temporais ou espaciais) ou por meio de modelos de estimativa para verificar e avaliar a disponibilidade desta energia ao longo do tempo em uma determinada região. Dos componentes da radiação solar as medidas da irradiância solar global ($I_g - W/m^2$) que é a soma da irradiância solar direta na horizontal ($I_b - W/m^2$) com a irradiância solar difusa na horizontal ($I_d - W/m^2$), é a irradiância monitorada com maior frequência nas estações solarimétricas, diferentemente das medidas de (I_b) que são importantes em projetos de conversão fotovoltaico e térmico (Jamil e Akhtar, 2017) e da (I_d) importante em estudos de energia passiva e biomassa (Dal Pai et al, 2016; Gueymard, 2008) com também no desenvolvimento de cultivares protegidos (Escobedo et al., 2013). Contudo as medidas de (I_b) que demandam de equipamentos custosos geralmente são calculadas pelo método da diferença entre (I_g) e (I_d) pela relação ($I_b = I_g - I_d$).

¹ Engenheiro Eletricista, Doutorando em Energia na Agricultura, FCA/UNESP/Botucatu.

² Engenheira Ambiental, Mestranda em Energia na Agricultura, FCA/UNESP/Botucatu.

³ Físico, Prof. Dr. Departamento de Engenharia Rural, FCA/UNESP/Botucatu.

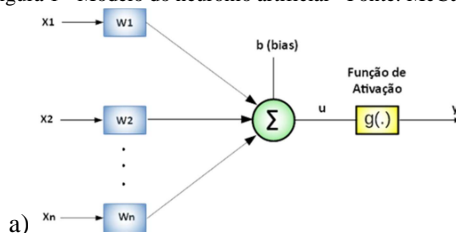
⁴ Físico, Prof. Dr. Departamento de Bioprocessos e Biotecnologia, FCA/UNESP/Botucatu.

As medidas da (I_d) também não são comuns mas são monitoradas nos últimos anos com custos menores (Dal Pai et al., 2016). No Brasil essas medidas são realizadas em poucas regiões (Marques Filho et al., 2016; Drechmer e Ricieri 2006; Codato et al., 2008; Furlan, et al., 2012; Soares et al., 2004) através do anel de sombreamento com baixo custo em equipamentos e operação. A proposta elaborada por Melo e Escobedo (1994) está em uma montagem simples onde o anel é fixo a um eixo no plano horizontal e inclinado na latitude local, e o eixo central do anel paralelo a direção polar com o sensor que transladando sobre a horizontal nesse plano, diferente de outras montagens (Drumonnond, 1956 e Robinson e Stoch, 1964; Martinez e Cúnsulo, 2017) que possuem restrições operacionais, de manutenção e custos elevados. Apesar da importância, medidas ou estimativas da radiação difusa ainda são escassos e poucas informações disponíveis, principalmente no Brasil, país com elevada disponibilidade de recurso solar. (Silva et al. 2008; Oliveira et al., 2002; Santos et al., 2016). Diferentemente das medidas (séries temporais ou espaciais) os modelos de estimativa necessitam de informações de outras medidas na superfície terrestre como modelos estatísticos e técnicas de aprendizado de máquinas. (Jamil e Akhatar, 2017; Khorasanizadeh et al., 2016; Ruiz-Arias et al., 2010). A maioria dos modelos para estimar a irradiação solar difusa ($H_d - MJ/m^2$) são estatísticos contudo para minimizar os efeitos das variações temporais é proposto a utilização de frações permitindo em partes minimizar alguns efeitos astronômicos e geográficos. Os primeiros trabalhos a correlacionaram a fração da irradiação solar difusa ($K_d = H_d/H_g$) onde ($H_g - MJ/m^2$) é a irradiação global, com a fração transmitida da irradiação solar global ou índice de claridade ($K_T = H_g/H_0$) onde ($H_0 - MJ/m^2$) é a irradiação no topo da atmosfera, foram propostos por Liu e Jordan (1960) a partir de uma base de dados destas variáveis na partição diária. Este modelo com apenas uma variável de referência é utilizado na literatura em partições horária, diária e mensal (Orgill e Hollands, 1977; Collares-Pereira e Rabl, 1979; Chandrasekaran and Kumar, 1994; Oliveira et al., 2002; Jacovides et al., 2006; Paleuscu e Blaga, 2016), como também outros com mais de uma variável (Reindl et al., 1990; Ridley et al., 2010; Furlan et al., 2012; Kuo et al., 2014). Outras modelos de estimativas são as Técnicas de Aprendizado de Máquinas (TAM) do tipo Redes Neurais Artificiais (RNA), Máquinas de Vetores de Suporte (SVM), e de técnicas híbridas com uso das RNA's e Sistema de Inferência Fuzzy (FIS) entre outras. Estas técnicas permitem utilizar diferentes parâmetros de grandeza físicas diferentes e através de algoritmos prever medidas a partir de uma base de dados conhecida (Khatib et al, 2012; Paulescu et al, 2016). Diante destas considerações e da importância de se conhecer as medidas da irradiação solar difusa na superfície horizontal pretende-se utilizar a (RNA/MLP) na estimativa da fração da irradiação solar difusa (K_d) na partição diária como base em dados climatológicos obtidos na cidade de Botucatu/SP e ao final compará-la com modelo estatístico (ME) de referência elaborado com a mesma base de dados e com apenas uma variável independente.

REDES NEURAIS ARTIFICIAIS (RNA)

A técnica de RNA são modelos computacionais que apresentam uma estrutura formada por um conjunto de elementos chamados neurônios, similar às do cérebro humano, distribuídos paralelamente e composta por camadas de entrada, de saída e camadas ocultas interligadas entre si por conexões. Modeladas matematicamente que através de algoritmos de aprendizado são capazes de analisar funções complexas e não lineares que adquirem a capacidade de aprendizado e adequação. São utilizadas para o reconhecimento e classificação de padrões, agrupamento, previsão e nos últimos anos na estimativas de variáveis astronômicas. A estrutura básica de uma rede neural é o neurônio artificial ilustrado pela Figura 1. É formada por um conjunto de conexões (*sinapses*) que recebem os valores de entrada (x_1, x_2, \dots, x_i) definidos por pesos sinápticos (w_1, w_2, \dots, w_i) que através de um combinador linear (Σ) executa a soma ponderada de suas entradas, produzindo uma saída chamada de potencial de ativação (u). Na sequência é transformada por uma função de ativação (f), que gera a saída (y) limitada por um intervalo de valores definido pelo limiar de ativação (*bias - b*) (Haykin, 2001).

Figura 1 - Modelo do neurônio artificial - Fonte: McCullock – Pitts (1943)



A RNA a ser utilizada é a Perceptron Multicamadas (MLP) caracterizada pela presença de pelo menos uma camada intermediária de neurônios, localizada entre as camadas de entrada e a camada de saída e são classificadas como redes do tipo *feedforward* (alimentação a frente ou única direção) de camadas múltiplas (Moustris et al., 2008) e treinada com o algoritmo *Levenberg-Marquadt* (Çelik et al., 2016). A principal característica de uma RNA está na capacidade de aprendizado proporcionado pela arquitetura e no processo de treinamento utilizado. O aprendizado é consequência das mudanças de valores dos pesos sinápticos, conectados aos neurônios entre as camadas com o objetivo de aproximar os valores da saída da rede com os valores desejados (Alam et al. 2009). Para utilização da RNA/MLP é importante a composição da estrutura, a definição do número de neurônios, das camadas ocultas, do algoritmo de treinamento, da função de transferência e do conjunto de dados que podem influenciar nos pesos sinápticos e nos resultados. Com uma arquitetura flexível e pouca complexidade nos algoritmos de treinamento e parametrização são as mais utilizadas na estimativa da irradiação solar. Na estimativa da irradiação global (H_g) (Elizondo et. al, 1994; Kalogirou, 2000; Tymvios et. al, 2005; Kumar et. al, 2015), com diversas variáveis climatológicas e geográficas como variáveis de entrada. Na irradiação solar direta (H_b) poucos trabalhos são encontrados na literatura (Alam et. al, 2006; Tomar et. al, 2012; Kaushika et al., 2014), assim como para irradiação solar difusa (H_d) (Soares et. al, 2004; Alam et. al, 2009; Ihya et. al, 2014; Boznar et. al, 2017).

MATERIAL E MÉTODOS

Localização e Clima

A Estação Meteorológica do Lageado está localizada no campus da UNESP – Universidade Estadual Paulista em Botucatu/SP/Brasil (latitude 22,85°S, longitude 48,45°O) e está a 786 m do nível do mar com uma elevada diferença entre a região mais baixa (400 a 500 m) e a mais alta (700 a 980) o que provoca mudanças na temperatura do ar e nos ventos (Santos e Escobedo, 2016). De acordo com a classificação de Koppen, o clima é do tipo Cwa, com bioma de cerrado e mata atlântica, temperado úmido, com inverno frio e seco (junho-agosto), verão quente e úmido (dezembro-fevereiro) com elevada precipitação (Escobedo et al. 2011; Teramoto e Escobedo, 2012). A temperatura do ar média anual é de 20,46°C, onde o mês mais quente é fevereiro (22,5°C) e o mais frio julho (16,7°C). A umidade relativa do ar varia de 61,20% (agosto) a 76,5% (janeiro). A evolução da nebulosidade segue a variação da temperatura sendo mais elevada nos meses do verão (61%) e menor nos meses de inverno (28%). A precipitação segue as evoluções da temperatura e da nebulosidade sendo constituído de dois intervalos distintos. O período chuvoso (outubro a março/primavera e verão) com maior precipitação no verão com média total de 304mm e o período seco (abril a setembro/outono e inverno) com menor valor no inverno com média de 38mm.

Obtenção da Base de dados

As medidas foram obtidas simultaneamente da (I_g), da (I_d), da (I_{bn}), irradiância de ondas longas (I_{wa}), temperatura (*temp*) e umidade (*rh*) no período de 2000 a 2006. Nas medidas instantâneas (W/m^2) da irradiância solar global (I_g) foi utilizado um piranômetro EPPLEY-PSP, e para medidas da irradiância solar direta normal (I_{bn}) foi usado um pireliômetro Eppley-NIP acoplado a um rastreador solar do tipo ST-3 e transformada para irradiação direta projetada na horizontal (I_b) através da Eq. (1): $I_b = I_{bn} \cdot \cos \theta_z$ (1) onde θ_z é o ângulo zenital (graus). Para medidas da irradiância solar difusa (I_d) um piranômetro EPPLEY-PSP, como o anel de sombreamento (Melo e Escobedo, 1994) com as seguintes especificações: 80 cm de diâmetro, 10 cm de largura e 1mm de espessura. Os valores da irradiação solar difusa (I_d) foram corrigidos isotropicamente para compensar a parcela de (I_d) obstruída pelo próprio anel, as quais depende de parâmetros geométricos como largura e raio, geográficos como latitude e os astronômicos como a declinação solar (Oliveira et. al, 2002). Para medidas de irradiância de ondas longas atmosféricas (I_{wa}), um pirgeômetro, modelo CG1 da Kipp e Zonen e para medidas de temperatura do ar (°C) e umidade relativa (%) um termógrafo e um higrógrafo da Filotécnica Milano, instalados dentro de um abrigo meteorológico. Todos os equipamentos são aferidos pelo método comparativo a cada dois anos, através de instrumentos adicionais específicos para esta finalidade. Na aquisição e armazenamento dos dados foi utilizado um sistema de aquisição automático digital “*Datalogger*” Campbell Cr23X da *Campbell Scientific*, operando na frequência de 1Hz, onde os valores foram digitalizados em intervalo de 5 segundos, armazenando médias em intervalos de 5 minutos, transmitidos para um computador através de um módulo de armazenamento SM – 192. Após obtenção das medidas instantaneamente (unidade - W/m^2), os dados foram posteriormente integrados

na partição diária (unidade - MJ/m^2) (Chaves e Escobedo, 2000). Todas as medidas realizadas passaram por um controle de qualidade, considerando a possibilidade das medições estarem sujeitas a erros sistêmicos ou aleatórios, seja pelas condições do dia da medida ou do equipamento e de valores discrepantes, adotando-se aqui alguns testes de procedimentos como limitação de irradiação global diária (H_g) em função da irradiação no topo da atmosfera (H_0), da fração transmitida da irradiação global (K_T) e da fração da irradiação difusa diária (K_d) (De Miguel et al. 2001; Lopez et al. 2004; Younes et al. 2005; Ruiz-Arias et al. 2010; Ihya et al, 20015; Pashiardis e Kalogirou, 2016). Foram retirados aproximadamente 8,5% do total de dados medidos do período.

Modelo Estatístico (ME)

Na elaboração do modelo estatístico os valores da fração da irradiação difusa ($K_d = H_d/H_g$) e da transmissividade atmosférica ($K_T = H_0/H_g$), calculados a partir de dados experimentais medidos foram correlacionados (Liu e Jordan, 1960) e suas equações de estimativa ajustadas por regressão polinomial entre as frações.

Elaboração e validação do Modelo Estatístico (ME)

Na concepção do ME os dados foram reorganizadas retirando-se uma parte para elaboração e outra para validação do modelo de acordo com Tabela 1. Para validação do modelo, os dados foram reorganizados em duas bases denominadas por Ano Típico (AT) e Ano Atípico (AAT) para irradiação solar difusa.

Tabela 1 - Dados para elaboração e validação modelo estatístico

Total de dados (2000 – 2006)	1.1 Dados para elaboração do ME	
2.348 (100%)	1.703 – 72,53%	
	1.2 Dados para validação do ME	
	Base AT	Base AAT
	324 – 13,79%	321 – 13,67%
	2.1 Dados para treinamento da rede MLP	
	Base - AT	Base - AAT
	2.024 – 86,20%	2.027 - 86,33%
	2.2 Dados para validação da rede MLP	
324 – 13,80%	321 – 13,68%	

Esta reorganização foi realizada através de relação estatística da média do valor de H_d de cada mês de toda a série assim como o intervalo de variação do desvio padrão. Comparou-se para cada mês de cada ano o valor de H_d , a média horária entre os anos, considerando o intervalo de variação do desvio padrão da média total com o valor de H_d , da média de cada ano. Após análise selecionou-se para cada mês da série os valores próximos a média (acima ou abaixo), como Ano Típico (AT) e mais afastados (acima ou abaixo), como Ano Atípico (AAT). Como descrito a seleção dos meses que constituem um Ano Típico (AT), é similar ao processo de seleção do Ano Meteorológico Típico (WMO, 1981). A Tabela 2, apresenta os meses e os anos que constituem um Ano Típico (AT) e um Ano Atípico (AAT) dos dados analisada de H_d . No destaque na Tabela 2, para o mês de janeiro, o AT foi 2003, enquanto Julho foi 2005.

Tabela 2 - Base Ano Típico (AT) e Ano Atípico (AAT)

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
AT	2003	2000	2001	2001	2000	2001	2005	2001	2006	2006	2005	2001
AAT	2002	2003	2000	2004	2004	2002	2004	2000	2005	2005	2001	2005

Realizada a fase de organização da base de dados, descreve-se na sequência a metodologia utilizada para treinamento e validação da técnica RNA/MLP e sua parametrização.

Técnicas de aprendizado de máquinas

No treinamento da RNA/MLP foram desenvolvidos *script's* para treinamento e validação, com diferentes funções e parametrização. A validação foi realizada com a base apresentada na Tabela 1. Diferentemente do modelo estatístico, que utiliza um conjunto de dados com as variáveis K_T e K_d para modelagem, no treinamento da RNA/MLP são definidos uma sequência de combinações com variáveis climatológicas determinadas a partir dos dados medidos e ou calculados, disponíveis como entrada para as redes de acordo com Tabela 3. Neste estudo optou-se em utilizar variáveis considerando sua importância, disponibilidade, período de medidas e características físicas em relação a variável a ser estimada. Outros arranjos foram analisados, porém em função das condições de seleção das variáveis, as mesmas não foram utilizadas.

As variáveis consistem em valores medidos de H_0 , irradiação no topo da atmosfera (MJ/m^2), de H_g , irradiação solar global (MJ/m^2), de H_{wa} , irradiação de ondas longas atmosféricas (MJ/m^2), *umid*, umidade (UR , %) e *tar*, temperatura (T , °C) e calculados como transmissividade atmosférica ou índice de claridade (K_T) e fração difusa (K_d), *decsol*, declinação solar (°C) e *nebul*, nebulosidade (%) calculada em função da proporcionalidade das medidas da irradiância solar direta ($(I_b - W/m^2)$ dos valores abaixo de $120 W/m^2$. Conforme a WMO (2008), em seu guia de instrumentos meteorológicos e métodos de observação, este valor foi adotado com limiar para o brilho do sol, ou seja, para valores inferiores a esses, considera-se sem insolação ou “nebuloso”.

Tabela 3 - Combinações e variáveis de entrada

Combinações	Variáveis de: Entrada								Saída
combinação 1	Kt								K _d
combinação 2	Kt	H ₀							
combinação 3	Kt	H ₀	H _g						
combinação 4	Kt	H ₀	H _g	H _{wa}					
combinação 5	Kt	H ₀	H _g	H _{wa}	decsol				
combinação 6	Kt	H ₀	H _g	H _{wa}	decsol	nebul			
combinação 7	Kt	H ₀	H _g	H _{wa}	decsol	nebul	umid		
combinação 8	Kt	H ₀	H _g	H _{wa}	decsol	nebul	umid	tar	

medido/calculado
 calculado
 medido

Definido as combinações das variáveis de entrada apresentam-se na sequência a metodologia para treinamento da rede.

Rede Perceptron Multicamadas (MLP)

Na implementação da técnica MLP os conjuntos de dados foram submetidos para execução no software MATLAB R2014 que pode ser descrito como uma ferramenta computacional com aplicação em diversas áreas. Permite o desenvolvimento e análise de sistemas matemático, estatísticos, de otimização e controle com uma linguagem de programação dinâmica com uma ampla disponibilidade de funções para serem implementadas nas técnicas de aprendizado de máquinas. O treinamento da rede MLP foi realizado com o algoritmo *Levenberg-Marquadt* (*trainlm*) e implementado pela função *feedforwardnet* e parametrizado pelo número de camadas ocultas, do número de neurônios, da quantidade máxima de épocas de treinamento e das funções de ativação. A definição do número de camadas foi realizado em duas fases, sendo a primeira com apenas uma camada escondida (n_1) e a segunda inserindo-se mais uma camada (n_2). O número de neurônios iniciais para cada camada escondida foram definidas em partes atendendo os valores mínimos proposto por Fletcher e Gloss (1993) e apresentados na Tabela 4. O número máximo de épocas foi fixado em mil (1.000) interações e a função de ativação Tangente Hiperbólica (*tansig*) definida conforme as características dos dados de entrada. A parametrização da primeira fase foi idêntica à segunda fase. Nas duas fases são realizados dez (10) treinamentos para cada combinação (Tabela 3) testando todas as topologias da Tabela 4. O critério de parada do treinamento é a validação cruzada (70% treinamento, 15% para validação e 15% para teste).

Tabela 4 – Intervalo de valores na parametrização da RNA/MLP

RNA/MLP	1	2	3	4	5	6
n° de neurônios (n_1 e n_2)	5	10	20	40	80	160

Em cada interações (épocas) os erros (EQM – Erro quadrático médio) entre os valores estimados e medidos da validação são avaliados e se forem diminuindo significa que a rede está aprendendo até que este erro comece a aumentar durante uma sequência de seis (6) interações (épocas) sucessivas. Caso este erro da validação no treinamento aumentar enquanto do treinamento continuar caindo é provável que a rede esteja com *overfitting*, condição que indica a parada do treinamento. Este critério minimiza o risco de *underfitting* da rede. Sendo definido a melhor rede em cada treinamento (menor EQM do treinamento) todas as combinações são validadas com os dados reorganizados (AT e AAT) em cada fase e escolhida a fase e a combinação que apresentar o melhor desempenho.

Indicadores de Validação

São utilizados o EQM - erro quadrático médio (*Mean Square Error*) que indica o erro médio entre os valores estimados e medidos, sendo os melhores resultados aquele com desvios menores mais próximos dos valores desejados. O indicador RMSE, raiz quadrada do erro quadrático médio (*Root Mean Square Error*) permite verificar o valor do erro produzido pelo modelo, sendo que valores menores, indicam um desempenho satisfatório para os modelos gerados. O R^2 , coeficiente de determinação, permite verificar o quanto a variação da variável dependente é explicada pela variação da variável independente através da equação do modelo, e r , coeficiente de correlação que permite

verificar o quanto o valor medido explica o valor estimado, ou seja, a correlação existente entre eles (Willmott, 1981; Stone, 1993; Escobedo et al. 2011).

Os indicadores são obtidos pelas seguintes relações:

$EQM = \left[\frac{\sum_{i=1}^N (Ei - Mi)^2}{N} \right]$	$RMSE = 100 * \frac{\left[\frac{\sum_{i=1}^N (Ei - Mi)^2}{N} \right]^{1/2}}{\bar{M}}$
$R^2 = \left(\frac{\sum_{i=1}^N (Mi - \bar{M}i)^2 * \sum_{i=1}^N (Ei - \bar{E}i)^2}{\sqrt{\sum_{i=1}^N (Mi - \bar{M}i)^2} * \sqrt{\sum_{i=1}^N (Ei - \bar{E}i)^2}} \right)^2$	$r = \left(\frac{\sum_{i=1}^N (Mi - \bar{M}i) * \sum_{i=1}^N (Ei - \bar{E}i)}{\sqrt{\sum_{i=1}^N (Mi - \bar{M}i)^2} * \sqrt{\sum_{i=1}^N (Ei - \bar{E}i)^2}} \right)$

onde Ei são os valores estimados, $\bar{E}i$, é a média dos valores estimados, Mi são os valores medidos, $\bar{M}i$ é a média dos valores medidos, N é o número de observações.

RESULTADOS E DISCUSSÕES

São apresentados os resultados obtidos para estimativa de K_d utilizando modelagem estatística com uma variável e na sequência o resultado obtido por meio das Rede Neural/MLP com variáveis climatológicas de referência e na sequência uma análise de desempenho dos modelos estatístico e da rede. Ao final os indicadores são comparados com trabalhos disponíveis na literatura.

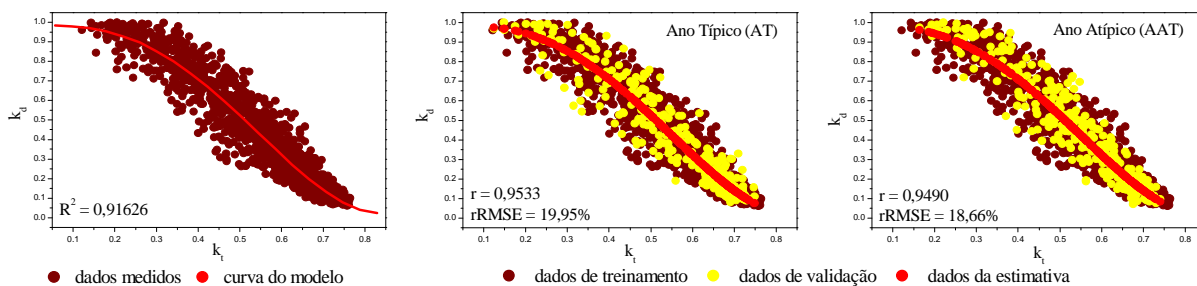
Modelo Estatístico

O modelo estatístico elaborado da correlação de K_d em função de K_T (Figura 2.a) é similar as curvas encontradas na literatura, assim como as equações de estimativa (Equação 1) ajustada por regressão polinomial de 4ª ordem e com coeficiente de determinação na ordem $R^2 = 0.9162$. Outras equações de ordens polinomiais foram elaboradas, mas apresentando coeficiente de determinação menores.

$$k_d = 0.981 + 0.0894 k_t - 0.462 k_t^2 - 5.881 k_t^3 + 5.585 k_t^4 \dots \dots \dots (1)$$

São apresentados na Figura 2 (b) e (c) as curvas de dispersão geradas na validação do modelo ME para Ano típico (AT) e atípico (AAT), com os valores dos dados medidos (marrom), dos dados de validação (amarelo) e da estimativa (vermelho).

Figura 2 - a) Curva de dispersão ME e validação para: b) AT; c) AAT



Os indicadores estatísticos RMSE e o r para Ano Típico (AT) e Atípico (AAT) resultantes da análise na validação dos modelos ME são apresentados na Tabela 5. O modelo ME na base AT apresentou uma correlação de ($r = 0.953$) e um índice de desempenho de ($RMSE = 19.95\%$) assim como na base de validação AAT, que também apresentou uma correlação de ($r = 0.949$) e um índice de desempenho de ($RMSE = 18.66\%$), mostrando que ambos os modelos podem ser utilizados na estimativa da fração da irradiação difusa na partição diária, considerando a variável disponível na maioria das estações meteorológicas e próximo aos valores encontrados na literatura em outras localidades (Tabela 5).

Tabela 5 – Índices estatísticos - outras localidades e ME para Botucatu

Autores	Localidade	r	rRMSE %
Khorasanizadeh et al. 2016	Irã	0.8767	
Bortolini et al. 2013	Madri	0.9628	
Shamshirband et al. 2016	Irã	0.8156	
Oliveira et al. 2002	São Paulo	0.90	
ME AT – Proposto 2017	Botucatu/São Paulo	0.9533	19.95
ME AAT – Proposto 2017	Botucatu/São Paulo	0.9490	18.66

Rede Neural Artificial MLP

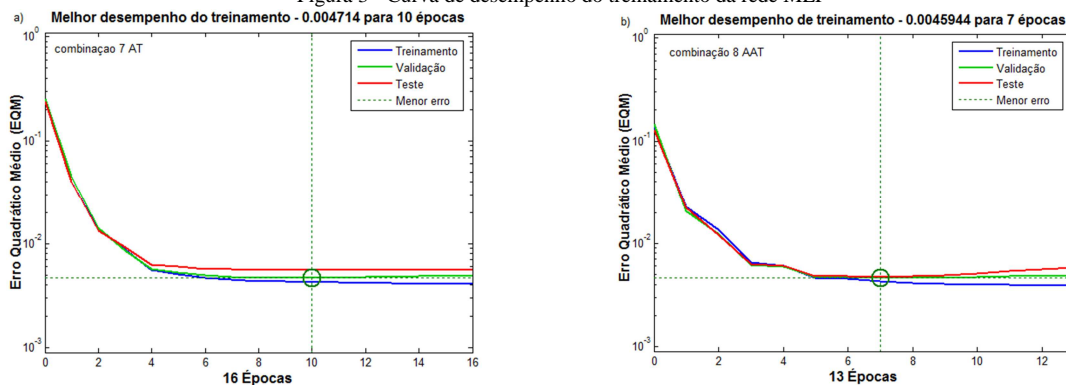
Os indicadores das combinações da técnica RNA/MLP são apresentados na Tabela 6 com os melhores resultados experimentados na parametrização e seus indicadores estatísticos após treinamento e validação realizados pelas redes MLP. Observa-se nos indicadores de (K_d) que o índice de correlação (r) entre os valores medidos e estimados nas duas bases de validação variou entre 0,9531 a 0,9649 mostrando a princípio uma boa correlação. Os índices de desempenho (RMSE) da RNA/MLP variou entre os valores 15,56% a 19,98% entre todas as combinações. Observa-se que a rede apresentou os melhores resultados de validação na base AT quando sete variáveis de entrada (combinação 7) participaram do treinamento, e quando foram utilizadas duas (2) camadas escondidas. Porém após os dez (10) treinamentos realizados a melhor combinação não foi a que apresentou o menor EQM de treinamento.

Tabela 6 - - Indicadores e configurações do ME e MLP

Modelo Estatístico			Indicadores Estatísticos											
			Para Ano típico (AT)					Para Ano típico (AAT)						
Variáveis de:			Validação					Validação						
Entrada	Saida													
Combinação 1	K_t	K_d												
							19.95	0.9533				18.66	0.9490	
Rede MLP			Para Ano típico (AT)					Para Ano típico (AAT)						
Variáveis de:			Treinamento					Validação						
Entrada	Saida		Épocas	MSE	RMSE (%)	MSE	RMSE (%)	r	Épocas	MSE	RMSE (%)	MSE	RMSE (%)	r
Combinação 1	K_t		19	0.0058	19.29	0.0066	19.98	0.9531	5	0.0057	19.56	0.0067	18.30	0.9510
Combinação 2	K_t, H_0		4	0.0054	18.68	0.0066	19.87	0.9537	7	0.0054	18.99	0.0065	18.01	0.9526
Combinação 3	K_t, H_0, H_2		4	0.0054	18.71	0.0065	19.77	0.9541	4	0.0054	19.03	0.0064	17.92	0.9531
Combinação 4	K_t, H_0, H_2, H_{va}		4	0.0054	18.66	0.0064	19.66	0.9547	8	0.0054	18.91	0.0063	17.79	0.9538
Combinação 5	$K_t, H_0, H_{gr}, H_{va}, d_{ccsol}$	K_d	7	0.0054	18.60	0.0063	19.44	0.9557	5	0.0053	18.78	0.0067	18.28	0.9511
Combinação 6	$K_t, H_0, H_2, H_{va}, d_{ccsol}, n_{debul}$		43	0.0043	16.60	0.0052	17.70	0.9634	10	0.0045	17.25	0.0055	16.59	0.9599
Combinação 7	$K_t, H_0, H_2, H_{va}, d_{ccsol}, n_{debul}, u_{mid}$		10	0.0045	17.12	0.0050	17.34	0.9649	17	0.0044	17.02	0.0056	16.65	0.9596
Combinação 8	$K_t, H_0, H_{gr}, H_{va}, d_{ccsol}, n_{debul}, u_{mid}, k_{ar}$		7	0.0039	15.91	0.0051	17.48	0.9643	7	0.0044	17.10	0.0048	15.56	0.9649

Os gráficos da Figura 3 (a) e (b) demonstram o comportamento das curvas do EQM e o número de épocas utilizadas na melhor combinação da validação para base AT e AAT de treinamento da rede MLP. Em destaque (linha tracejada e círculo verde) o ponto com o menor erro após o treinamento. O amortecimento suave das curvas e a quantidade de épocas caracterizam um treinamento com pouca dificuldade e baixa complexidade de aprendizagem. Estas características devem-se principalmente pela comportamento numérico dos pontos e das variáveis de entrada.

Figura 3 - Curva de desempenho do treinamento da rede MLP

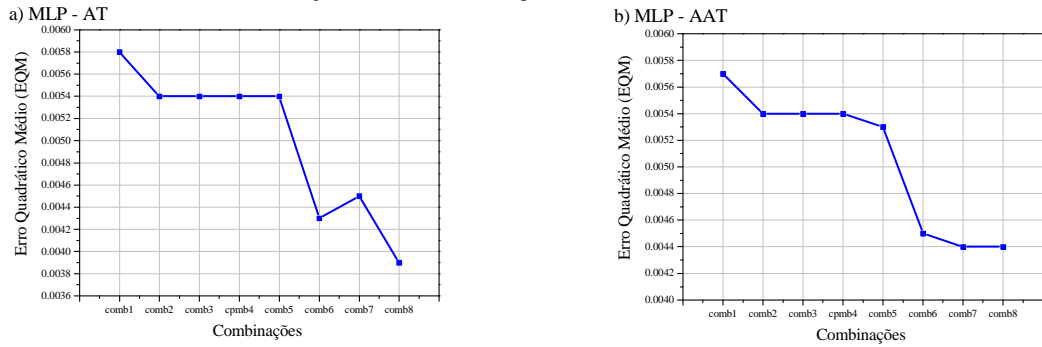


Através dos gráficos da Figura 4 (a) e (b) é possível verificar a importância de outras variáveis de entrada no treinamento tendo em vista que com a inserção de outras variáveis no treinamento das redes a uma diminuição do EQM e conseqüentemente um melhor desempenho no treinamento. Outras parametrizações de ajustamento foram testadas mas não resultaram em melhoria dos indicadores.

Analisando os indicadores da Tabela 6 a RNA/MLP apresenta um desempenho melhor que (ME). Em média uma redução de aproximadamente 17,5%. A inserção de outras variáveis no treinamento contribuí para um melhor desempenho conforme pode ser observado nos gráficos da Figura 4 (a) e (b) com uma diminuição do EQM de treinamento.

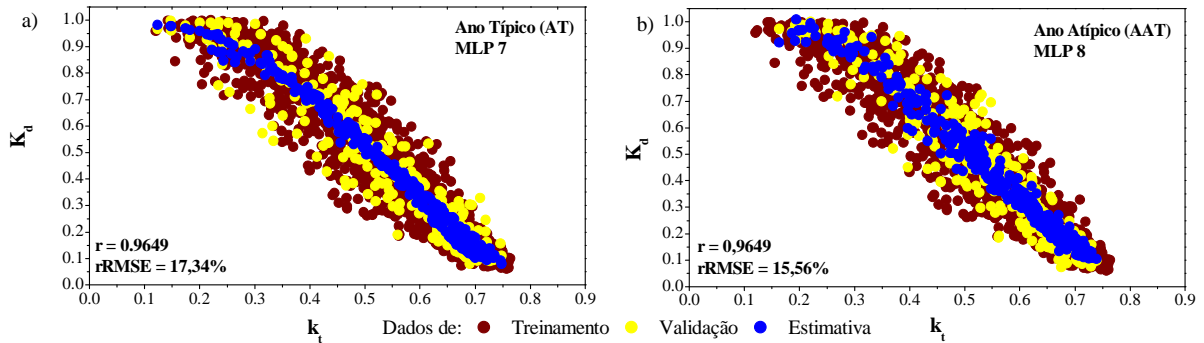
Para um melhor entendimento de exatidão entre os valores medidos e estimados, a Figura 5 (a), (b) mostra a dispersão apresentada com a RNA/MLP e sua melhor combinação para estimativa da fração da irradiação difusa nas duas bases AT e AAT. Importante ressaltar que os resultados encontrados refere-se às combinações proposta e que a alteração dos parâmetros de entrada e da quantidade podem influenciar na configuração das parametrizações e dos indicadores de desempenho da rede.

Figura 4 - Curva de desempenho do treinamento



Dependendo do algoritmo de treinamento parâmetros como a taxa para aprendizagem e do termo momentum que visam acelerar a convergência ao menor erro entre o valor medido e estimado. Os resultados encontrados são para as combinações e variáveis da Tabela 3 sendo que outras combinações podem ser testadas verificando o comportamento e desempenho das técnicas em seus treinamentos. Com estas mudanças também serão necessários outras parametrizações como na escolha do número de camadas, de neurônios e da função de ativação sendo também possível avaliar se a inserção de cada variável irá impactar no desempenho da rede com valores mais significativos para estimativa de k_d .

Figura 5 - Curva de dispersão de K_d x K_T - melhores combinações de AT e AAT



São apresentados na Tabela 7 alguns indicadores do coeficiente de correlação (r) e/ou do índice de desempenho (RMSE) de alguns trabalhos da literatura que utiliza a RNA/MLP e o melhor resultado da rede para partição diária encontrado neste trabalho.

Tabela 7 – Índices estatísticos - outras localidade e TAM para Boutcatu

Autores	Localidade	Técnicas	Variáveis de entrada	r	RMSE
Elminir et. al (2007)	Cairo/Egito	Perceptron	Radiação global e extraterrestre e brilho solar.		2.52 – 5.96
Xue (2017)	Chengdu/China	Perceptron	Mês do ano, duração da luz do sol, temperatura média, precipitação, velocidade do vento, umidade relativa e radiação solar global diária	0.9340	
Ihya et. al (2014)	Marrocos	Perceptron	Fração da irradiação global e fotoperíodo	0.9596	18.73
Jiang (2008)	Zhengzhou, China	Perceptron	Índice de claridade	0.9591	
Modelo MLP (AAT)	Botucatu/SP	MLP	Fração da irradiação global, Radiação no topo da atmosfera, Global e de Ondas longas, declinação solar, nebulosidade, umidade e temperatura.	0.9649	15.56

Com os resultados observa-se que a rede MLP proporcionou um melhor desempenho na estimativa da fração da irradiação difusa (k_d) com o uso de variáveis climatológicas e meteorológicas, como também melhor que o modelo estatístico.

CONCLUSÕES

Os resultados apresentados nos mostra que o modelo estatístico pode ser utilizado para estimativa da fração transmitida da irradiação solar difusa (k_d) com uma equação de estimativa similar aos apresentados na literatura com um coeficiente de determinação próximo ($R^2 = 0.9163$) mesmo com as diferenças das variações astronômicas e climáticas de cada localidade. O valor encontrado do coeficiente de determinação R^2 (próximo a 1) denota que existe uma elevada correlação entre a fração da irradiação solar difusa e irradiação solar global. Com a utilização RNA/MLP observou-se uma melhora em comparação ao modelo estatístico. Apesar dos modelos estatísticos apresentados na literatura terem apenas uma variável dependente (k_T) não se pode afirmar que apenas esta variável

influencia a estimativa de k_d . As diversas condições atmosféricas, geográficas e meteorológicas que acontecem ao longo do dia como a nebulosidade, a umidade, a temperatura, além de outros constituintes com poeiras e partículas em suspensão na atmosfera, podem também influenciar a variabilidade de k_d . O resultado mostrou que existe a possibilidade de se obter indicadores melhores com mais de uma variável, e que kd também está fortemente correlacionado com outras variáveis climatológicas. Por fim, outro fato relevante é que os resultados podem influenciar diretamente os estudos e os projetos, como por exemplo nos investimentos de uma nova planta fotovoltaica, no desenvolvimento de um novo cultivar e no controle de constituintes atmosféricos.

REFERÊNCIAS

- Alam, S., Kaushik, S. C.; Garg, S. N.. Computation of beam solar radiation at normal incidence using artificial neural network. *Renewable Energy*. v.31, p.1483–1491, 2006.
- Alam, S.; Kaushik, S.C.; Garg, S. N.. Assessment of diffuse solar energy under general sky condition using artificial neural network. *Applied Energy*, vol. 86, p. 554–564, 2009.
- Bakirci, K.. Models for the estimation of diffuse solar radiation for typical cities in Turkey. *Energy*. vol. 82, p. 827 – 838, 2015.
- Boland, J.; Scott, L.; Luther, M.. Modelling the diffuse fraction of global solar radiation on a horizontal surface. *Environmetrics*, vol. 12, p. 103 –117, 2001.
- Boznar, M. Z.; Bostijan, G.; Oliveira, A. P.; Soares, J.; Mlakar, P.. Spatially transferable regional model for half-hourly values of diffuse solar radiation for general sky conditions based on perceptron artificial neural networks. *Renewable Energy*, vol. 103, p. 794 – 810, 2017.
- Çelik, O.; Teke, A.; Bas, Yildirim, H. B.. The optimized artificial neural network model with LevenbergeMarquardt algorithm for global solar radiation estimation in Eastern Mediterranean. *Journal of Cleaner Production*, vol. 116, pp. 1 -12, 2016.
- Chandrasekaran, J.; Kumar, S.. Hourly diffuse fraction correlation at a tropical location. *Solar Energy*, vol. 53, no 6, p. 505 – 510, 1994.
- Chaves, M., Escobedo, J. F.. A software to process daily solar radiation data, *Renewable Energy*, vol.19(1), pp. 339 – 344, 2000.
- Xue, Xinhua. Prediction of daily diffuse solar radiation using artificial neural networks. *International Journal of Hydrogen Energy*. vol. 42, p. 28214 – 28221, 2017.
- Codato, G.; Oliveira, A. P.; Soares, J.; Escobedo, J. F.; Gomes, E. N.; Pai, A. D.. Global and diffuse solar irradiances in urban and rural areas in southeast Brazil. *Theoretical and Applied Climatology*, vol. 93, pp. 57–73, 2008.
- Collares-Pereira, M.; Rabl, A.. The average distribution on solar radiation – correlations between diffuse and hemispherical and between daily and hourly insolation values. *Solar Energy*, vol. 22, p. 155 - 164, 1979.
- Dal Pai, A.; Escobedo, J. F.; Dal Pai, E.; Oliveira, A. P.; Soares, J. R.; Codato, G.. MEO shadowing method for measuring diffuse solar irradiance: corrections based on sky cover. *Renewable Energy*, vol. 99, p.754 - 76, 2016.
- De Miguel, A.; Bilbao, J.; Aguiar, R.; Kambezidis, H.; Negro, E.. Diffuse solar irradiation model evaluation in the north Mediterranean Belt Area. *Solar Energy*, vol. 70, n. 2, p. 143 – 153, 2001.
- Drechmer, P. A. O.; Ricieri, R. P.. Irradiação global, direta e difusa, para a região d Irradiação global, direta e difusa, para a região de para a região de Cascavel, Estado e Cascavel, Estado do Paraná do Paraná. *Acta Scientiarum Technology*, vol. 28, n. 1 pp. 73 – 77, 2006.
- Drummond, A. J., 1956, *Arch. Met. Geoph. B. Serie B*. vol. 7, Issue 3 - 4, p. 413 – 436, 1956.
- Elizondo, D.; Hoogenboom, G.; McClendon, R. W.. Development of a neural network model to predict daily solar radiation. *Agricultural and Forest Meteorology*, vol. 71, p. 115 - 132, 1994.
- Elminir, H. K.; Azzam, Y. A.; Younes, F. I.. Prediction of hourly and daily diffuse fraction using neural network, as compared to linear regression models. *Energy*, vol. 32, p. 1513 – 1523, 2007.
- Escobedo, J. F.; Dal Pai, A.; Lenz, L.C.; Prado, M. B.. Efeito da cobertura do céu no aumento das frações solar difusa e par-difusa em estufa de polietileno. *Avances en Energías Renovables y Medio Ambiente*, Vol. 17, pp.02.19-02.25, 2013.
- Escobedo, J. F.; Gomes, E. N.; Oliveira, A. P.; Soares, J.. Ratios of UV, PAR and NIR components to global solar radiation measured at Botucatu site in Brazil. *Renewable Energy*, v.36, p.169-178, 2011.
- Fletcher D.S., Goss, E. . Forecasting with neural network: an application using bankruptcy data. *Information Management*, vol.24, p.159–67, 1993.
- Furlan, C.; Oliveria, A. P.; Soares, J.; Codato, G.; Escobedo, J. F.. The role of clouds in improving the regression model for hourly values of diffuse solar radiation. *Applied Energy*, vol. 92, p. 240 – 254, 2012.
- Gueymard, C.. REST2: high-performance solar radiation model for cloudless sky irradiance, illuminance, and photosynthetically active radiation e validation with a benchmark dataset, *Solar Energy*, vol. 82, pp. 272 – 285, 2008.
- Haykin, S. *Redes neurais: princípios e prática*. 2.ed. Porto Alegre, Bookman, 2001.

- Ihya, B.; Mechaqrane, A.; Tadili, R.; Bargach, M. N.. Prediction of hourly and daily diffuse solar fraction in the city of Fez (Morocco). *Theor Appl Climatol*, vol. 120, pp.737 – 749, 2014.
- Jacovides, C.P.; Tymvios, F. S.; Assimakopoulos, V. D.; Kaltsounides, N. A.. Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation. *Renewable Energy*, vol. 31, p. 2492 – 2504, 2006.
- Jamil, B.; Akhtar, N.. Estimation of diffuse solar radiation in humid-subtropical climatic region of India: Comparison of diffuse fraction and diffusion coefficient models. *Energy*, vol. 131, p. 149 – 164, 2017.
- Kalogirou, S. Applications of artificial neural-networks for energy systems. *Applied Energy*, vol. 67, p. 17 – 35, 2000.
- Kaushika, N.D.; Tomar, R.K.;Kaushik, S.C..Artificial neural network model based on interrelationship of direct, diffuse and global solar radiations. *Solar Energy*, vol. 103, pp. 327–342, 2014.
- Khatib, T.; Mohamed, A.; Sopian, K.; Mahmoud, M.. Assessment of Artificial Neural Networks for Hourly Solar Radiation Prediction. *International Journal of Photoenergy*, vol. 2012, p. 1 - 8, 2012.
- Khorasanizadeh, H.; Mohammadi, K.; Goudarzi, N.. Prediction of horizontal diffuse solar radiation using clearness index based empirical models: A case study. *International Journal of Hydrogen Energy*, vol. 41, p. 21888 – 21898, 2016.
- 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*, vol. 52, p. 1294 – 1299, 2015.
- Kuo, C.W., Chang, W.C., Chang, K.C.. Modeling the hourly solar diffuse fraction in Taiwan. *Renewable Energy* 66, 56–61. 2014
- Liu, B.H.Y.; Jordan, R. C.. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar Energy*, vol. 4, pp. 1, 1960.
- Lopez, G.L.; Muneer, T.; Claywell, R.. Assessment of four shadow band correction models using beam normal irradiance data from the United Kingdom and Israel. *Energy Conversion and Management*, vol. 45, p.1963 – 1979, 2004.
- Marques Filho, E. P.; Oliveira, A. P.; Willian A. Vita, W. A.;Mesquita, F. L. L.;Codato, G.; Escobedo, J. F.;Cassol, M.;França, J. R. A..Global, diffuse and direct solar radiation at the surface in the city of Rio de Janeiro: Observational characterization and empirical modeling. *Renewable Energy*, vol. 91, p. 64 – 74, 2016.
- Martinez, C.; Cúnsulo, M.. Seguidor automático para bloqueador de radiación solar. *Avances en Energías Renovables y Medio Ambiente*, 2017.
- Marzouq, M.; Fadili, H. E.; Lakhliai, Z.; Zenkouar, K.. A review of solar radiation prediction using artificial neural networks. *International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS)*. vol. 5 p. 325 – 345, 2017.
- Melo, J. M. D.; Escobedo, J. F.. Medida da radiação solar difusa. In: *Energias Limpas em progresso. VII Congresso Ibérico de Energia Solar*,Vigo, Espanha. *Anais International Solar Energy Society*, vol. 1, 1994.
- Miguel de Simón-Martín, M.; Díez-Mediavilla, M.; Alonso-Tristán, C.. Shadow-band radiometer measurement of diffuse solar irradiance: Calculation of geometrical and total correction factors. *Solar Energy*, vol. 139, p. 85 – 99, 2016.
- Moustris, K.; Paliatsos, A. G.; Bloutsos, A.; Nikolaidis, K.; Koronaki, I.; Kavadias, K.. Use of neural networks for the creation of hourly global and diffuse solar irradiance data at representative locations in Greece. *Renewable Energy*, vol. 33, p 928 – 932, 2008.
- Oliveira, A. P.; Escobedo, J. F.; Machado, A. J.; Soares, J.. Correlation models of diffuse solar-radiation applied to the city of São Paulo, Brazil. *Applied Energy*, vol. 71, p. 59 – 73, 2002.
- Orgill, J. F.; Hollands, G.T.. Correlation equation for hourly diffuse radiation on a horizontal surface. *Solar Energy*, vol. 9, p. 357 – 359, 1977.
- Pashiardis, S.; Kalogirou, S.A.. Quality control of solar shortwave and terrestrial longwave radiation for surface radiation measurements at two sites in Cyprus. *Renewable Energy*, vol. 96, p. 1015 – 1033, 2016.
- Paulescu, E.; Blaga, R.. Regression models for hourly diffuse solar radiation. *Solar Energy*, vol. 125, p. 111 – 124, 2016.
- Reindl, D. T.; Beckman W. A.; Duffie, J. A.. Diffuse fraction correlations. *Solar Energy*, v.45, p.1 – 7, 1990.
- Ridley, B.; Boland, J.; Lauret, P..Modelling of diffuse solar fraction with multiple predictors. *Renewable Energy*, vol. 35, p. 478–483, 2010.
- Robinson, H., Stoch, L. Sky radiation and measurements and corrections. *Journal of Applied Meteorology*, v.3, pp.179-81, 1964.
- Ruiz-Arias, J.A.; Alsamamra, H.; Tovar-Pescador, J.; Pozo-Vázquez, D.. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions.. *Energy Conversion and Management*, vol. 51, p. 881 – 893, 2010.
- Shamshirband, S.;Mohammadi, K.; Khorasanizadeh, H.; LipYee, P.; Lee, M.; Petković, D.; Zalnezhad, E.. Estimating the diffuse solar radiation using a coupled support vector machine–wavelet transform model. *Renewable and Sustainable Energy Reviews*, vol. 56, p.428–435, 2016.

- Silva, F. R., Oliveira, H. S. M., Marinho, G. S.. Análise das componentes global e difusa da radiação solar em Natal-RN entre 2007 e 2008. II Congresso Brasileiro de Energia Solar e III Conferência Regional Latino-Americana da ISES - Florianópolis, 18 a 21 de novembro de 2008
- 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*, vol. 79, p. 201-214, 2004.
- Stone, R.J.. Improved statistical procedure for the evaluation of solar radiation estimation models. *Solar Energy*, vol. 51, n 04, p. 289 – 291, 1993.
- Teke, A.; Yildirim, H. B.; Celik, O.. Evaluation and performance comparison of diferente models for the estimation of solar radiation. *Renewable and Sustainable Energy Reviews*, vol. 50, p. 1097 – 1107, 2015.
- Teramoto, E. T.; Escobedo, J. F.. Análise da frequência anual das condições de céu em Botucatu, São Paulo. *Revista Brasileira Engenharia Agrícola Ambiental*. v.16, n. 9, p.985-992, 2012.
- Tomar, R. K.; Kaushika, N. D., Kaushik, S. C. Artificial neural network based computational model for the prediction of direct solar radiation in Indian zone. *Journal of Renewable and Sustainable Energy*. Vol. 4, p.063 – 146, 2012.
- Tymvios, F. S.; Jacovides, C.P.; Michaelides, S.C.; Scouteli, C.. Comparative study of Ångström's and artificial neural networks' methodologies in estimating global solar radiation. *Solar Energy*, vol. 78, p. 752-762, 2005.
- Willmott, C. J. On the validation of models. *Physical Geography*, vol. 2, p. 184-194, 1982.
- WMO – World Meteorological Organization. Guide to meteorological Instruments and Methods of Observation. WMO-nº8, Seventh Edition, p.1 - 681, Geneva, Switzerland, 2008.
- WMO - 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, p. 298, 1981.
- Younes, S.; Claywell, R.; Muneer, T.. Quality control of solar radiation data: Present status and proposed new approaches. *Energy*, vol. 30, p. 1533 – 1549, 2005.

Multilayer Perceptron Network (MLP) in the estimation of the fraction of diffuse solar irradiation in the daily partition

ABSTRACT: This work proposes the use of Artificial Neural Networks (ANN / MLP) in the estimation of the diffuse fraction (K_d) of solar irradiation from a set of combinations of climatological variables. The data are obtained from 2000 to 2006 at the Meteorological Station of Universidade Estadual Paulista, Campus Botucatu / SP. To verify the performance of the techniques, a statistical model (ME) using only the transmitted fraction of the global solar irradiation (K_T) as reference variable was elaborated. The input variables for ANN / MLP training comprise measurements / calculated from eight astronomical and meteorological variables. For validation the total data base is reorganized forming two bases called the Typical Year (AT) and Atypical Year (AAT). Performance is evaluated by the correlation coefficient (r) and the precision index (RMSE). The statistical model (ME) developed and adjusted by the 4th order polynomial regression presents a coefficient of determination (R^2) of 0.91626 which shows that K_d is well correlated with K_T statistically. The performance index of the statistical model (ME) in the validation with the two bases (AT and AAT) was $RMSE_{AT} = 19.95\%$ and $RMSE_{AAT} = 18.66\%$, which shows a good approximation of the estimated values and the correlation of $r_{AT} = 0.9533$ and $r_{AAT} = 0.9490$ which shows the importance of the independent variables for dependent variable. With the (ANN / MLP) the performance indicators (RMSE) presented a reduction in relation to the (ME) in both validation bases in approximately 20%, which shows that the insertion of variables contributes in the estimates of (K_d) with the climatic and geographical conditions used.

Keywords: Solar radiation; Diffuse radiation; Multilayer Perceptron Network.