

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”

FACULDADE DE CIÊNCIAS AGRONÔMICAS

CAMPUS DE BOTUCATU

**ESTIMATIVA DA RADIAÇÃO SOLAR ULTRAVIOLETA EM
BOTUCATU/SP/BRASIL UTILIZANDO TÉCNICAS DE
APRENDIZADO DE MÁQUINA**

THIAGO DO NASCIMENTO SANTANA DE ALMEIDA

Dissertação apresentada à Faculdade de Ciências Agronômicas da Unesp – Campus de Botucatu, para a obtenção do título de Mestre em Agronomia (Energia na Agricultura).

BOTUCATU-SP

Fevereiro – 2013

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”

FACULDADE DE CIÊNCIAS AGRONÔMICAS

CAMPUS DE BOTUCATU

**ESTIMATIVA DA RADIAÇÃO SOLAR ULTRAVIOLETA EM
BOTUCATU/SP/BRASIL UTILIZANDO TÉCNICAS DE
APRENDIZADO DE MÁQUINA**

THIAGO DO NASCIMENTO SANTANA DE ALMEIDA

Orientador: Prof. Dr. João Francisco Escobedo

Dissertação apresentada à Faculdade de Ciências Agronômicas da Unesp – Campus de Botucatu, para a obtenção do título de Mestre em Agronomia (Energia na Agricultura).

BOTUCATU-SP

Fevereiro – 2013

FICHA CATALOGRÁFICA

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉCNICA DE AQUISIÇÃO E TRATAMENTO DA INFORMAÇÃO - SERVIÇO TÉCNICO DE BIBLIOTECA E DOCUMENTAÇÃO - UNESP - FCA - LAGEADO - BOTUCATU (SP)

A447e Almeida, Thiago do Nascimento Santana de, 1985-
Estimativa da radiação solar ultravioleta em Botucatu/SP/Brasil utilizando técnicas de aprendizado de máquina / Thiago do Nascimento Santana de Almeida. - Botucatu : [s.n.], 2013
viii, 57 f. : il. color., gráfs., tabs., fots.

Dissertação (Mestrado) - Universidade Estadual Paulista, Faculdade de Ciências Agronômicas, Botucatu, 2013
Orientador: João Francisco Escobedo
Inclui bibliografia

1. Máquina de vetores de suporte. 2. Método da função de base radial. 3. Modelos matemáticos. 4. Radiação solar. 5. Redes neurais (Computação). I. Escobedo, João Francisco. II. Universidade Estadual Paulista "Júlio de Mesquita Filho" (Campus de Botucatu). Faculdade de Ciências Agronômicas. III. Título.

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
FACULDADE DE CIÊNCIAS AGRONÔMICAS
CAMPUS DE BOTUCATU

CERTIFICADO DE APROVAÇÃO

TÍTULO: "ESTIMATIVA DA RADIAÇÃO SOLAR ULTRAVIOLETA EM
BOTUCATU/SP/BRASIL UTILIZANDO TÉCNICAS DE APRENDIZADO
DE MÁQUINA"

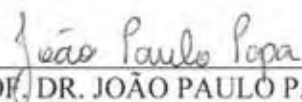
ALUNO: THIAGO DO NASCIMENTO SANTANA DE ALMEIDA

ORIENTADOR: PROF. DR. JOÃO FRANCISCO ESCOBEDO

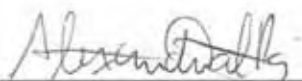
Aprovado pela Comissão Examinadora



PROF. DR. JOÃO FRANCISCO ESCOBEDO



PROF. DR. JOÃO PAULO PAPA



PROF. DR. ALEXANDRE DAL PAI

Data da Realização: 13 de fevereiro de 2013

AGRADECIMENTOS

Agradeço a Deus, a meus pais Adil e Dijanira e a meus irmãos Igor e Viviane, que me proporcionaram condições para trilhar e vencer essa jornada.

À minha namorada Andreia por me apoiar sempre que precisei;

E principalmente aos meus amigos por estarem sempre ao meu lado me auxiliando e ensinando tudo o que foi necessário.

DEDICATÓRIA

Inicialmente dedico este trabalho ao prof. Alexandre Dal Pai, meu orientador da graduação e hoje um amigo que confiou em mim, me indicou para ao seu orientador e me enfiou nessa roubada.

Ao prof. Escobedo por me orientar e me ensinar, não somente como ser um aluno de mestrado, mas também suas experiências e as mais inimagináveis histórias vividas por ele... Repetidas vezes...

Ao prof. Eduardo que me ensinou praticamente tudo que sei sobre radiação solar, origin, whisky, músicos a pé... ...

E principalmente aos meus amigos, Douglas que me deu uma mão absurda na hora de ensinar a máquina a aprender, Ludmila minha companheira de aventuras, Érico que me mostrou o que eu precisava escrever exatamente do modo que eu não deveria escrever, Enzo por nos mostrar o caminho da paz e também aos demais Escobedo's Rangers, Cícero Fábio e Ricardo.

Agradeço a todos vocês que estavam sempre ao meu lado pra me auxiliar em tudo que precisei e por me atrapalhar nos poucos momentos que precisava de concentração propondo jogos, bebedeiras, baladas, ouvindo músicas, mostrando promoções na internet, estourando pipoca, contando piadas, compartilhando fotos inadequadas para menores...

...sem vocês isso jamais seria tão bom quanto foi!

go go **ESCOBEDO RANGERS** !

SUMÁRIO

RESUMO	1
SUMMARY	3
1 INTRODUÇÃO.....	4
2 REVISÃO BIBLIOGRÁFICA.....	7
2.1 Radiação Solar Ultravioleta.....	7
2.1.1 Modelos Estatísticos de Estimativa da UV.....	9
2.2 Redes Neurais Artificiais.....	10
2.2.1 Histórico	11
2.2.2 Perceptron	12
2.2.3 Perceptron de Múltiplas Camadas.....	13
2.2.4 Algoritmos de aprendizagem	14
2.3 Máquina de Vetores de Suporte.....	15
2.4 Função de Base Radial.....	18
2.5 Trabalhos Relacionados.....	21
3 MATERIAL E MÉTODOS.....	27
3.1 Localização e Clima.....	27
3.2 Aquisição dos Dados.....	28
3.3 Tratamentos dos Dados.....	30
3.4 Composição dos Agrupamentos.....	31
3.5 Técnicas de Aprendizado de Máquina Utilizadas.....	35
3.6 Utilização das Técnicas de Aprendizado de Máquina.....	35
3.7 Tipo de Aprendizado.....	37
4 RESULTADOS E DISCUSSÃO.....	38
4.1 Comparação com a Literatura.....	46
4.1.1 Trabalhos que utilizaram TAM.....	46
4.1.2 Trabalhos que utilizaram modelos estatísticos para Botucatu	47
5 CONCLUSÃO.....	48
6 REFERÊNCIAS.....	50

ÍNDICE DE FIGURAS

Figura 1. Espectro Solar.	8
Figura 2. Neurônio biológico e Neurônio artificial.	11
Figura 3. Perceptron.	12
Figura 4. Exemplo de uma RNA com 4 entradas 2 camadas ocultas e 2 saídas.	13
Figura 5. Funcionamento da SVM.	15
Figura 6. Aumento de dimensionalidade da SVM.	16
Figura 7. Simulação de uma SVM no modelo de neurônios de uma RNA.	17
Figura 8. Função Multiquadrática.	19
Figura 9. Função Multiquadrática Recíproca.	19
Figura 10. Função Multiquadrática Recíproca Inversa.	19
Figura 11. Função Gaussiana.	20
Figura 12. Função Secante Hiperbólica.	20
Figura 13: Médias diárias mensais dos parâmetros meteorológicos.	28
Figura 14. Estação de Radiometria Solar de Botucatu.	30
Figura 15: Planilha de dados para inserção na FannTool.	35
Figura 16: Planilha de dados para inserção na LibSVM.	36
Figura 17. Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 1. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.	39
Figura 18. Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 2. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.	41
Figura 19. Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 3. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.	42
Figura 20: Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 4. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.	43
Figura 21. Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 5. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.	44

ÍNDICE DE TABELAS

Tabela 1: Validação dos modelos propostos por Leal et al. e Barbero et al.	22
Tabela 2: Validação dos modelos propostos por Erusiafe & Chendo; AbdulAzeez;	23
Tabela 3: Validação dos modelos gerados por CHENG et al.	24
Tabela 4: Validação dos modelos propostos por Lopes et al.	25
Tabela 5: Validação dos modelos propostos por Soares et al.	26
Tabela 6: Equipamentos agrometeorológicos utilizados e suas características.....	29
Tabela 7: Característica detalhada dos equipamentos utilizados para medição de radiação solar.....	29
Tabela 8: Dias retirados da base de dados.....	31
Tabela 9: Grupos de variáveis utilizadas como entrada para as técnicas de aprendizado. ..	32
Tabela 10: Avaliação dos modelos gerados.	45
Tabela 11: Modelos estatísticos gerados para a cidade de Botucatu	47

LISTA DE ABREVIATURAS E SÍMBOLOS

AOD	→	Profundidade ótica de aerossol
E_0	→	Fator de correção da excentricidade da órbita terrestre
h	→	Ângulo horário
H_0	→	Radiação no topo da atmosfera
H_{0UV}	→	Radiação ultravioleta no topo da atmosfera
H_d	→	Radiação difusa
H_G	→	Radiação solar global
H_{PAR}	→	Radiação solar fotossinteticamente ativa
H_{UV}	→	Radiação solar ultravioleta
IV	→	Infravermelha
K_t	→	Transmissividade Atmosférica
K_{tUV}	→	Transmissividade da radiação ultravioleta
K_{UV}	→	Fração da radiação ultravioleta
MAE	→	Mean absolute error (Erro médio absoluto)
MBE	→	Mean bias error (Erro médio padrão)
MLP	→	Multilayer perceptron (Perceptron de múltiplas camadas)
m_r	→	Massa ótica relativa
NSE	→	Nash-Sutcliffe coefficients (Coeficiente de Nash-Sutcliffe)
PAR	→	Photosynthetically active radiation (Radiação fotossinteticamente ativa)
R^2	→	Coeficiente de determinação
RBF	→	Radial basis function (Função de base radial)
RMSE	→	Root mean square error (Erro médio quadrático)
RNA	→	Redes neurais artificiais
SVM	→	Support vector machine (Máquina de Vetores de Suporte)
T_{max}	→	Temperatura máxima
T_{min}	→	Temperatura mínima
T_{med}	→	Temperatura média
TAM	→	Técnica de aprendizado de máquina
T_d	→	Temperatura do ponto de orvalho
UV	→	Ultravioleta
w	→	Peso associado a cada dado de entrada de uma RNA
$\bar{Y}$	→	Média dos valores medidos

Y_C	➔	Valores de radiação calculados
$\bar{Y}_c$	➔	Média dos valores calculados
Y_m	➔	Valores de radiação medidos
ϕ	➔	Latitude local
δ	➔	Declinação solar
φ	➔	Razão de nebulosidade
Δ	➔	Índice de brilho de céu
ε	➔	Índice de claridade de Perez
θ	➔	Bias
θ_z	➔	Ângulo zenital

RESUMO

O presente trabalho tem como objetivo avaliar a estimação da radiação solar ultravioleta diária (UV) utilizando técnicas de aprendizado de máquina em Botucatu/SP/Brasil. Para a geração dos modelos foram utilizadas as redes neurais artificiais com função linear (RNA), a máquina de vetores de suporte com função linear (SVM-Linear) e a máquina de vetores de suporte com função RBF (SVM-RBF). Como entrada, para cada uma das técnicas, foram testados cinco grupos contendo diferentes variáveis meteorológicas medidas como rotina na estação de radiometria solar de Botucatu. A maior precisão na estimação da UV foi obtida utilizando a SVM-RBF e, como variáveis de entrada, foram utilizados dados de insolação, umidade relativa, temperatura do ar, precipitação, fotoperíodo, radiação solar no topo da atmosfera, radiação ultravioleta no topo da atmosfera, radiação solar global, transmissividade atmosférica e massa óptica relativa. Este modelo apresentou $MBE = 0,321\%$; $RMSE = 5,712\%$; $d = 0.991$ e $R^2 = 0.969$, porém o conjunto de entradas utilizando apenas a radiação global, radiação solar no topo da atmosfera e radiação ultravioleta no topo da atmosfera merece atenção por apresentar resultados próximos a partir de apenas uma variável medida (radiação global) e duas calculadas. Este grupo apresentou $MBE = 1,614\%$; $RMSE = 6,749\%$; $d = 0,989$ e $R^2 =$

0,959. Tais resultados foram comparados com modelos encontrados na literatura e se mostraram equivalentes aos que utilizaram técnicas de aprendizado de máquina para estimar a radiação UV em outras cidades e superiores aos modelos estatísticos gerados com o intuito de estimar a UV para a cidade de Botucatu.

Palavras chave: Função de Base Radial. Modelos de Estimativa. Redes Neurais Artificiais. Máquina de Vetores de Suporte.

ESTIMATIVE OF ULTRAVIOLET SOLAR RADIATION IN BOTUCATU/SP/BRAZIL

UTILIZING MACHINE LEARNING TECHNIQUES. Botucatu, 2013. XXXp. Dissertação (Mestrado em Agronomia / Energia na Agricultura) – Faculdade de Ciências Agronômicas, Universidade Estadual Paulista.

Author: THIAGO DO NASCIMENTO SANTANA DE ALMEIDA

Adviser: JOÃO FRANCISCO ESCOBEDO

SUMMARY

In this paper is evaluated the estimation of daily solar ultraviolet radiation (UV) using machine learning techniques in Botucatu / SP / Brazil. To develop the model was utilized the artificial neural networks with linear function, the support vector machine with linear function and with RBF function. As input to each of the techniques, were tested five groups containing different weather variables measured as routine in Botucatu radiometry solar station. A more precise estimation of UV was obtained using the SVM-RBF and as input variables, we used data of insolation, relative humidity, air temperature, precipitation, photoperiod, solar radiation at the top of the atmosphere, ultraviolet radiation at the top of atmosphere, solar radiation, atmospheric transmissivity and relative optical mass. This model presented MBE = 0.321%; RMSE = 5.712%; $d = 0.991$ and $R^2 = 0.969$, but the set of entries using only the global radiation, solar radiation at the top of the atmosphere and ultraviolet radiation at the top of the atmosphere deserves attention for presenting similar results from only one measured variable (global solar radiation) and two calculated. This group presented MBE = 1.614%; RMSE = 6.749%; $d = 0.989$ and $R^2 = 0.959$. These results were compared to models found in the literature and proved equivalent to that used machine learning techniques to estimate the UV radiation in other cities and higher than the statistical models generated in order to estimate the UV to the Botucatu city.

Keywords: Radial Basis Function. Estimative Models. Artificial Neural Network. Support Vector Machine.

1 INTRODUÇÃO

O sol representa a principal fonte de energia do planeta Terra, pois praticamente todas as outras utilizadas pelo homem tem em sua origem a influência dela. São exemplos: a energia de biomassa, que tem em sua formação a energia solar no processo de fotossíntese, e a energia eólica, onde os ventos são originados do aquecimento da superfície terrestre pela energia solar.

A energia solar chega até superfície terrestre na forma de ondas eletromagnéticas (radiação), sendo que as características destas ondas variam de acordo com seu comprimento de onda (λ). Quanto menor for o comprimento de onda, maior é sua frequência e, conseqüentemente, maior é a energia transportada. Convencionalmente, a energia solar que incide sobre a superfície terrestre é chamada de radiação solar global. Ela é formada pelas as faixas espectrais do ultravioleta (100-400 nm), fotossinteticamente ativa ou visível (400-750 nm) e infravermelho (750-4000 nm) (ESCOBEDO et al., 2011).

Em media, a radiação global é composta por 5% da ultravioleta (UV), 49% da fotossinteticamente ativa (PAR) e 46% da infravermelho (IV) (ESCOBEDO et al., 2009). Dentre as três faixas espectrais, a UV representa a menor parcela da radiação solar global, porém requer atenção devido a seus efeitos em organismos vivos.

Considerando a influência da radiação solar ultravioleta sobre os organismos vivos, além do efeito em materiais sintéticos, estudos buscando conhecer sua variação temporal e espacial tornam-se interessantes (PORFIRIO et al., 2012; CUI et al., 2008; FIOLETOV et al., 2003; FOYO-MORENO et al., 1998; ESCOBEDO et al. 2011). Isso pode ser verificado na literatura a partir de vários trabalhos desenvolvidos nesse sentido:

- Em seres humanos, a UV pode agir como um fator cancerígeno ou causar reações de fotosensibilidade na pele (BACHELOR; BOWDEN, 2004; AGAR et al., 2004);
- Em organismos aquáticos como plantas de águas rasas, peixes e crustáceos, pode prejudicar o desenvolvimento ao causar danos nas células e tecidos (ROSEMA et al., 2002; PIENITZ; VICENT, 2000; CLAIR et al., 2001; HARRISON; SMITH, 2009; NAZARI et al., 2010);
- Nas plantas terrestres, induz mudanças estruturais e fisiológicas podendo causar redução da área foliar e redução na absorção de luz visível (TOSSERAMS; ROZEMA, 1995; HAO et al. 2000; BOEGER; POULSON, 2006);
- Fotodecomposição de águas residuais e de contaminantes industriais mediante o uso de concentradores solares (SOMMER et al., 1996; BLATCHLEY et al., 1998 ; JIMÉNEZ et al., 2000);
- É reconhecido também seu efeito como germicida além da degradação em materiais sintéticos como plástico e borracha (WHO, 2006).

No Brasil existem poucas localidades que monitoram a UV como rotina e, dentre as localidades que a monitoram, poucas possuem séries climatológicas longas (com mais de 15 anos). Uma alternativa para esta ausência de monitoramento é a sua estimativa através de modelos.

Em relação aos modelos estatísticos de estimativa da radiação UV, são comumente encontrados na literatura dois principais tipos. O primeiro estima a UV em função de dados de radiação solar global (CAÑADA et al., 2003; ROOBA, 2004; OGUNJOBI; KIM, 2004; ESCOBEDO et al., 2009) e o segundo estima a UV a partir do índice de transmissividade atmosférica K_{TUV} , sendo que o K_{TUV} é estimado a partir da equação $K_{TUV} = UV/UV_0$, onde UV_0 é a radiação solar ultravioleta extraterrestre. Esse

índice expressa a variação da UV em função do clima, eliminando a dependência astronômica e geográfica.

Outra alternativa é a utilização de técnicas de aprendizado de máquina (TAM), uma subárea da inteligência artificial, utilizando dados de outros parâmetros meteorológicos, o que nos permitiria não apenas prever, mas também estimar valores perdidos em bases de dados (LEAL *et al.*, 2011; BARBERO *et al.*, 2006; FONSECA JUNIOR *et al.*, 2011; VILLÀN *et al.*, 2010).

As TAM baseiam-se em um conjunto de variáveis que servirão de entrada e, a partir delas, obtêm-se conclusões genéricas utilizando o método de indução para que problemas complexos de diversas áreas de conhecimento sejam resolvidos.

Ao comparar o aprendizado de máquina com o aprendizado natural, as variáveis de entrada representarão o ambiente externo, onde podemos recolher informações de uma determinada área de conhecimento. A análise que fazemos sobre os novos dados coletados é feita, na técnica utilizada, sob a forma de modificações dos seus parâmetros livres, onde o algoritmo gera hipóteses buscando aquela que apresente uma melhor capacidade de descrever o domínio. As conclusões sobre a nova situação serão o resultado de como iremos responder ao ambiente após as análises ou, no caso da técnica, do resultado da interação dos dados de entrada com os parâmetros livres já modificados.

As principais vantagens obtidas ao utilizar as TAM são: fácil utilização e ótimos resultados, muitas vezes superiores aos modelos estatísticos (CHEN *et al.*, 2011; WU; LIU, 2012).

Assim, com base nas informações expostas, o objetivo deste trabalho é estimar e avaliar a estimativa da radiação solar ultravioleta diária em Botucatu/SP/Brasil, utilizando técnicas de aprendizado de máquina, a partir de variáveis meteorológicas calculadas ou medidas como rotina na maioria das estações climatológicas no Brasil.

2 REVISÃO BIBLIOGRÁFICA

2.1 Radiação Solar Ultravioleta

A radiação solar ultravioleta (UV) é definida como a porção do espectro eletromagnético com comprimento de onda entre 0,1 a 0,4 μm . Essa radiação corresponde a aproximadamente 5,7% da radiação incidente no topo da atmosfera e a 4% da radiação global (IQBAL, 1983; ELHADIDY et al., 1990; ROOBA, 2004; JACOVIDES et al., 2006; ESCOBEDO et al., 2012). A UV é dividida em três faixas, de acordo com o comprimento de onda e, conseqüentemente, com seu potencial energético: UVC (100-280 nm), UVB (280-315 nm) e UVA (315-400 nm), sendo que a maior parcela da radiação UV extraterrestre (UV_0) corresponde a UVA (68%), seguida pela UVC (17%) e por último a UV-B (15%) (IQBAL, 1983) (Figura 1).

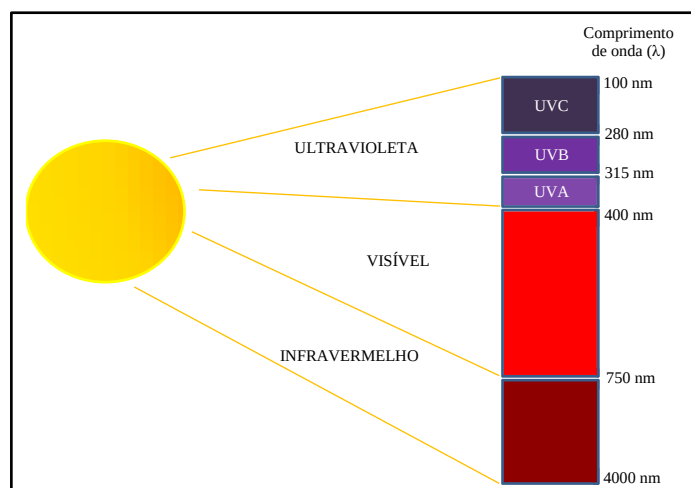


Figura 1. Espectro Solar.

Ao atingir a atmosfera, a radiação UV sofre uma atenuação pelo ozônio que varia de acordo com seu comprimento de onda. A UVA praticamente não sofre interação com a camada de ozônio, porém parte da UVB e toda a UVC são absorvidas ainda na estratosfera. De toda a radiação ultravioleta medida na superfície terrestre, a fração correspondente à UVA representa aproximadamente 90% do total e a UVB 10% (KORONAKIS et al, 2002).

Os níveis de incidência da radiação solar UV sobre a superfície terrestre dependem da combinação da influência de vários fatores, sendo as condições de nebulosidade e a concentração de aerossóis e de ozônio na atmosfera os principais (JOSEFSSON; LANDELIUS, 2000; FOYO-MORENO et al., 2003; VILLÁN et al., 2010). A atenuação da UV por estes constituintes atmosféricos pode ocorrer pelos processos físicos de espalhamento, reflexão e absorção.

Devido a sua característica de espalhamento, os efeitos da UVB e UVA tornam-se similares e, mudanças na cobertura de nuvem podem alterar significativamente a incidência da UV na superfície terrestre. Porém, a análise da atenuação causada por nuvens requer informações instantâneas, como, por exemplo, a espessura, a posição, o número de camadas, etc. Tais informações podem ser obtidas a partir de instrumentos específicos ou através de satélites (SANTOS et al., 2011; JOSEFSSON; LANDELIUS, 2000; CALBÓ et al., 2005; SABBURG; PARISI, 2005; FOYO-MORENO et al., 2003).

Além da camada de ozônio e do efeito causados pelas nuvens, a UV também é fortemente atenuada por aerossóis providos da queima de biomassa. Essa

interação pode reduzir sua incidência de 15% a 35% (LATHA et al., 2005; KALASHNIKOVA et al., 2007). Kylling et al. (1998) observou que sob condições de céu limpo, a UV pode sofrer decréscimo de 5% a 35%, dependendo da profundidade ótica de aerosol (AOD) e do albedo.

Devido a grande variação climática em diferentes localidades, e a escassez de séries climáticas longas, vários estudos foram desenvolvidos com o intuito de criar diferentes modelos visando estimar a radiação UV. Esses modelos podem ser estatísticos ou paramétricos, sendo este último de difícil utilização, pois exige um avançado conhecimento no processo de transferência atmosférica.

2.1.1 Modelos Estatísticos de Estimativa da UV

Vários autores têm proposto diferentes modelos que relacionam o K_{TUV} com os principais parâmetros atmosféricos que influenciam a transmissividade (MARTINEZ-LOZANO et al., 1999; FOYO-MORENO et al., 1999; CAÑADA et al., 2003; OGUNJOBI & KIM, 2004; VARO et al., 2005; BARBERO et al., 2006). Dois dos parâmetros mais utilizados são a massa ótica relativa do ar (m_r) e o índice de claridade K_t . A m_r é definida como a razão entre o caminho ótico numa trajetória oblíqua e o caminho ótico na direção vertical (zenital). Ela representa o caminho ótico atmosférico percorrido pela radiação solar, em uma determinada direção em relação ao zênite local, do topo da atmosfera até a superfície terrestre (KASTEN & YOUNG, 1989). O K_t (transmissividade atmosférica), razão entre a radiação solar global e a radiação solar extraterrestre, é utilizado principalmente para avaliar as condições atmosféricas em relação à concentração de nuvens (KUDISH; IANETZ, 1996; LI; LAM, 2000; ESCOBEDO et al., 2009; OKOGBUE et al., 2009). Assim, a maioria dos modelos do segundo tipo expressa a UV pela equação $K_{TUV} = aK_T^b m_r^c$, onde “a”, “b” e “c” são constantes, K_T é a transmissividade atmosférica, adimensional, e “ m_r ” é a massa ótica relativa do ar, adimensional.

Na literatura encontram-se vários modelos tradicionais de estimativas para diversos parâmetros agrometeorológicos. Porém, nos últimos anos, o desenvolvimento de modelos criados a partir de técnicas de aprendizado de máquina (TAM) se mostra cada vez mais frequente, como por exemplos os trabalhos de Leal et al. (2011), Barbero et al. (2006), Erusiafe; Chendo (2010) e Abdulazeez (2011). Como

exemplo de TAM pode-se citar as redes neurais artificiais e a máquina de vetores de suporte.

2.2 Redes Neurais Artificiais

Uma das técnicas de aprendizado de máquina mais difundidas, sem dúvidas, é a Rede Neural Artificial (RNA).

Esta técnica recebeu este nome por sua capacidade de “aprender”, ou seja, adquirir conhecimento sobre o meio externo através da associação de padrões de exemplos anteriores, da mesma forma que o cérebro humano.

Em uma RNA existe uma coleção de neurônios, divididos em camadas e conectados por ligações, dando a característica de um processamento paralelo. Segundo Souza (2008), estes neurônios são considerados os elementos básicos de uma RNA e são constituídos de:

- Pesos sinápticos, que representarão a importância de determinada entrada;
- Uma regra de propagação, que definirá a combinação das entradas no neurônio;
- Uma função de ativação que define o resultado da regra de propagação sobre o nível de ativação do neurônio, sendo as funções de limiar, linear e sigmoidal as mais utilizadas.

A primeira camada de uma RNA é formada por pseudoneurônios com o único intuito de receber os dados de entrada. Após definidas as entradas, são associados pesos de acordo com a importância determinada pela rede. Além dos pesos, existe outro parâmetro livre: o *bias*, que tem a propriedade de aumentar ou diminuir a influência do valor de entrada. Consideram-se, estes parâmetros livres, os responsáveis por simular a capacidade de armazenar o conhecimento adquirido e torná-lo disponível para uso futuro. Posteriormente será gerada uma saída, com sua amplitude restringida pela função de ativação utilizada.

Por definição, para o funcionamento das técnicas de aprendizado faz-se necessário que o conjunto de variáveis de entrada seja dividido em dois subconjuntos: de treinamento e de teste. O subconjunto de treino é utilizado para o

aprendizado, onde a técnica desenvolverá a capacidade de generalizar os resultados, ou seja, desenvolver a capacidade de produzir um resultado adequado para dados diferentes dos que foram utilizados no treinamento. O subconjunto de teste é utilizado para medir o grau de efetividade aprendido pela técnica utilizada na predição da classe desejada (LORENA; CARVALHO, 2007).

2.2.1 Histórico

Os primeiros estudos relativos a redes neurais artificiais datam do ano de 1943, no trabalho de McCulloch (um psiquiatra e neuroanatomista) e Pitts (um matemático), onde foi proposto um cálculo lógico que unificaria os estudos da neurofisiologia e da lógica matemática. Este cálculo foi inserido em um modelo de neurônio artificial que assumia um comportamento de “tudo ou nada” (Figura 2). Ao se conectar vários desses neurônios e ajustar suas conexões sinápticas seria possível computar qualquer função computável (PARREIRAS, 2003).

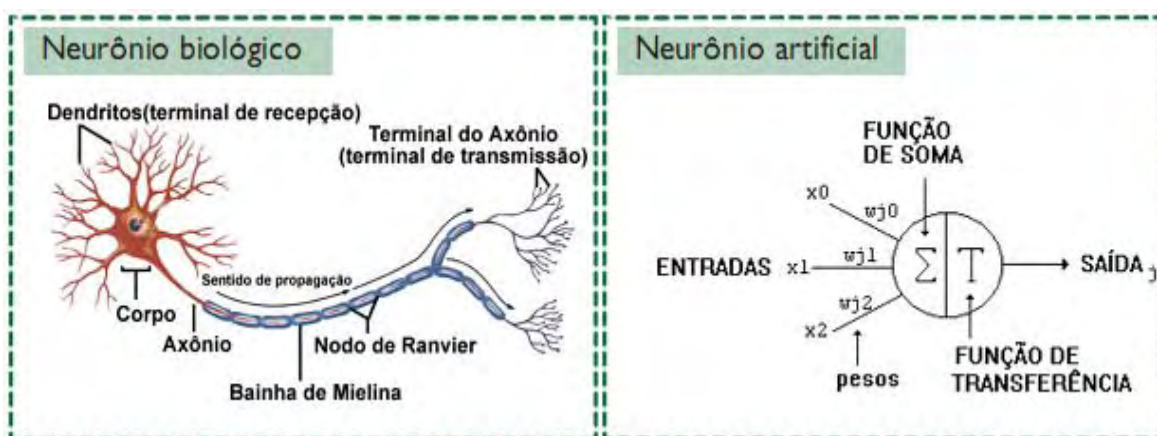


Figura 2. Neurônio biológico e Neurônio artificial.

A partir deste trabalho foram gerados muitos outros, com o intuito de analisar e melhorar as RNA.

Em 1958 Rosenblatt propôs o perceptron, que viria a ser a primeira técnica de aprendizagem supervisionada; em 1962 Widrow desenvolveu a primeira rede de múltiplas camadas e em 1986 Rumelhart, Hilton e Williams demonstraram que algoritmos de retropropagação poderiam ser utilizados em técnicas de aprendizado de máquina com grande eficiência.

Obviamente outros trabalhos foram desenvolvidos, como os de Broomhead e Lowe de 1988 que anexaram às RNA funções de base radial e os trabalhos de Vapnik, no início dos anos 90, que implementaram uma poderosa rede supervisionada chamada de máquina de vetores de suporte. Entretanto uma bibliografia extensa encontra-se no trabalho de Haykin (2001).

2.2.2 Perceptron

O perceptron é a forma mais simples de uma RNA. Ele é formado por um conjunto de neurônios, associado a uma função de ativação e um algoritmo de aprendizagem (Figura 3).

Sua topologia é descrita em três camadas, sendo que apenas uma apresenta a capacidade de ajustar os parâmetros livres (pesos - w e *bias* - θ). As outras duas são as camadas de entrada e saída (RIBEIRO, 2003).

Por definição, ajustando seus parâmetros livres, o perceptron consegue resolver qualquer problema de classificação, desde que este seja linearmente separável.

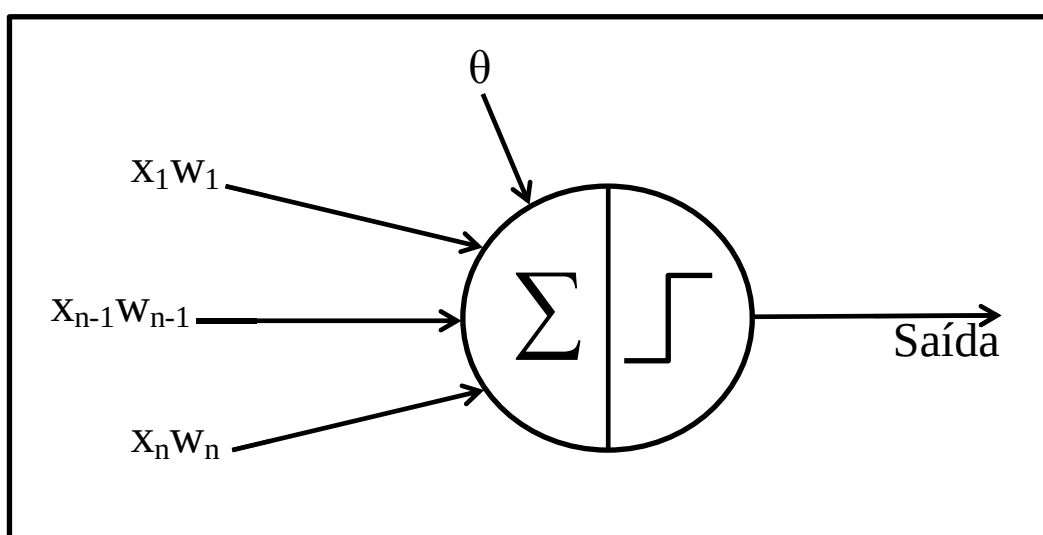


Figura 3. Perceptron.

2.2.3 Perceptron de Múltiplas Camadas

A Multilayer Perceptron, MLP ou Perceptron de Múltiplas Camadas é caracterizada por apresentar mais de uma camada e arquitetura alimentada a diante (feedforward) não-recorrente, ou seja, a saída da rede não serve de alimentação para uma camada anterior (FERREIRA, 2004; GUARNIERI, 2006) (Figura 4).

Sua formação dá-se por várias camadas fortemente conectadas contendo neurônios anexados a uma função de ativação:

- A primeira camada é chamada de entrada que, do mesmo modo que a Perceptron, somente recebe e distribui os dados para as camadas seguintes;
- Uma ou mais camadas intermediárias (podendo ser chamadas de escondidas ou ocultas) que realizam o processamento das entradas. Receberam este nome por não tornar acessível ao usuário o comportamento dos dados nesta fase;
- Uma camada de saída, responsável por apresentar o resultado gerado.

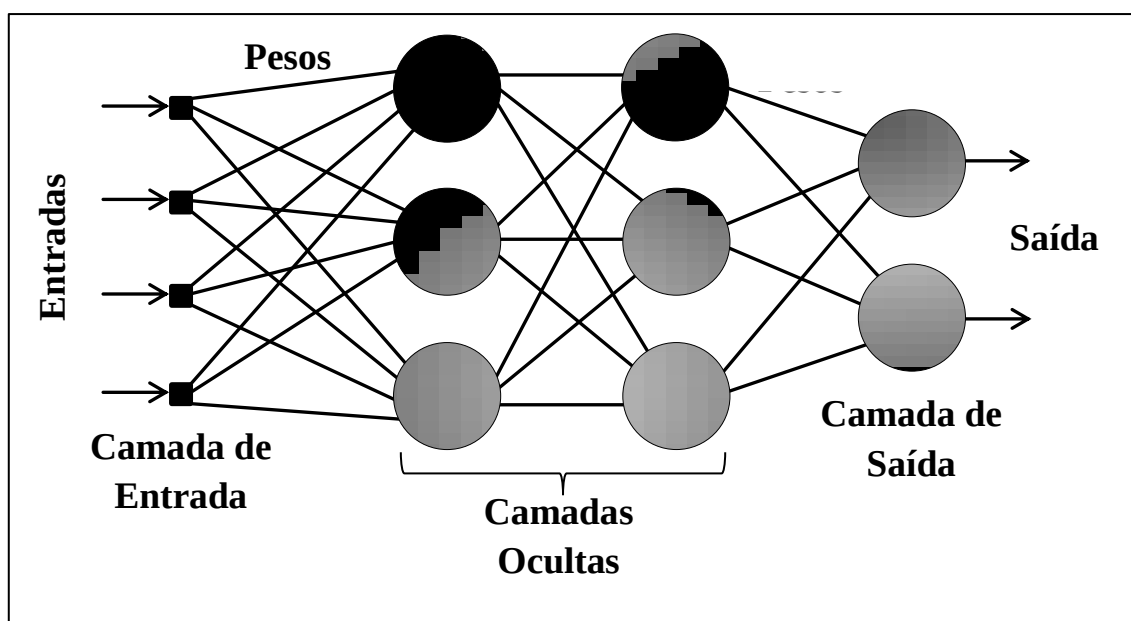


Figura 4. Exemplo de uma RNA com 4 entradas 2 camadas ocultas e 2 saídas.

2.2.4 Algoritmos de aprendizagem

É durante a fase de treinamento que os pesos da RNA convergem para determinados valores, variando de acordo com as interações. Essa convergência se dá a partir do ajuste dos pesos sinápticos, feita por um algoritmo de aprendizagem. Desta maneira que a aplicação dos dados de entrada produzirão as saídas necessárias (GUARNIERI, 2006).

O aprendizado de uma TAM pode ser classificado como:

- não supervisionado: onde o sistema não tem conhecimento sobre os resultados desejados, gerando-os a partir das características abstraídas das variáveis de entrada;
- supervisionado: onde as saídas são conhecidas por meio de rótulos ou classes, o que permite ao sistema convergir sua saída para o resultado disponibilizado, alterando seus parâmetros a cada iteração, até que haja uma estabilização em um valor mínimo de erro. (DORVLO *et al.*, 2002).

Na aprendizagem supervisionada, o algoritmo mais utilizado em redes neurais é o algoritmo de retropropagação (Backpropagation). Seu funcionamento baseia-se em um passo para frente (propagação) e um passo para trás (retro-propagação).

O algoritmo inicia atribuindo valores aleatórios para os pesos sinápticos, que serão associados a um valor de entrada, para depois processá-los camada a camada, até obter o valor de saída. A partir do valor de saída calcula-se o erro entre ele e o valor esperado, gerando o erro médio quadrado. Esse o passo é chamado de propagação.

No passo de retro-propagação é feito o ajuste dos pesos, baseado no erro médio quadrado, obtido anteriormente. São calculados os termos relacionados com os pesos da última camada da rede, em seguida os termos que envolvam os pesos da penúltima camada e assim por diante, até a primeira camada da rede, daí o nome de retropropagação.

2.3 Máquina de Vetores de Suporte

Dentre as técnicas de aprendizado de máquina disponíveis, a máquina de vetores de suporte (SVM ou *Support Vector Machine*) tem se destacado, resolvendo problemas de análise de expressões, reconhecimento de padrões em imagens e sons, regressão linear, entre outros, obtendo ótimos resultados com um baixo tempo de treinamento e associados a resultados muitas vezes superiores às redes neurais artificiais devido a sua alta capacidade de generalização. (SCHÖLKOPF; SMOLA, 2002; CUCKER; ZHOU, 2007; STEINWART; CHRISTMANN, 2008). São frequentemente indicados, para esta técnica, procedimentos de regressão tradicionais/estatística, pois os processos de derivação de uma função apresentam o menor desvio entre as respostas estimadas e observadas para todos os exemplos de treinamento (BASAK *et al.*, 2007).

A SVM descende da teoria do aprendizado estatístico apresentado no algoritmo "*Generalized Portrait*" de Vapnik, Lerner e Chervonenkis (FERRÃO, 2007) e seu funcionamento se baseia em vetores de suporte (exemplos de dados de treinamento mais próximos de uma classe diferente), sendo que a partir deles cria-se um hiperplano como superfície de decisão com a máxima separação entre exemplos de diferentes classes e alta capacidade de generalização (Figura 5).

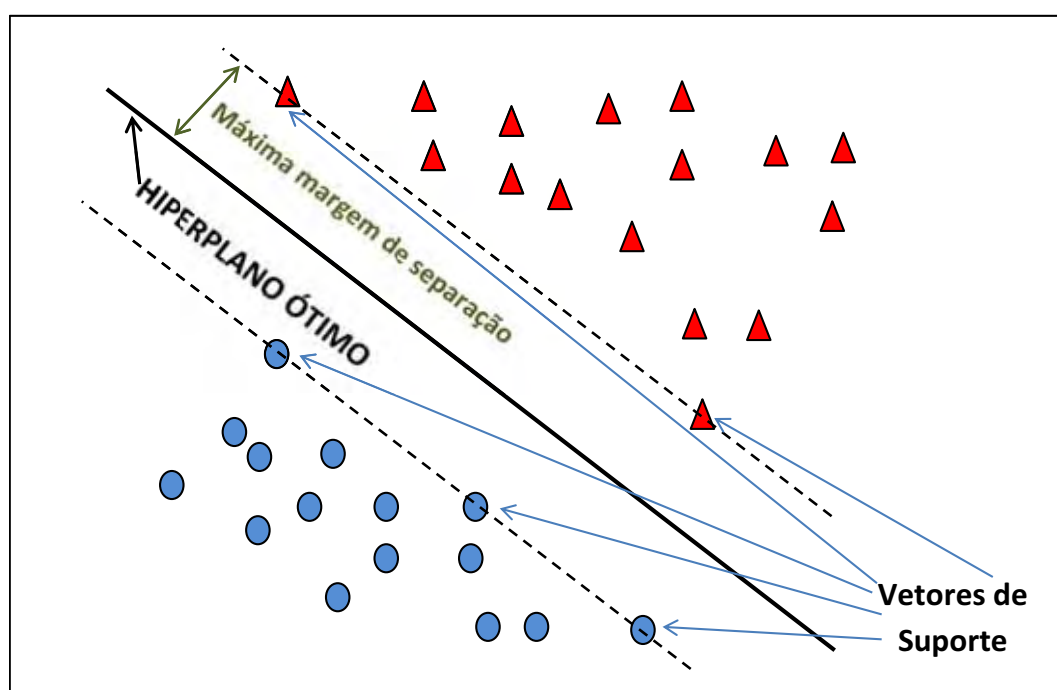


Figura 5. Funcionamento da SVM.

Segundo Vapnik (1998), para as técnicas de aprendizado estatístico deseja-se que as seguintes características estejam presentes:

- Um conjunto flexível e grande o suficiente de funções indicadoras disponíveis, para representar o comportamento do conjunto de dados. Essas funções indicadoras são responsáveis por tentar simular o comportamento dos dados.
- Um princípio indutivo, capaz de associar o conjunto de treinamento à função do sistema. Por exemplo: regularização, minimização do risco empírico, minimização do risco estrutural e inferência Bayesiana.
- Um algoritmo de aprendizagem, que indique como implementar o princípio indutivo e selecionar a melhor função dentro do universo das funções existentes.

O princípio indutivo utilizado pela SVM é a minimização do risco estrutural, o que significa que ela buscará reduzir juntamente os erros do conjunto de treinamento e do conjunto de teste, resultando em uma maior generalização e, normalmente, um maior acerto na fase de teste (ANDREOLA, 2009).

Em 1963, o objetivo inicial da SVM foi encontrar um hiperplano ótimo (com máxima distância entre classes) em função de vetores de suporte. Esse método permitiria uma classificação linear dos dados. Em 1992 surge uma implementação onde o espaço de entrada é transformado em um espaço de maior dimensão, tornando linear a separação dos dados que antes não podiam ser separados desta maneira (Figura 6). Em 1995 foram introduzidas variáveis de folga, também com o intuito de tornar a técnica mais flexível e aumentar sua capacidade de generalização (ALES, 2008).

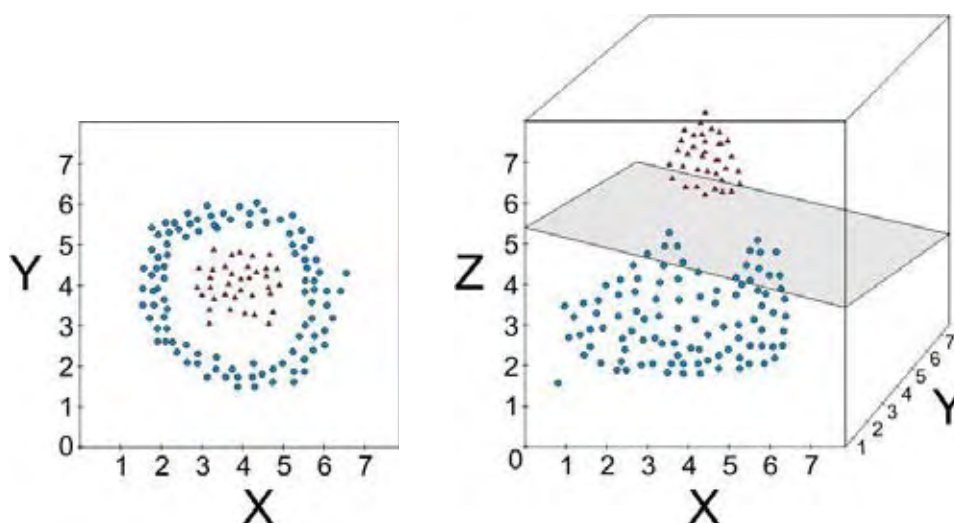


Figura 6. Aumento de dimensionalidade da SVM.

Para realizar uma regressão, a SVM tenta minimizar o erro generalizado ao invés de minimizar o erro de treinamento observado, obtendo assim um desempenho generalizado. Este erro generalizado é a combinação do erro de treinamento e de um termo de regularização que controla a complexidade do espaço hipótese (BASAK *et al.*, 2007).

O aprendizado da máquina de vetores de suporte, como o das redes neurais artificiais, é gerado a partir de variáveis que servirão de entrada para a técnica. Ao analisar estas variáveis, o algoritmo gera hipóteses buscando aquela que apresente melhor capacidade de descrever o domínio. Ao comparar a SVM a uma RNA é possível considerar o número de nós da camada de entrada é igual à dimensão do vetor de entrada. A quantidade de vetores-suporte determina a quantidade de nós na camada escondida e o nó de saída constrói uma função linear no espaço característico, determinada por uma transformação não-linear (Figura 7) (SEMOLINI, 2002):

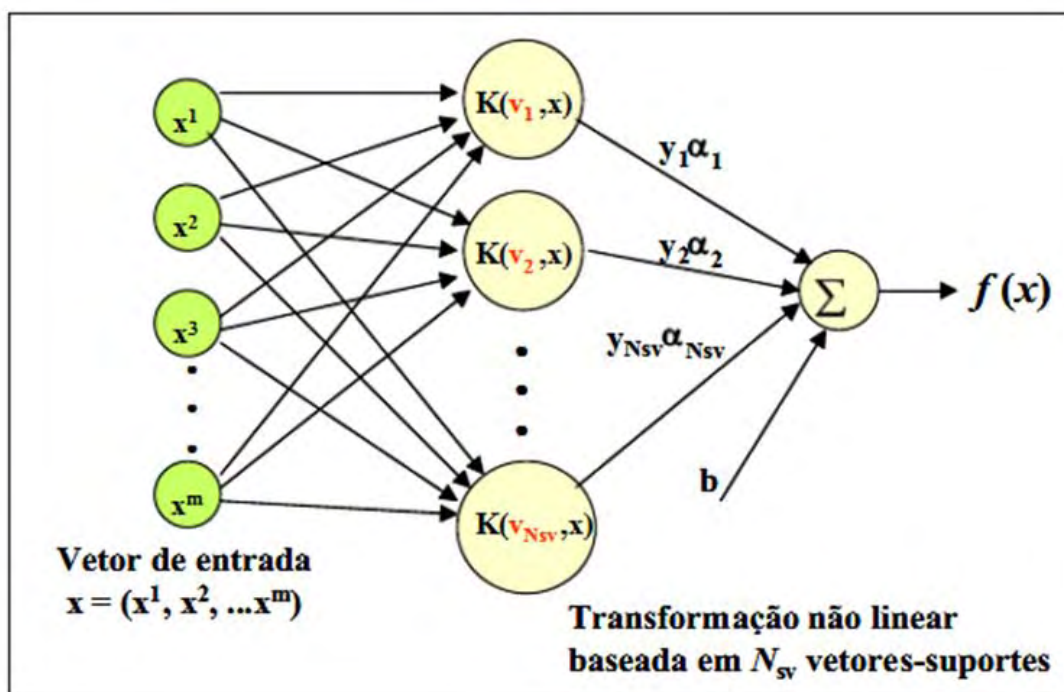


Figura 7. Simulação de uma SVM no modelo de neurônios de uma RNA.

2.4 Função de Base Radial

Segundo Tiago; Leitão (2002), são chamadas Funções de Base Radial (RBFs) as funções que apresentam simetria radial, ou seja, dependem apenas da distância entre o centro da função e um ponto x .

A utilização das RBF como função de ativação em redes neurais iniciou-se em meados dos anos 80, com o intuito de suprimir algumas deficiências das redes MLP. O emprego das RBF proporcionou uma maior simplicidade no processo de treinamento, a eficiência computacional aumentou e houve uma maior eficiência na aproximação de funções (POWELL, 1985). Isso acontece devido a sua capacidade de transformar um espaço de menor dimensão em um espaço de maior dimensão (OLIVEIRA, 2011).

O aprendizado de uma rede RBF vem do ajuste de uma superfície não-linear ao conjunto de dados, em um espaço multidimensional, através de algum critério estatístico. Ao usar esta superfície multidimensional para interpolar pontos diferentes dos encontrados no conjunto de treino, mas que estejam em sua vizinhança, acontece a generalização do modelo. (CARVALHO; BRAGA, 2004).

Porém, segundo Altran (2005), deve-se atentar às características de cada função, pois nem todas as funções de base radial são adequadas ao uso em técnicas de aprendizado de máquina. Algumas funções podem apresentar menor tempo computacional, enquanto outras uma maior capacidade de generalização, entretanto as funções multiquadráticas e splines de placas finas, por exemplo, não são adequadas por serem ilimitadas e não diferenciáveis, respectivamente.

Como principais exemplos de RBF pode-se citar as seguintes funções:

- Multiquadrática: $\phi(x) = \sqrt{x^2 + c^2}$, $c > 0$.

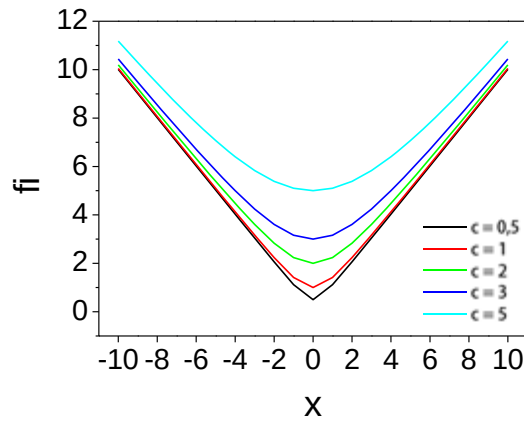


Figura 8. Função Multiquadrática.

- Multiquadrática recíproca: $\phi(x) = \frac{1}{\sqrt{x^2 + c^2}}$, $c > 0$.

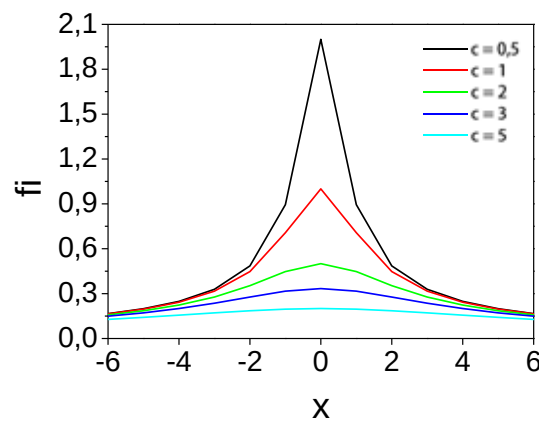


Figura 9. Função Multiquadrática Recíproca.

- Multiquadrática recíproca inversa: $\phi(x) = \frac{1}{c} - \frac{1}{\sqrt{x^2 + c^2}}$, $c > 0$.

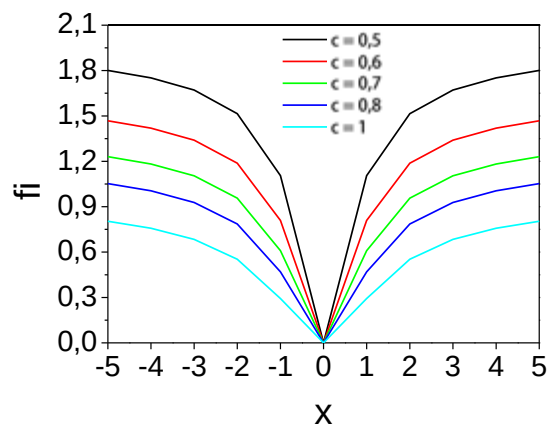


Figura 10. Função Multiquadrática Recíproca Inversa.

- Gaussiana: $\phi(x) = \exp(-cx^2)$, $c > 0$.

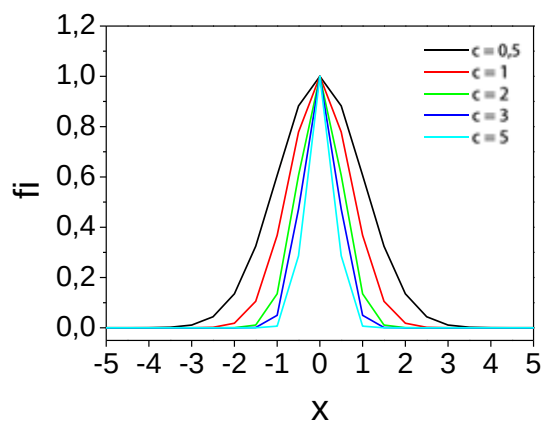


Figura 11. Função Gaussiana.

- Secante Hiperbólica: $\phi(x) = \frac{2}{e^{cx} + e^{-cx}}$, $c > 0$.

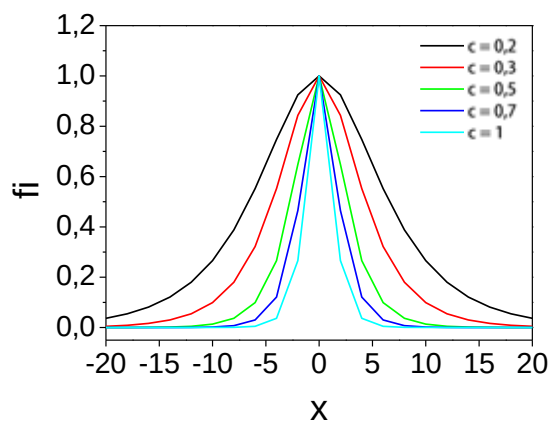


Figura 12. Função Secante Hiperbólica.

Segundo a análise de Dias (2006), em geral, as redes RBF vêm obtendo melhores resultados, para problemas de classificação, do que as redes de retropropagação (backpropagation), porém seu aprendizado exige uma quantidade maior de dados para a fase de treinamento.

2.5 Trabalhos Relacionados

Como mencionado anteriormente, o conhecimento dos valores de radiação solar ultravioleta são de grande importância para várias áreas de conhecimento e diferentes modelos são encontrados ao redor do mundo. Entretanto, a grande maioria dos trabalhos envolvendo radiação solar (UV e outras) possuem como metodologia os modelos tradicionais, já bem difundidos.

Neste meio, trabalhos contendo técnicas de aprendizado de máquina ainda são escassos, mas devido aos resultados promissores e a facilidade de utilização, este número está crescendo rapidamente.

Como exemplo, para a estimação da radiação ultravioleta, podemos citar os trabalhos de Leal et al. (2011), que propuseram modelos para as cidades de Pesqueira/PE e Araripina/PE e Barbero et al. (2006), que propuseram modelos para as cidades de Almería, na Espanha e Golden, Colorado nos EUA. Ambos utilizaram as RNA e compararam os resultados com modelos estatísticos. Na Tabela 1 é apresentada a validação dos modelos propostos pelos dois autores:

Tabela 1: Validação dos modelos propostos por Leal et al. e Barbero et al.

CIDADE	MODELO	MBE(%)	RMSE(%)	R ²
PESQUEIRA	$H_{UV} = f(H_G)$	-0,4	4,6	0,96
	$K_{UV} = f(K_t)$	-0,5	4,6	0,88
	$K_{t_{UV}} = f(K_t; m_r)$	-0,5	4,5	--
	RNA 1: $\{K_t\}$	-1,6	4,8	--
	RNA 2: $\{K_t, m_r\}$	-1,3	4,6	--
ARARIPINA	$H_{UV} = f(H_G)$	0,3	4,4	0,96
	$K_{UV} = f(K_t)$	-0,1	4,3	0,93
	$K_{t_{UV}} = f(K_t; m_r)$	-0,1	4,4	--
	RNA 1: $\{K_t\}$	2,0	4,9	--
	RNA 2: $\{K_t, m_r\}$	0,7	4,6	--
ALMERÍA	$H_{UV} = f(H_G)$	1,0	6,0	--
	$H_{UV} = f(K_t; m_r)$	0,0	4,7	--
	RNA 1: $\{m_r, H_G, K_t\}$	0,0	4,4	--
GOLDEN (COMPLETO)	$H_{UV} = f(H_G)$	0,5	6,9	--
	$H_{UV} = f(K_t; m_r)$	-1,0	5,5	--
	RNA: $\{m_r, H_G, K_t\}$	-3,0	6,5	--
GOLDEN (SEM NEVE)	$H_{UV} = f(H_G)$	1,0	6,4	--
	$H_{UV} = f(K_t; m_r)$	-0,1	4,6	--
	RNA: $\{m_r, H_G, K_t\}$	-2,0	5,1	--

Por ser uma radiação medida como rotina em várias localidades, modelos para estimação da radiação global são mais comumente encontrados na literatura. Para estimar esta radiação as técnicas de aprendizado de máquina também se mostram eficientes. Erusiafe; Chendo (2010) propuseram um modelo para a cidade de Akoka, na Nigéria, utilizando como entrada das RNA imagens de satélite. No trabalho de Abdulazeez (2011) que também utilizou as RNA, foram utilizados dados de K_t , Razão de Insolação e Temperatura média da cidade de Gusau, Nigéria. Na Tabela 2 é apresentada a validação dos modelos propostos para a radiação global:

Tabela 2: Validação dos modelos propostos por Erusiafe & Chendo; AbdulAzeez;

CIDADE	MBE (relativo)	RMSE (relativo)	R ²
Akoka (10h-11h)	-0,1202	0,0349	0,9324
Akoka (11h-12h)	-0,0405	0,0276	0,911
Akoka (12h-13h)	0,0128	0,0356	0,9186
Gusau	--	0,0028	0,9996

Deng et al. (2010) estimou a radiação Global utilizando RNA. Em seu trabalho foram desenvolvidos 14 modelos para 10 cidades da China. Entre esses modelos o que apresentou melhores resultados teve como entrada dados de dia juliano, latitude, longitude, brilho solar, temperatura média, precipitação, umidade relativa e pressão atmosférica e, posteriormente, foi comparado com o modelo de Angstrom. A RNA apresentou resultados de RMSE(%) entre 1,77 a 3,30 e R² entre 0,863 e 0,951 enquanto o modelo de Angstrom apresentou RMSE(%) entre 1,77 e 2,64 e R² entre 0,832 e 0,942.

Wu; Liu (2012) também desenvolveram modelos para a radiação global, porém utilizando a SVM. Em seu trabalho foram gerados 5 modelos para 24 cidades da China. Os modelos criados tinham como entrada as seguintes variáveis: modelo 1: Tmax-Tmin; modelo 2: Tmax, Tmin; modelo 3: Tmax-Tmin,Tmed; modelo 4: Tmax, Tmin, Tmed; modelo 5: Tmax, Tmin, Tmax-Tmin. Para o modelo 1 a SVM gerou um MBE variando de -0,116 MJm⁻² a 0,195 MJm⁻² e um RMSE entre 0,823 MJm⁻² e 4,203 MJm⁻²; para o modelo 2 o MBE apresentado foi entre -0,174 MJm⁻² e 0,246 MJm⁻² e o RMSE entre 0,86 MJm⁻² e 5,271 MJm⁻²; o modelo 3 apresentou um MBE entre -0,135 MJm⁻² e 0,199 MJm⁻² e RMSE entre 0,81 MJm⁻² e 4,357 MJm⁻²; o modelo 4 apresentou MBE entre -0,587 MJm⁻² e 0,249 MJm⁻² e RMSE variando entre 1,017 MJm⁻² e 8,523 MJm⁻²; o último modelo apresentou MBE entre -0,289 MJm⁻² e 0,563 MJm⁻² e RMSE variando entre 0,956 MJm⁻² e 15,072 MJm⁻².

Outro exemplo de uso da SVM é o trabalho de Chen et al. (2011) que criou 7 modelos para estimativa de radiação global, baseados em temperatura, para a cidade de Chongqing / China. Cada um dos modelos foi repetido três vezes, sendo a primeira para SVM com função de ativação linear, a segunda para função de ativação

polinomial e a terceira para a SVM com função de ativação RBF. Na Tabela 3 é apresentada a validação dos modelos propostos:

Tabela 3: Validação dos modelos gerados por CHENG et al.

Modelo	Entradas	RMSE (%)	NSE
SVM-Linear 1	Tmax	19,43	0,8828
SVM-Linear 2	Tmin	24,00	0,8307
SVM-Linear 3	Tmax-Tmin	15,65	0,9173
SVM-Linear 4	Tmax, Tmin	10,11	0,9427
SVM-Linear 5	Tmax, Tmax-Tmin	11,94	0,9359
SVM-Linear 6	Tmin, Tmax-Tmin	12,17	0,9350
SVM-Linear 7	Tmax, Tmin, Tmax-Tmin	10,64	0,9410
SVM – Polin. 1	Tmax	18,90	0,9984
SVM – Polin. 2	Tmin	23,00	0,9979
SVM – Polin. 3	Tmax-Tmin	11,18	0,9984
SVM – Polin. 4	Tmax, Tmin	9,00	0,9990
SVM – Polin. 5	Tmax, Tmax-Tmin	10,50	0,9989
SVM – Polin. 6	Tmin, Tmax-Tmin	9,91	0,9989
SVM – Polin. 7	Tmax, Tmin, Tmax-Tmin	9,82	0,9989
SVM – RBF. 1	Tmax	18,23	0,9985
SVM – RBF. 2	Tmin	24,26	0,9979
SVM – RBF. 3	Tmax-Tmin	12,37	0,9985
SVM – RBF. 4	Tmax, Tmin	9,43	0,9989
SVM – RBF. 5	Tmax, Tmax-Tmin	12,04	0,9986
SVM – RBF. 6	Tmin, Tmax-Tmin	11,23	0,9983
SVM – RBF. 7	Tmax, Tmin, Tmax-Tmin	11,21	0,9987

Lópes et al. (2001) utilizou as RNA para estimar a radiação PAR para seis localidades na Espanha. Foram testadas várias combinações de entradas para a RNA sendo que foram selecionadas as duas que obtiveram os melhores resultados sendo uma baseada em informações radiométricas e meteorológicas e outra baseada no brilho

solar e em informações meteorológicas. Estes dois modelos foram comparados com modelos estatísticos. Na Tabela 4 é apresentada a validação dos modelos propostos:

Tabela 4: Validação dos modelos propostos por Lopes et al.

CIDADE	MODELO	MBE(%)	RMSE(%)	R ²
ALMERÍA	$H_{PAR} = f(H_G, \Delta, T_d, \varepsilon, \cos\theta_z)$	0,9	2,7	0,998
	$H_{PAR} = f(H_G, K_t, \cos\theta_z)$	1,4	3,0	0,998
	RNA 1: $\{H_G, \cos\theta_z\}$	0,0	2,0	0,999
	RNA 2: $\{n, \cos\theta_z\}$	0,1	8,5	0,977
GRANADA	$H_{PAR} = f(H_G, \Delta, T_d, \varepsilon, \cos\theta_z)$	-3,4	4,8	0,998
	$H_{PAR} = f(H_G, K_t, \cos\theta_z)$	-0,7	2,7	0,998
	RNA 1: $\{H_G, \cos\theta_z\}$	-2,2	3,7	0,998
DESERT ROCK	$H_{PAR} = f(H_G, \Delta, T_d, \varepsilon, \cos\theta_z)$	-4,4	5,1	0,998
	$H_{PAR} = f(H_G, K_t, \cos\theta_z)$	-5,3	6,0	0,997
	RNA 1: $\{H_G, \cos\theta_z\}$	-3,7	4,5	0,998
	RNA 2: $\{n, \cos\theta_z\}$	10,4	14,3	0,976
BONDVILLE	$H_{PAR} = f(H_G, \Delta, T_d, \varepsilon, \cos\theta_z)$	-3,1	5,5	0,997
	$H_{PAR} = f(H_G, K_t, \cos\theta_z)$	-3,5	5,9	0,996
	RNA 1: $\{H_G, \cos\theta_z\}$	-3,2	5,3	0,997
	RNA 2: $\{n, \cos\theta_z\}$	1,3	18,8	0,945
TABLE MONTAIN	$H_{PAR} = f(H_G, \Delta, T_d, \varepsilon, \cos\theta_z)$	0,6	4,8	0,995
	$H_{PAR} = f(H_G, K_t, \cos\theta_z)$	-0,1	4,8	0,995
	RNA 1: $\{H_G, \cos\theta_z\}$	1,0	4,5	0,996
	RNA 2: $\{n, \cos\theta_z\}$	-2,1	15,3	0,948
ABISKO	$H_{PAR} = f(H_G, K_t, \cos\theta_z)$	6,0	9,5	0,993
	RNA 1: $\{H_G, \cos\theta_z\}$	4,0	9,2	0,991

Sendo que: Δ representa o índice de brilho de céu, T_d representa a temperatura do ponto de orvalho, ε representa o índice de claridade de Perez e $\cos\theta_z$ representa o cosseno do ângulo zenital.

Soares et al. (2004) utilizaram as RNA para modelagem da radiação difusa no estado de São Paulo/SP e comparou com um modelo estatístico. Como entrada da rede foram utilizados dados de Radiação solar no topo da atmosfera, radiação global, radiação de ondas longas, umidade relativa, vapor d'água, ângulo de elevação solar, ângulo zenital e ângulo de azimuth. Na Tabela 5 é apresentada a validação dos modelos propostos:

Tabela 5: Validação dos modelos propostos por Soares et al.

MODELO	MBE (MJm^{-2})	RMSE (MJm^{-2})	r
$H_d = f(K_t)$	-0,0169	0,193	0,91
RNA	0,0116	0,121	0,94

Fonseca Junior (2011) et al. utilizaram as RNA e a SVM para estimar a insolação em 25 localidades no Japão. As duas técnicas geraram resultados similares para todas as localidades: RMSE variando entre 0,105 kWh/m² e 0,115 kWh/m² e um MAE entre 0,06 kWh/m² e 0,075 kWh/m².

Zanetti et. al. (2008) utilizaram também as RNA para estimar a evapotranspiração para 17 cidades do estado do Rio de Janeiro. Como entrada foram utilizadas as coordenadas geográficas e dados de temperatura e, como referência, foi calculada a evapotranspiração pelo método de Penman-Monteith. O RMSE encontrado variou de 0,2 a 1,2 (mm d^{-1})² e o R² variou de 0,6125 a 0,8466.

3 MATERIAL E MÉTODOS

3.1 Localização e Clima

Botucatu está situada na região centro-sul do estado de São Paulo, em uma região caracterizada pela produção de cana-de-açúcar. Nesta região, entre junho e setembro, é possível verificar um aumento na concentração de aerossóis em suspensão na atmosfera, devido ao início da colheita de cana-de-açúcar e queimadas feitas nas lavouras para facilitar a colheita dos colmos. Neste período, aproximadamente 60% do material particulado PM_{10} (diâmetro $>10 \mu m$) e 64% do carbono negro em suspensão na atmosfera são oriundos da queima de cana-de-açúcar (LARA et al., 2005). Segundo Codato et al. (2008), a maior concentração de aerossóis na atmosfera sobre a região de Botucatu é observada no mês de setembro ($PM_{10} = 70 \mu g/m^3$). E a menor concentração ocorre no mês de maio ($PM_{10} = 28 \mu g/m^3$), resultante basicamente da emissão feita por indústrias e veículos automotores.

O clima de Botucatu é composto de inverno frio e seco e verão quente e úmido. Os meses de fevereiro e julho são o mais quente e o mais frio do ano, respectivamente, com temperaturas médias de $23,21^\circ C$ e $17,05^\circ C$, enquanto os meses de

fevereiro e agosto são o mais e o menos úmido, com percentuais de 77,07% e 61,72%, respectivamente (Figura 13a).

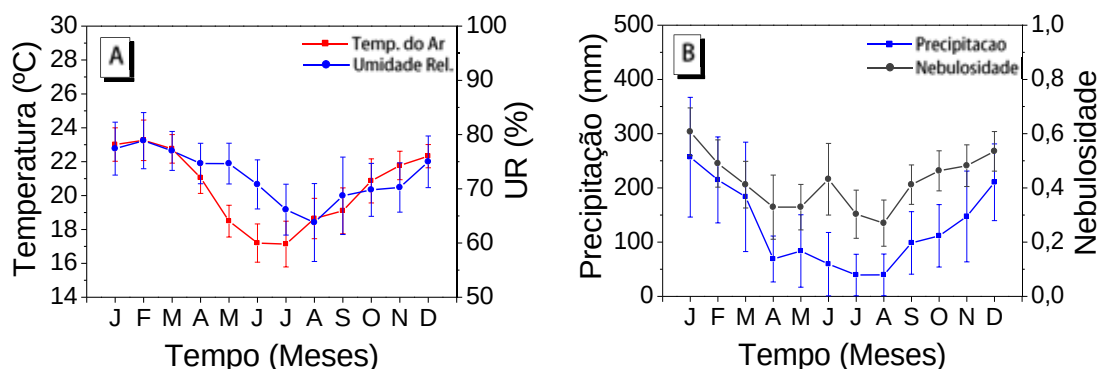


Figura 13: Médias diárias mensais dos parâmetros meteorológicos A) Temperatura do ar e umidade relativa, B) Precipitação e nebulosidade em Botucatu/SP, período de 1971 a 2000.

O ciclo anual de precipitação é constituído de dois períodos bem distintos: o primeiro, de outubro a março, corresponde ao período chuvoso, onde ocorre 75 a 90% do total anual, e o segundo, de abril a setembro, corresponde ao período seco, onde os valores de precipitação encontram-se abaixo de 100 mm (Figura 13b). No primeiro, a ocorrência de chuvas e a elevada nebulosidade são resultantes da formação da Zona de Convergência do Atlântico Sul e dos Sistemas Frontais (CARVALHO et al., 2004). No segundo, a nebulosidade e as chuvas que ocorrem são de média a baixa intensidade e resultantes da passagem de frentes frias (CPTEC, 2010). O maior valor de precipitação acumulada ocorre no mês de janeiro (288,22 mm), onde a razão de nebulosidade é 0,61, e o menor em agosto (40,69 mm), onde a razão de nebulosidade é 0,27. A razão de nebulosidade aqui citada é definida pela equação $\phi = 1 - (n/N)$ e expressa o número de horas que o sol ficou encoberto por nuvens numa razão de insolação n/N , onde n é o número de horas de brilho solar e N é o fotoperíodo.

3.2 Aquisição dos Dados

Os dados utilizados foram obtidos na Estação de Radiometria Solar (Figura 14), localizada na Faculdade de Ciências Agrônômicas de Botucatu, UNESP, com coordenadas geográficas: latitude de 22°54' S, longitude de 48°27' W e altitude de 786m e medidos no período de janeiro de 2000 a dezembro de 2007. Os equipamentos utilizados, bem como suas características são apresentados nas Tabelas 6 e 7. Para a aquisição dos

dados, utilizou-se um Datalogger Campbell 23X, operando na frequência de 0,2Hz e armazenando médias de 5 minutos. Os valores das radiações global e ultravioleta, em MJ.m⁻², foram gerados a partir da integração diária da irradiância (W.m⁻²).

Tabela 6: Equipamentos agrometeorológicos utilizados e suas características.

INSTRUMENTAÇÃO	
PARÂMETRO	MODELO
Insolação	Heliógrafo - Campbell Strokes
Umidade Relativa	Higrógrafo - Filotécnica Milano
Temperatura do ar	Termógrafo - Filotécnica Milano
Precipitação	Pluviógrafo Ota Keiki Seisakusho Co Ltda.

Tabela 7: Característica detalhada dos equipamentos utilizados para medição de radiação solar.

RADIAÇÃO	GLOBAL	UV
Marca	Eppley	Kipp & Zonen
Modelo	PSP	CUV-3
Fator de Calibração	7,45μV/Wm ⁻²	312μV/Wm ⁻²
Alcance Espectral	295-2800nm	290-400nm
Tempo de resposta	1s	5ms
Linearidade	±0,5% (de 0 a 2800 W/m ²)	< 1%
Co-seno	±1% (0<Z<70°) ±3% (70°≤Z<80°)	<±10%
Resposta à temperatura	±1% de -20°C a 40°C	<±0,1/K



Figura 14. Estação de Radiometria Solar de Botucatu.

3.3 Tratamentos dos Dados

Os dados coletados foram submetidos a um controle de qualidade objetivando eliminar valores falsos e ruídos, garantindo assim a qualidade das medidas.

Após a coleta, a primeira filtragem é realizada para detectar dias que não apresentaram medidas por motivos de queda de energia, por exemplo. Depois desse primeiro estágio, são retirados os valores detectados antes do nascer e depois do por do sol e, por último, os dados são integrados no dia, gerando valores diários. Procedimento este realizado através de software computacional, de maneira rápida e eficaz.

Foram mantidos somente os dias em que todos os parâmetros utilizados apresentaram medidas confiáveis. Após a filtragem, 201 dias foram retirados, o que corresponde a 7% do agrupamento total. A descrição do número de dias retirados é apresentada na tabela 8.

Tabela 8: Dias retirados da base de dados.

	2000	2001	2002	2003	2004	2005	2006	2007	Total
Jan	--	--	4	7	--	--	1	--	12
Fev	--	--	3	2	--	--	--	--	5
Mar	--	--	3	--	1	--	--	16	20
Abr	--	--	--	--	2	7	--	--	9
Mai	--	1	--	--	--	--	13	11	25
Jun	--	1	--	1	1	--	13	7	23
Jul	--	1	--	--	--	3	14	--	18
Ago	--	13	--	--	1	--	2	1	17
Set	1	28	1	--	1	--	--	--	31
Out	1	13	3	--	--	--	--	--	17
Nov	--	--	1	--	5	--	5	1	12
Dez	--	8	--	2	1	1	--	--	12
Total	2	65	15	12	12	11	48	36	201

3.4 Composição dos Agrupamentos

Os dados das variáveis meteorológicas utilizados foram divididos em dois agrupamentos. O primeiro agrupamento foi composto por 75% dos dados, período de janeiro de 2000 a dezembro de 2005, o qual foi utilizado para o treinamento da técnica. O segundo agrupamento, que compreende o período de janeiro de 2006 a dezembro de 2007, foi utilizado para testar a eficiência do modelo desenvolvido na estimativa da radiação solar ultravioleta. Após a divisão da base de dados foram criados 5 grupos, com diferentes combinações de variáveis para a alimentação das técnicas de aprendizado de máquina, apresentados na tabela 9.

Tabela 9: Grupos de variáveis utilizadas como entrada para as técnicas de aprendizado.

GRUPOS	COMPONENTES
Grupo 1	Insolação, umidade relativa, temperatura do ar, precipitação, fotoperíodo, radiação no topo da atmosfera e radiação ultravioleta no topo da atmosfera.
Grupo 2	Radiação solar global.
Grupo 3	Radiação solar no topo da atmosfera, radiação solar ultravioleta no topo da atmosfera e radiação solar global.
Grupo 4	Radiação solar no topo da atmosfera, radiação solar ultravioleta no topo da atmosfera, radiação solar global, transmissividade atmosférica (Kt) e a massa ótica relativa.
Grupo 5	Insolação, umidade relativa, temperatura do ar, precipitação, fotoperíodo, radiação solar no topo da atmosfera, radiação solar ultravioleta no topo da atmosfera, radiação solar global, transmissividade atmosférica (Kt) e massa ótica relativa.

O primeiro grupo foi criado com o intuito de estimar a radiação solar ultravioleta sem a dependência da medição de dados de radiação solar. Assim, este grupo utilizou as variáveis: insolação, umidade relativa, temperatura do ar, precipitação, fotoperíodo, radiação no topo da atmosfera e radiação ultravioleta no topo da atmosfera, sendo que os dois últimos parâmetros foram calculados a partir das equações 1 e 2, respectivamente (IQBAL, 1983):

$$H_0 = 10443E_0 \sin(\phi) \sin(\delta) \left[\frac{\pi h^*}{180} - \tan(h) \right] \quad (1)$$

$$H_{0UV} = 0,057059 H_0 \quad (2)$$

onde E_0 é o fator de correção da excentricidade da órbita terrestre, h é o ângulo horário, expresso em graus; h^* é o ângulo horário, expresso em radianos; ϕ é a latitude local, expressa em graus; δ é a declinação solar, expressa em graus.

O grupo 2 utilizou apenas dados da radiação solar global, sendo que a maior vantagem deste grupo é a estimativa da radiação solar ultravioleta a partir de uma única variável de entrada.

O terceiro grupo é composto pela radiação solar no topo da atmosfera, radiação solar ultravioleta no topo da atmosfera e radiação solar global. A utilização da radiação solar no topo da atmosfera permite às técnicas analisar a transmissividade atmosférica da radiação global generalizando para a ultravioleta, podendo ser utilizado em locais com condições climáticas semelhantes.

Em relação ao terceiro grupo, o quarto contém duas variáveis a mais, ambas calculadas: a transmissividade atmosférica (K_t), que permite às técnicas avaliar as condições de cobertura do céu por nuvens, e a massa ótica relativa, que fornece referências às técnicas sobre o caminho óptico atmosférico percorrido pela radiação solar relacionado à movimentação aparente do Sol. As duas variáveis foram calculadas, respectivamente, pelas equações 3 e 4 (IQBAL, 1983; KASTEN; YOUNG, 1989).

$$K_t = \frac{H_G}{H_0} \quad (3)$$

$$m_r = \frac{1}{[\cos(\theta_z) + 0,050572(96,0795 - \theta_z)^{-1,6364}]} \quad (4)$$

onde θ_z é o ângulo zenital, expresso em radianos.

O quinto e último grupo reúne todas as variáveis apresentadas anteriormente, ou seja, é composto por: insolação, umidade relativa, temperatura do ar, precipitação, fotoperíodo, radiação solar no topo da atmosfera, radiação solar ultravioleta no topo da atmosfera, radiação solar global, transmissividade atmosférica e massa ótica relativa.

Para avaliação do desempenho da estimação da radiação solar UV diária pelas TAM, através de cada um dos 5 grupo de variáveis, foram utilizados os seguintes índices estatísticos: coeficiente de determinação (R^2), mean bias error (MBE), root mean square error (RMSE) e “d” de Willmott, calculados a partir das equações 5, 6, 7 e 8 (IQBAL, 1983; WILLMOTT et al., 1985).

$$R^2 = \frac{\sum (Y_c - \bar{Y})^2}{\sum (Y_m - \bar{Y})^2} \quad (5)$$

$$MBE = \frac{\sum (Y_c - Y_m)}{n} \quad (6)$$

$$RMSE = \frac{\sqrt{\sum (Y_c - Y_m)^2}}{n} \quad (7)$$

$$d = 1 - \left[\frac{\sum_{i=1}^n (Y_c - Y_m)^2}{\sum_{i=1}^n (|Y_c - \bar{Y}| + |Y_c - \bar{Y}_c|)^2} \right] \quad (8)$$

onde Y_c são os valores de radiação calculados, Y_m os medidos, $\bar{Y}$ a média dos valores medidos, $\bar{Y}_c$ a média dos valores calculados e n o número de amostras.

Após a validação dos dados, os resultados foram comparados com os trabalhos de Leal (2011) e Barbero (2006) que utilizaram a RNA para estimar a UV e com os trabalhos de Corrêa (2012) e Escobedo (2009) que propuseram modelos estatísticos para a cidade de Botucatu. Esses modelos estatísticos foram utilizados para estimar a radiação UV para os anos de 2006 e 2007 e posteriormente validados a partir dos mesmos índices estatísticos.

3.5 Técnicas de Aprendizado de Máquina Utilizadas

No presente trabalho foram utilizadas duas TAM em dois sistemas operacionais distintos: As redes neurais artificiais em ambiente Windows 7 Ultimate® e a máquina de vetores de suporte em ambiente Linux Umbutu 2.6.32-21®.

A RNA utilizada foi a FannTool 1.1, uma compilação gratuita e livre, disponível no site: <http://code.google.com/p/fanntool/>.

A SVM utilizada foi a LibSVM 3.1, uma compilação também gratuita e livre, disponível no site: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.

3.6 Utilização das Técnicas de Aprendizado de Máquina

O procedimento para uso de ambas as TAM é similar, diferindo somente na organização dos dados de entrada. Faz-se importante ressaltar que ambos os conjuntos (treinamento e teste) devem seguir os padrões exigidos pela TAM utilizada.

Para a FannTool os dados devem estar em um documento de texto, a primeira linha deve conter o número de linhas do documento, o número de variáveis de entrada e o número de variáveis de saída, separados por um espaço. As próximas linhas devem ser alternadas entre os dados de entrada (separados por espaço e com casas decimais separadas por ponto e não por vírgula) e os dados de saída (Figura 15).

```

10 6 1
0 91.44068 9.5 13.3778 42.83006 2.44384
0.37826
0 01.44668 9.8 13.37246 42.77467 2.44068
0.42856
4.9 82.10944 24.3 13.36661 42.80127 2.4422
0.86623
2.6 86.51877 22 13.36025 42.70928 2.43695
0.71223
0 01.46788 24.3 13.35341 42.76424 2.44008
0.40256
0.7 90.09057 3.3 13.34607 42.74258 2.43885
0.68856
6 79.92995 5 13.33824 42.71879 2.43749
0.97894
6.7 78.58798 0 13.32994 42.69287 2.43601
0.97803
5.4 81.08021 2.3 13.32117 42.68332 2.43147
0.91721
10.5 71.31714 0 13.31193 42.68212 2.43255
1.1705
7.9 78.3016 3.3 13.30223 42.60078 2.43076
0.97811
8.3 75.55596 0 13.29209 42.57119 2.42907
1.08734
6.9 78.20456 0 13.2815 42.53552 2.42703
1.10399
5.9 80.12166 0 13.27048 42.41276 2.42003
1.04083
6 79.92995 0 13.25903 42.45733 2.42257
0.99827
12 68.42736 0 13.24716 42.41101 2.41993
1.20561
10.8 71.01766 0 13.22221 42.32112 2.4148
1.17392
4.8 82.54853 0.5 13.20914 42.26872 2.41181
1.06058
5.9 80.42558 9.5 13.19569 42.1418 2.40457
0.97227
10.7 71.43555 0 13.18186 42.16488 2.40589
1.20531
9.8 71.16094 0 13.16767 42.10936 2.40272
1.16423

```

Figura 15: Planilha de dados para inserção na FannTool.

Para a libSVM o documento deve iniciar com os dados de saída seguido dos dados de entrada numerados e seguidos de dois pontos (:)

Ex. saída 1:entrada1 2:entrada2...

Como na FannTool, os dados devem estar em um documento de texto, separados por um espaço e as casas decimais são separadas por ponto e não vírgula (Figura 16).

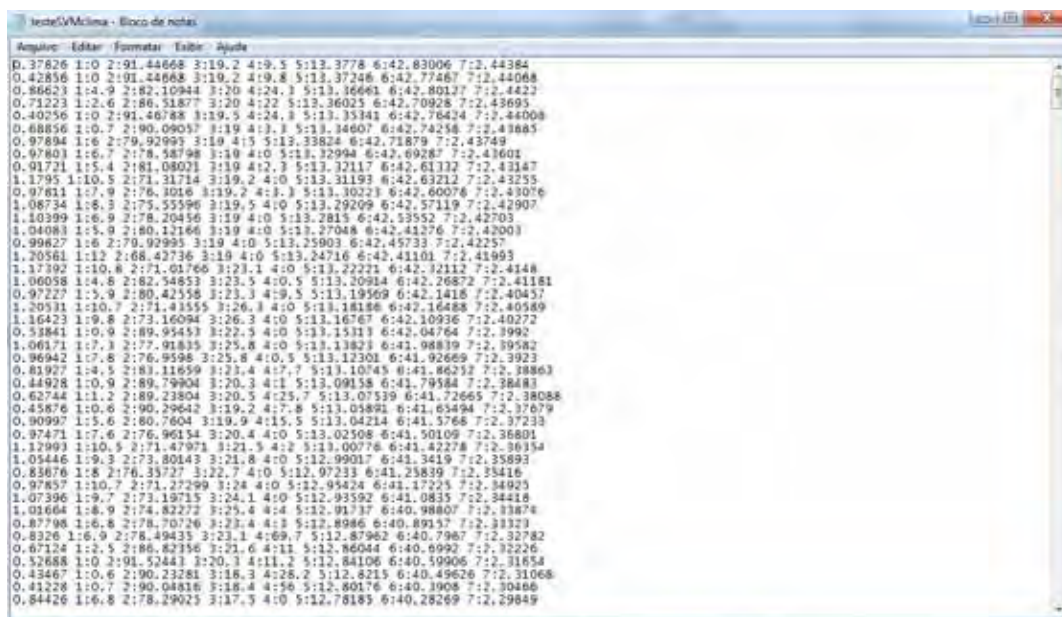


Figura 16: Planilha de dados para inserção na LibSVM.

Com os dados já preparados deve-se iniciar a TAM desejada e adicionar os conjuntos de treinamento no local indicado;

Os resultados variarão de acordo com o ajuste dos parâmetros de cada TAM, através de tentativa e erro;

Após o ajuste dos parâmetros basta iniciar o treinamento para que a TAM devolva um arquivo contendo o resultado do treinamento. Esse arquivo contém todo o comportamento que a TAM foi capaz de abstrair dos dados de entrada;

Após esses passos inicia-se a segunda fase, a fase de testes. Aqui se faz necessário carregar o arquivo que a TAM gerou juntamente com o conjunto de dados de teste e iniciar os testes. Esse procedimento gerará um documento que irá conter os resultados estimados.

3.7 Tipo de Aprendizado

O processo de aprendizagem utilizado foi o supervisionado, ou seja, os dados de treinamento compostos pelos agrupamentos apresentados na tabela 9, foram inseridos na máquina juntamente com sua respectiva saída (a radiação solar ultravioleta medida na estação de radiometria solar), para o ajuste dos parâmetros livres do sistema. A vantagem deste método é que, ao incluir as saídas desejadas, os acertos serão valorizados e os erros penalizados, aumentando assim a acurácia da predição.

4 RESULTADOS E DISCUSSÃO

Para cada grupo foram gerados dois gráficos, sendo que o primeiro demonstra o comportamento dos dados medidos e estimados em função do tempo, em dias corridos, sendo que o primeiro valor (1) corresponde ao dia 01/01/2006 e o último valor (643) ao dia 31/12/2007. O segundo gráfico de cada grupo apresenta a correlação entre os dados medidos e os dados estimados pela técnica em questão. Este gráfico contém uma reta vermelha referente à regressão linear e uma reta azul de 45° que simula a correlação ideal dos dados. Para a validação dos modelos foram utilizados os indicativos estatísticos MBE e RMSE, em porcentagem, além dos índices d de Wilmott e R^2 . Ao final, na Tabela 10, são apresentados os resultados gerados para cada um dos grupos.

As Figuras 17a, 17c e 17e apresentam a curva da radiação solar ultravioleta medida na estação radiométrica associada às curvas da radiação solar ultravioleta diária estimada através da RNA, SVM-Linear e SVM-RBF, respectivamente, para o grupo 1, em função do tempo (dias corridos) e as Figuras 17b, 17d e 17f apresentam as curvas dos valores estimados contra os valores medidos.

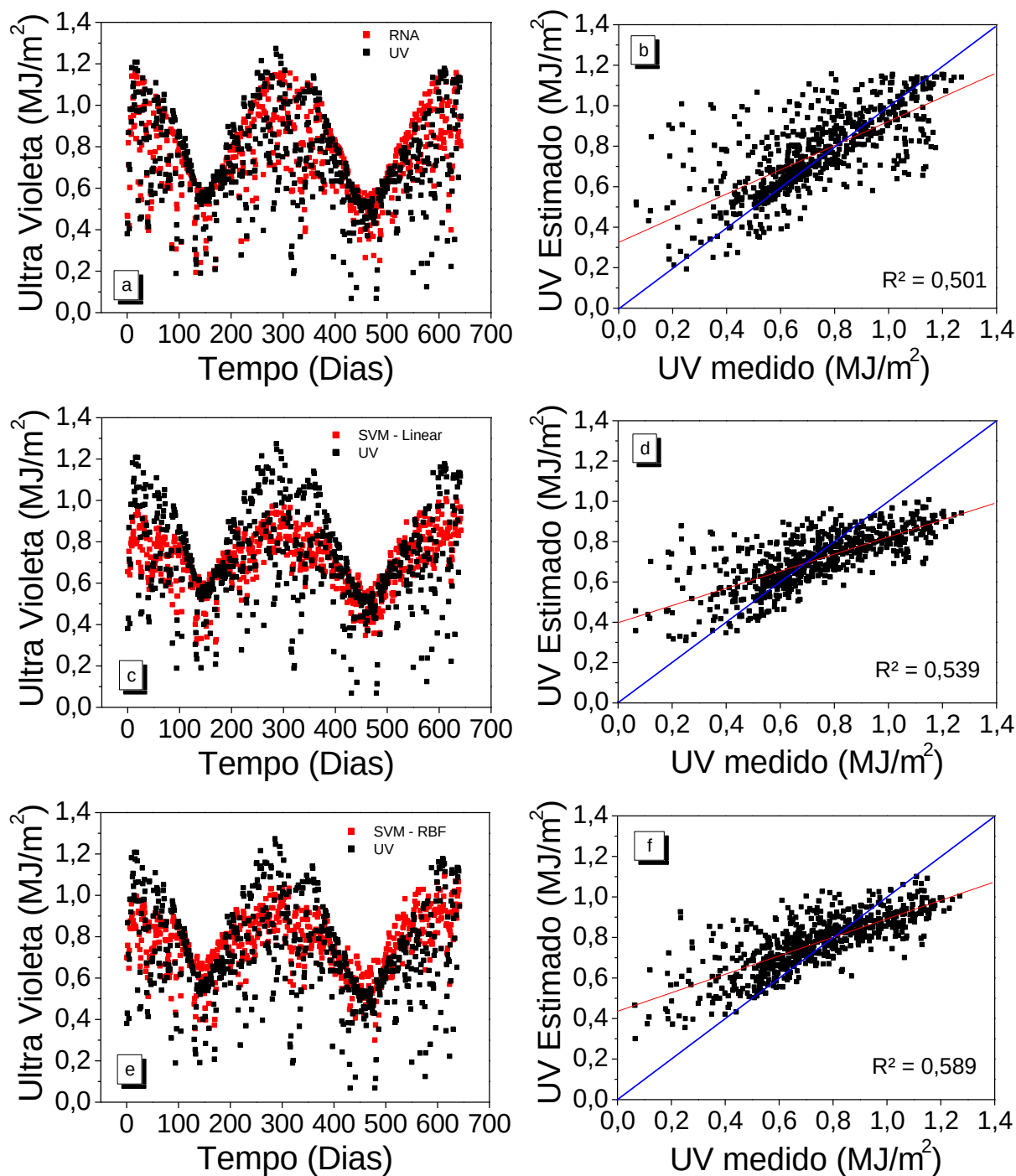


Figura 17. Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 1. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.

Verifica-se na Figura 17 que a estimação da radiação solar UV a partir do grupo 1 apresentou baixa precisão para as três técnicas, com tendência a subestimar os valores durante o verão e superestimá-los durante o inverno. Observou-se para este grupo um coeficiente de determinação R^2 inferior a 60% (Tabela 10). A baixa precisão na estimação pode ser explicada pela baixa correlação entre as variáveis de entrada com a radiação solar

UV. A transmissividade atmosférica da radiação solar UV é influenciada principalmente pelas condições de nebulosidade, concentração de aerossóis e ozônio presentes na atmosfera (FOYO-MORENO *et al.*, 2003; VILLÁN *et al.*, 2010). Considerando apenas o efeito das nuvens, a máxima atenuação da radiação solar UV ocorre na presença de nuvens espessas e baixas e com ocorrência de precipitação (JOSEFSSON; LANDELIUS, 2000). Dentre as variáveis utilizadas, a umidade relativa do ar, precipitação e insolação indicam indiretamente as condições de nebulosidade, porém, a influência do ozônio e dos aerossóis na transmissividade atmosférica da radiação solar UV pouco pôde ser evidenciada através das variáveis de entrada utilizadas.

Em relação ao grupo 2, conforme Figura 18, verifica-se uma maior precisão na estimativa quando comparado ao grupo 1. O coeficiente de determinação R^2 para este grupo foi de aproximadamente 91% (Tabela 10).

Verifica-se também, conforme Figuras 18a, 18c e 18e que houve superestimação dos valores para o ano de 2007. Possivelmente, as condições atmosféricas para este ano foram atípicas em relação às dos 6 anos utilizados no agrupamento de treino, havendo uma menor transmissividade atmosférica da radiação solar UV. Uma das possíveis razões seria o aumento na concentração de aerossóis e de ozônio na atmosfera durante o período seco do ano, associado a condições de baixa nebulosidade e umidade do ar, diminuindo os valores de UV. Como as TAM estimaram a UV somente a partir da radiação global, o valor estimado foi elevado quando comparado ao valor medido (real).

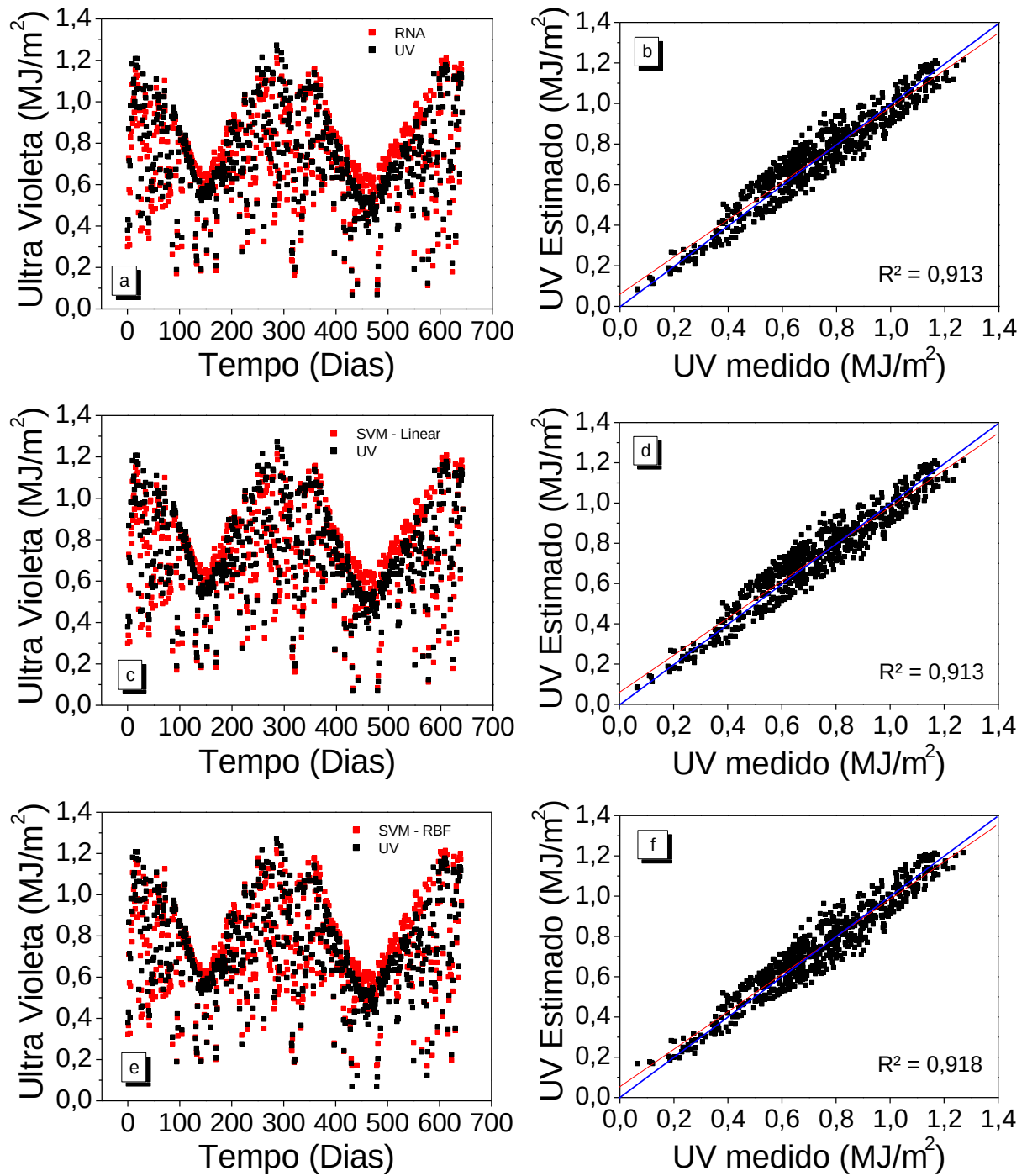


Figura 18. Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 2. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.

No grupo 3, verifica-se que o acréscimo das radiações solar do topo da atmosfera e UV do topo da atmosfera, junto da radiação solar global, permitiu uma maior precisão na estimação quando comparado ao grupo 2 (Figura 19).

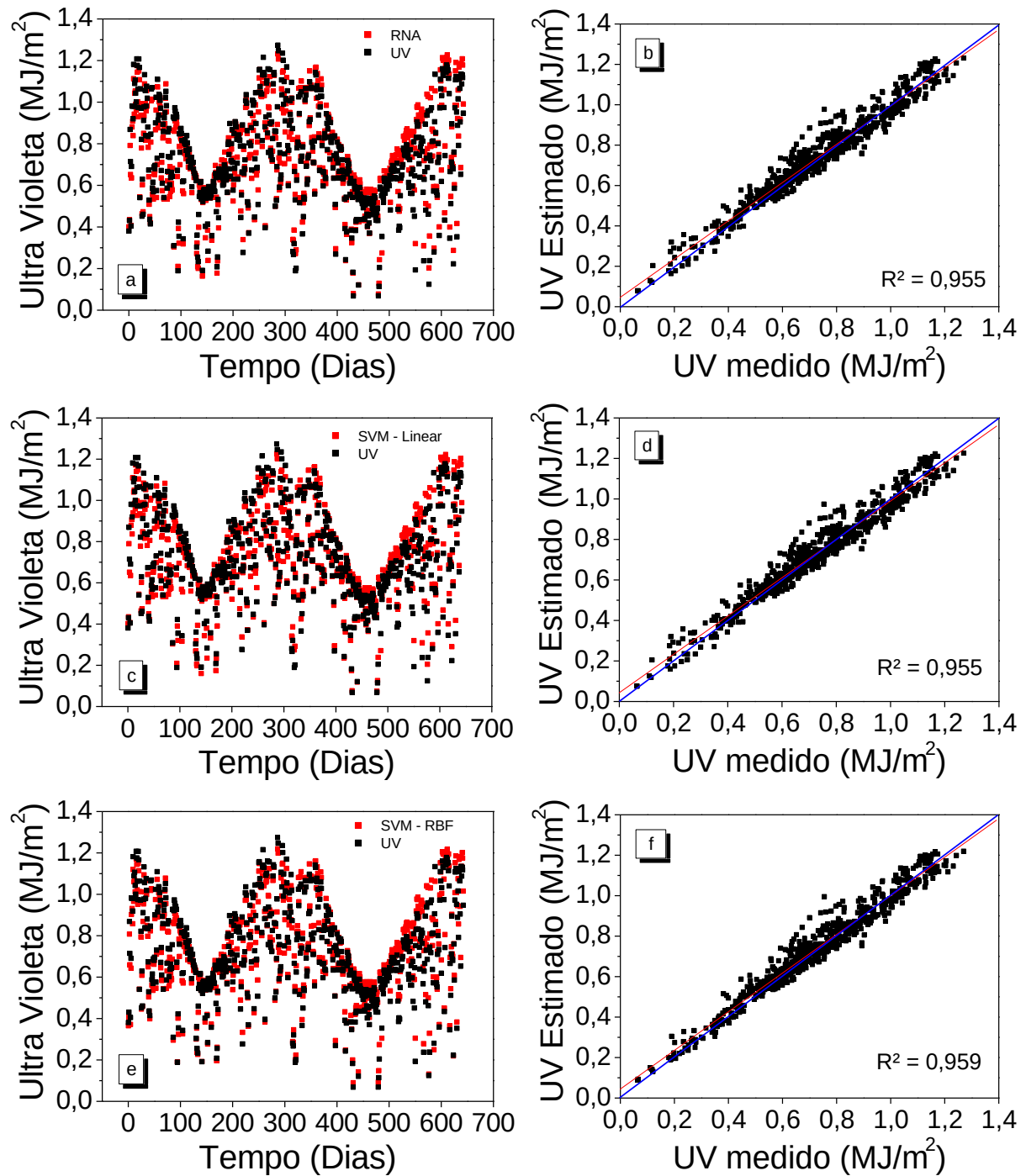


Figura 19. Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 3. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.

O coeficiente de determinação R^2 para este grupo foi superior a 95%, para as 3 técnicas, valores estes superiores em 5% ao grupo 2, (Tabela 10). O acréscimo das duas variáveis permitiu às técnicas estimar o padrão de variação da transmissividade atmosférica da radiação solar UV, associando este ao padrão da radiação global. Também foi

observada a superestimação nos valores da UV em 2007, porém em menor escala se comparado ao grupo 2.

Para o grupo 4, o acréscimo da massa ótica relativa e da transmissividade atmosférica proporcionou uma pequena melhoria na precisão de estimação em relação ao grupo 3, conforme Figura 20.

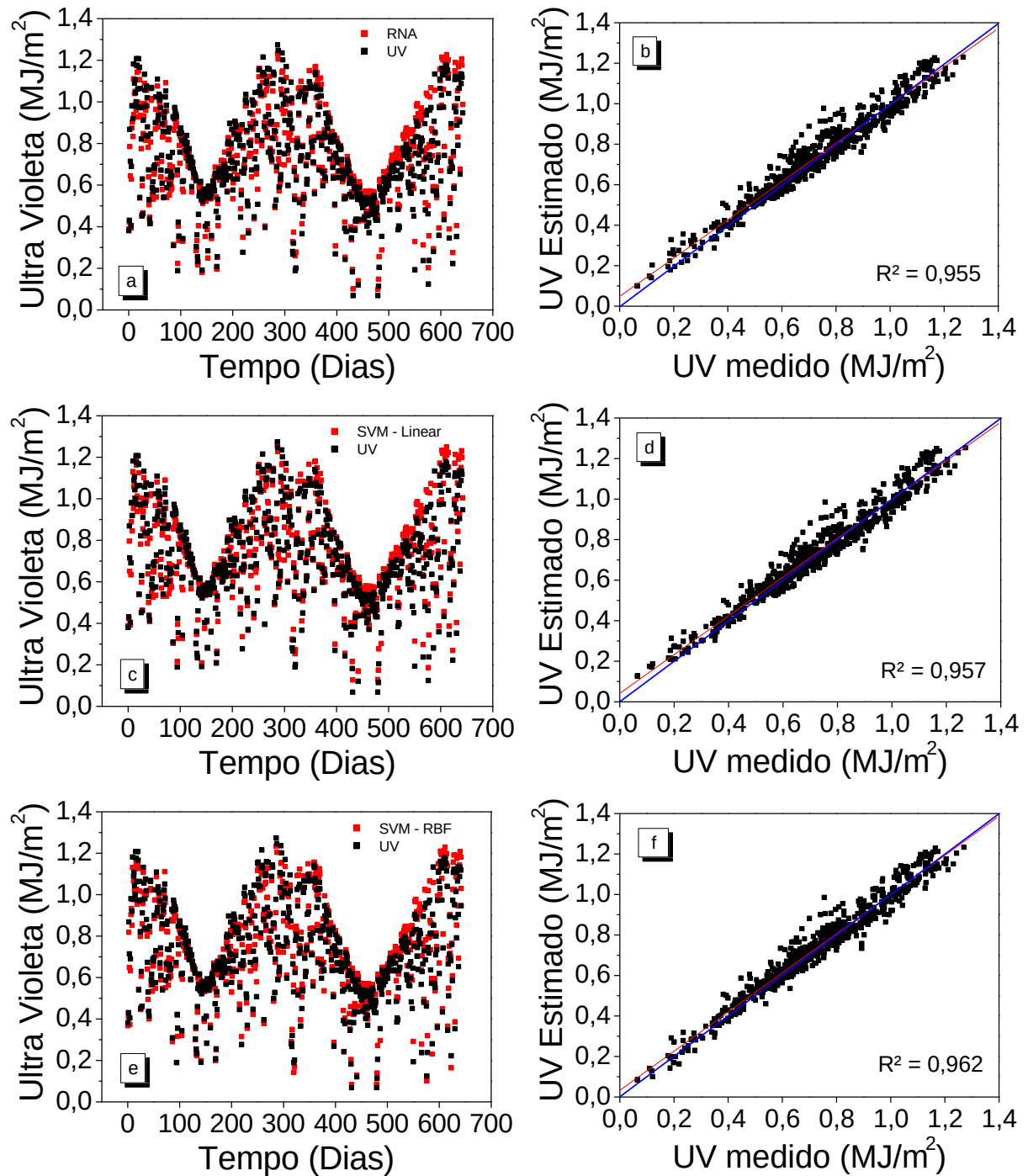


Figura 20: Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 4. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.

O coeficiente de determinação R^2 foi de 95% para a RNA e SVM-Linear e 96% para a SVM-RBF (Tabela 10). Considerando que a radiação global e a radiação solar no topo da atmosfera são precursores do Kt, o acréscimo deste pouco influenciou na precisão sendo que a pequena melhoria foi, possivelmente, influência da massa ótica.

Dentre os grupos de variáveis utilizados, verifica-se no grupo 5 o melhor desempenho na estimativa da radiação solar UV para as três técnicas (Figura 21).

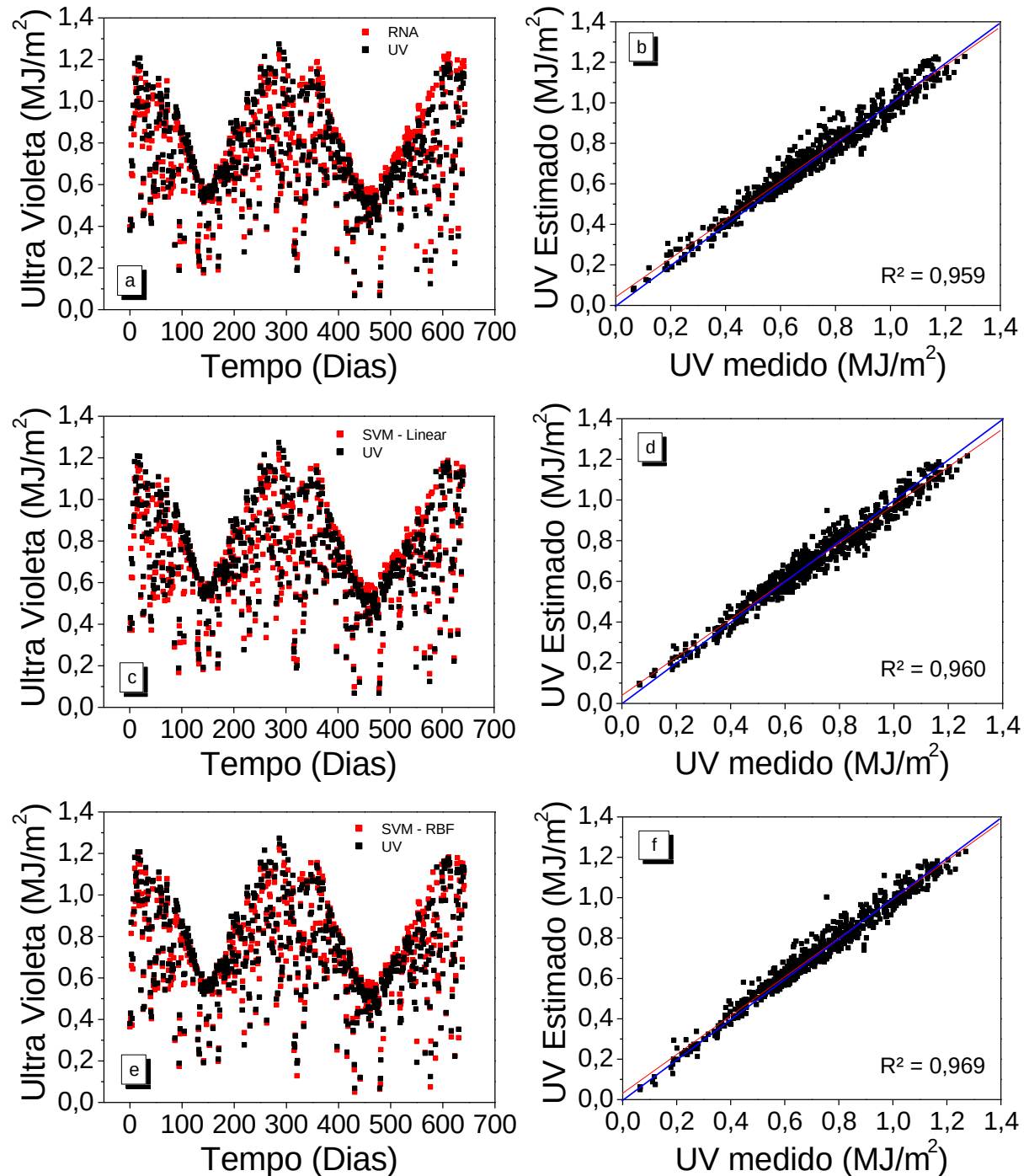


Figura 21. Resultados gerados pela RNA (a), SVM-Linear (c) e SVM-RBF (e) para o grupo 5. b),d),f) correlação entre os dados medidos e estimados, respectivamente, para as três técnicas.

Os índices estatísticos R^2 e “d” foram elevados (superiores a 95% e 0,98 respectivamente) e os índices MBE e RMSE foram baixos quando comparado aos grupos anteriores (inferiores a 0,98% e 6,7%), indicando maior precisão (Tabela 10). Tanto a análise pontual, quanto a geral destes modelos se mostram mais precisas e é possível notar uma redução na tendência de superestimação observada nos demais grupos para o ano de 2007. Podemos deduzir que a associação de todos os parâmetros meteorológicos utilizados permitiu às técnicas uma visão ampla sobre a influência do clima na radiação solar UV, o que não foi possível nos modelos anteriores.

A validação completa de todos os modelos gerados é apresentada na tabela 10:

Tabela 10: Avaliação dos modelos gerados.

ARQUITETURA	MBE (%)	RMSE (%)	D DE WILMOTT	R^2
RNA – Grupo 1	4,50	23,77	0,830	0,501
RNA – Grupo 2	0,33	9,55	0,977	0,913
RNA – Grupo 3	1,09	6,97	0,988	0,955
RNA – Grupo 4	1,17	7,01	0,988	0,955
RNA – Grupo 5	0,86	6,62	0,989	0,959
SVM – Linear – 1	-2,73	22,74	0,784	0,539
SVM – Linear – 2	0,19	9,54	0,977	0,913
SVM – Linear – 3	0,57	6,89	0,988	0,955
SVM – Linear – 4	1,32	6,89	0,988	0,957
SVM – Linear – 5	-0,99	6,58	0,989	0,960
SVM – RBF – 1	5,56	22,23	0,807	0,589
SVM – RBF – 2	0,46	9,311	0,978	0,918
SVM – RBF – 3	1,61	6,75	0,989	0,959
SVM – RBF – 4	1,25	6,46	0,990	0,962
SVM – RBF – 5	0,32	5,71	0,992	0,969

Nota-se que, para todas as TAM, os melhores resultados foram obtidos no grupo 5, com exceção do MBE para a RNA e SVM-Linear, obtido pelo grupo 2.

Com exceção do primeiro agrupamento, os índices MBE e RMSE, obtidos para todas as TAM, foram inferiores a 2% e 10%, respectivamente, e os índices d de Willmott e R^2 superiores a 0,97 e 0,91, respectivamente.

Considerando apenas o agrupamento 5, por apresentar os melhores resultados, a divergência entre os valores dos índices estatísticos das três TAM, é menor que 1%, demonstrando a similar eficácia das três técnicas para a estimativa da radiação solar UV para a cidade de Botucatu.

4.1 Comparação com a Literatura

Com o objetivo de verificar se os resultados obtidos foram satisfatórios, foi feita uma comparação entre os modelos gerados (Tabela 10), trabalhos encontrados na literatura que estimaram UV utilizando TAM (Tabela 1) e trabalhos que utilizaram modelos estatísticos para estimar a UV para a cidade Botucatu.

4.1.1 Trabalhos que utilizaram TAM

Dentre os resultados obtidos para a cidade de Pesqueira, uma cidade de clima semiárido, com verão quente e seco e inverno ameno e chuvoso, os modelos gerados pela RNA apresentaram MBE variando entre -1,3% e -1,6% e RMSE entre 4,6% e 4,8%. Os modelos gerados pela RNA para a cidade de Araripina, uma cidade de clima semiárido, com verão quente e seco e inverno ameno e chuvoso, apresentaram MBE entre 0,7% e 2,0% e RMSE entre 4,6% e 4,9%. Para Almeria, na Espanha, uma cidade litorânea com verão quente e inverno ameno, a RNA gerou um MBE = 0,0% e RMSE = 4,4% e os modelos para a cidade de Golden, Colorado/EUA, uma cidade de clima continental úmido, com verão ameno e inverno frio e com presença de neve, a RNA gerou MBE variando entre -2,0% e -3,0% e RMSE entre 5,1% e 6,5%.

Considerando somente os grupos 2, 3 e 4 por utilizarem somente dados de radiação como os modelos acima, para Botucatu, uma cidade com verão quente e úmido e inverno frio e seco, a RNA gerou MBE variando entre 0,3% e 1,1% e RMSE entre 7,0 e 9,5%; a SVM-Linear apresentou MBE variando entre 0,2% e 1,3% e RMSE entre 6,9% e 9,5% e a SVM-RBF apresentou MBE variando entre 0,5% e 1,6% e RMSE entre 6,5% e 9,3%.

Os três modelos citados para a cidade de Botucatu apresentaram MBE ligeiramente inferior e RMSE ligeiramente superior aos modelos encontrados na literatura, o que comprova a eficácia das técnicas na geração dos modelos para esta cidade.

A divergência dos resultados pode ser resultante da diferença entre os dados de entrada utilizados pelas TAM, pela filtragem aplicada aos mesmos durante a preparação para o uso e devido à diferença entre o clima das cidades.

4.1.2 Trabalhos que utilizaram modelos estatísticos para Botucatu

Para a cidade de Botucatu foram criados alguns modelos estatísticos com o objetivo de estimar a radiação UV. Corrêa et al. (2012) propôs dois modelos para estimá-la, ambos a partir da equação de Angström-Prescott, onde foi feita uma correlação com o número de horas de brilho solar (n) e com o fotoperíodo (N) formando a razão de insolação (n/N) e Escobedo et al. (2009) propôs um modelo a partir da radiação global, de acordo com a cobertura de céu. Tais modelos, acompanhados de suas respectivas validações, são apresentados na tabela 11.

Tabela 11: Modelos estatísticos gerados para a cidade de Botucatu

MODELO	EQUAÇÃO	MBE (%)	RMSE (%)	d	R ²
Corrêa 1	$0,234 + 0,279 (n/N)$	3,42	28,15	0,756	0,330
Corrêa 2	$0,190 + 0,590 (n/N) - 0,316 (n/N)^2$	3,42	29,37	0,756	0,319
Escobedo	$0.0416(H_G)$	1,51	9,85	0,977	0,913

Considerando somente o agrupamento 3 por utilizar H_G , H_0 e H_{0UV} , sendo somente a H_G medida, a RNA apresentou MBE = 1,09%, RMSE = 6,97%, d de Willmott = 0,988 e $R^2 = 0,954$; a SVM-linear apresentou MBE = 0,57%, RMSE = 6,89% d de Willmott = 0,989 e $R^2 = 0,955$ e a SVM-RBF apresentou MBE = 1,61%, RMSE = 6,75% d de Willmott = 0,989 e $R^2 = 0,959$.

Ao analisar a tabela 11 pode-se notar que todos os índices gerados a partir dos modelos estatísticos foram inferiores aos índices gerados a partir dos modelos que utilizaram TAM, demonstrando que a eficácia das técnicas na geração dos modelos para a cidade de Botucatu é também superior à eficácia dos modelos tradicionais.

5 CONCLUSÃO

Este trabalho permite concluir que as técnicas de aprendizado de máquina são eficazes alternativas para a estimativa de séries temporais da radiação solar ultravioleta.

As TAM se mostraram capacitadas para esta tarefa atingindo resultados próximos entre elas, contudo as redes neurais artificiais apresentaram os piores resultados (maior $R^2 = 0.959$) quando comparados com a máquina de vetores de suporte utilizando função linear (maior $R^2 = 0.960$) e a vetores de suporte utilizando funções de base radial (maior $R^2 = 0.969$).

Conclui-se também que a escolha de variáveis para a alimentação das TAM está fortemente conectada aos resultados gerados e, por isso, merece grande atenção. Os melhores resultados foram observados no grupo 5. Este grupo foi composto por uma gama de variáveis que permitiram às técnicas abstrair características não só do comportamento da radiação como também da variação climática da cidade de Botucatu.

A TAM que fez melhor uso deste grupo foi a SVM-RBF, provavelmente devido ao ajuste das funções de base radial ser maior do que o da função linear. Para o grupo 5 a SVM-RBF apresentou resultados de MBE = 0,31%, RMSE = 5,71%, d de Willmott = 0,992 e $R^2 = 0,969$. O mesmo agrupamento foi responsável, também, pelos melhores resultados das outras duas técnicas: SVM-linear com resultados de MBE = -0,99%,

RMSE = 6,58%, d de Willmott = 0,989 e $R^2 = 0,960$ e RNA com resultados de MBE = 0,86%, RMSE = 6,62%, d de Willmott = 0,989 e $R^2 = 0,959$.

Contudo, a estimação da radiação solar UV a partir de dados de radiação solar global, radiação solar do topo da atmosfera e radiação UV no topo da atmosfera (parâmetros utilizados no grupo 3) apresenta desempenho próximo ao do grupo 5 para todas as técnicas destacando-se por permitir a estimação necessitando da medição de apenas uma variável: a RNA apresentou MBE = 1,09%, RMSE = 6,97%, d de Willmott = 0,988 e $R^2 = 0,954$; a SVM-linear apresentou MBE = 0,57%, RMSE = 6,89% d de Willmott = 0,989 e $R^2 = 0,955$ e a SVM-RBF apresentou MBE = 1,61%, RMSE = 6,75% d de Willmott = 0,989 e $R^2 = 0,959$.

A comparação entre trabalhos encontrados na literatura mostrou que os resultados gerados foram similares aos obtidos em outras cidades que também utilizaram TAM e superiores aos resultados obtidos por modelos estatísticos para a cidade de Botucatu.

Uma possibilidade para a melhoria da precisão na estimação da UV, utilizando TAM, seria a inserção de dados de ozônio e aerossóis medidos por satélite, visto que estes componentes influenciam diretamente o valor da radiação em questão. Porém como um grande número de variáveis não representa necessariamente uma melhoria nos resultados, faz-se viável a aplicação de técnicas de seleção de características, responsáveis por selecionar as variáveis que virão a influenciar positivamente na estimação e, conseqüentemente, melhorar a eficácia computacional.

6 REFERÊNCIAS

- ABDULAZEEZ, M. A. Artificial neural network estimation of global solar radiation using meteorological parameters in Gusau, Nigeria. **Arch. Appl. Sci. Res**, v.3 p. 586-595, 2011.
- AGAR, N. S.; HALLIDAY, G. M.; BARNETSON, R. C.; ANANTHASWAMY, H. N.; WHEELER, M.; JONES A. M. The basal layer in human squamous tumors harbours more UVA than UVB finger-print mutations: A role for UVA in human skin carcinogenesis, **Proceedings of the National Academy of Sciences**, USA, 101, 4954–4959, 2004.
- ALES, V. T. **O Algoritmo Sequential Minimal Optimisation para resolução do problema de support vector machine**: Uma técnica para reconhecimento de padrões. 2008. 150 p. Dissertação (Mestrado em Ciências da Computação). Universidade Federal do Paraná, Curitiba, 2008.
- ALTRAN, A. B. **Aplicação de funções de base radial em problemas de previsão de cargas elétricas via redes neurais artificiais**. 2005. 112p. Dissertação (Mestrado em Engenharia Elétrica). Universidade Estadual Paulista, Ilha Solteira, 2005.
- ANDREOLA, R.; HAERTEL, V. Support Vector Machines na Classificação de Imagens Hiperespectrais. In: XIV Simpósio Brasileiro de Sensoriamento Remoto, 2009, Natal. **Anais XIV Simpósio Brasileiro de Sensoriamento Remoto**, Natal, Brasil. p. 25-30, 2009.
- BACHELOR, M. A.; BOWDEN, G. T. UVA-mediated activation of signalling pathways involved in skin tumour promotion and progression, **Seminars in Cancer Biology**, 14, 131–138, 2004.
- BARBERO F. J., LÓPEZ, G., BATLLES F. J. Determination of daily solar ultraviolet radiation using statistical models and artificial neural networks. **Ann.Geophys.**, 24, 2105–2114, 2006.
- BASAK, D.; PAL, S.; PATRANABIS, D. C. Support vector regression. **Neural Information Processing**. v.11. n.10, p.203-224, 2007.
- BLATCHLEY, E. R., QO-QUANG, Z., JANEX, M. L.; LAÎNÉ, J. M. Process modeling of ultraviolet disinfection. **Water Science and Technology**. v.38, n.6, p.63-69, 1998.
- BOEGER, M. R. T.; POULSON, M. Efeitos da radiação ultravioleta-B sobre a morfologia foliar de *Arabidopsisthaliana* (L.) Heynh. (Brassicaceae). **Acta bot. bras.** 20(2): 329-338, 2006.
- CALBÓ, J.; PAGÈS, D.; GONZÁLEZ, J. P. Empirical studies of cloud effectson UV radiation: a review. **Reviews of Geophysics**. v. 43, p. 1-28, 2005.
- CAÑADA, J.; PEDROS, G.; BOSCA, J. V. Relationships between UV (0.290-.385 μm) and broad band solar radiation hourly values in Valencia and Córdoba, Spain. **Energy**, v. 28, p. 199-217, 2003.

CARVALHO, D. H. D.; BRAGA, A. P. Ajuste da generalização em redes de base radial: uma abordagem multi-objetivo para estimação de parâmetros. In: Simpósio Brasileiro de Redes Neurais, 2004, São Luis, MA. **Anais do VIII Simpósio Brasileiro de Redes Neurais**, 2004.

CARVALHO, L. M. V.; JONES, C.; LIEBMANN, B. The South Atlantic Convergence Zone: Intensity, Form, Persistence, and Relationships with Intraseasonal to Interannual Activity and Extreme Rainfall. **Journal of Climate**, v.17, p.88-108, 2004.

CHEN J.-L.; LIU, H.-B.; WU, W.; XIE, D.-T. Estimation of monthly solar radiation from measured temperatures using support vector machines – A case study. **Renewable Energy**, v. 36, p.413-420, 2011.

CLAIR, T. A., J. EHRMAN, I. KACZMARSKA, A. LOCKE, D. W. TARASICK, K. E. DAY, AND G. MAILLET. Will reduced summer UV-B levels affect zooplankton populations of temperate humic and clearwater lakes. **Hydrobiologia**, v. 462, p. 75-89, 2001.

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. **Theory and Appl. Climat.**, v. 93, p. 57-73, 2008.

CORRÊA F. H. P.; ESCOBEDO, J; DAL PAI, A.; TERAMOTO, E. T. Modelo de Angström- Prescott para estimativa das radiações UV, PAR e IV em Botucatu/SP/Brasil. **Avances en Energías Renovables y Medio Ambiente**, v. 16, p. 61-68, 2012.

CPTEC - Centro de Previsão do Tempo e Estudos Climáticos/INPE. Climanálise - Boletim de monitoramento e análise climática, 2010. Disponível em: <<http://climanalise.cptec.inpe.br/~rclimanl/boletim/>>. Acesso em 24 Jan. 2010.

CUCKER, F.; ZHOU, D. X. Learning Theory. An approximation theory viewpoint. **Cambridge University Press**, Cambridge, 2007.

CUI, X.; GU, S.; ZHAO, X.; WU, J.; KATO, T.; TANG, Y. Diurnal and seasonal variations of UV radiation on the northern edge of the Qinghai-Tibetan Plateau. **Agricultural and Forest Meteorology**. v.48. 144-151, 2008.

DENG, F.; SU, G.; LIU, C.; WANG, Z. Global solar radiation modeling using the artificial neural network technique. In. Asia-Pacific energy engineering conference, 2010, Chengdu, Anais **APPEEC Asia-Pacific energy engineering conference Chengdu, China**, p. 1-5, 2010.

DIAS, G. J. C.; **Planejamento estratégico de um centro de distribuição: Uma aplicação de redes neurais artificiais de funções de bases radiais para previsão de séries temporais**. 2006. 158p. Dissertação (Mestrado em Métodos Numéricos em Engenharia), Universidade Federal do Paraná, Curitiba, 2006.

DORVLO, A. S. S.; JERVASE, J. A.; AL-LAWATI, A. Solar radiation estimation using artificial neural network. **Applied Energy**, v.71 p. 307-319, 2002.

ELHADIDY MA, ABDEL-NABI DY, KRUSS PD. Ultraviolet solar radiation at Dhahram, Saudi Arabia. **Solar Energy**, v. 44, p. 315–9, 1990.

ERUSIAFE, N. E.; CHENDO, M. A. C. Application of artificial neural network in estimating hourly global solar radiation from satellite images. In International Conference on Solar Heating, Cooling and Buildings, 2010, Graz. Anais **EuroSun 2010 - International Conference on Solar Heating, Cooling and Buildings**, Graz, Austria, 2010. CDROM.

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**, v.86 p. 299-309, 2009.

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.

ESCOBEDO, J.; RODRIGUES, D.; DAL PAI, A.; OLIVEIRA, A. P.; SOARES, J. Irradiações UV, PAR e IV médias mensais em Botucatu/SP/Brasil. **Avances en Energías Renovables y Medio Ambiente**, v. 16, p. 85-92, 2012.

FERRÃO, M. F.; MELLO, C.; BORIN, A.; MERETTO, D. A.; POPPI, R. J. LS-SVM: Uma nova ferramenta quimiométrica para regressão multivariada. Comparação de modelos de regressão LS-SVM e PLS na quantificação de adulterantes em leite em pó empregando NIR. **Rev. Quim. Nova**, v. 30, n. 4, p. 852-859, 2007.

FERREIRA, A. A. **Comparação de arquiteturas de redes neurais para sistemas de reconhecimento de padrões em narizes artificiais**. 2004. 107p. Dissertação (Mestrado em Ciências da Computação), Universidade Federal de Pernambuco, Recife, 2004.

FIOLETOV, V. E.; KERR, J. B.; MCARTHUR, L. J. B.; WARDLE, D. I.; MATHEWS, T. W. Estimating UV Index Climatology over Canada. **Journal of Applied Meteorology**, v.42, p.417-433, 2003.

FONSECA JUNIOR, F. G. S.; OOZEKI, T.; TAKASHIMA, T.; OGIMOTO, K. Analysis of the use of support vector regression and neural networks to forecast insolation for 25 locations in Japan. **Anais ISES Solar World Congress**. 2011.

FOYO-MORENO, I.; ALADOS, I.; OLMO, F. J.; ALADOS-ARBOLEDAS, L. The influence of cloudiness on UV global irradiance (295-385 nm). **Agricultural and Forest Meteorology**, v. 120, p. 101-111, 2003.

FOYO-MORENO, I.; VIDA, J.; ALADOS-ARBOLEDAS, L. A Simple All Weather Model to Estimate Ultraviolet Solar Radiation (290-385 nm). **Journal of Applied Meteorology**, v.38 p.1020-1027, 1998.

FOYO-MORENO, I.; VIDA, J.; ALADOS-ARBOLEDAS, L. A simple all weather model to estimate ultraviolet solar radiation (290-385 nm). **Journal of Applied Meteorology**, Lawrence, v. 38, p. 1021-1026, 1999.

GUARNIERI, R. A. **Emprego de redes neurais artificiais e regressão linear múltipla no refinamento das previsões de radiação solar do modelo ETA**. 2006. Dissertação (Mestrado em Meteorologia), Instituto Nacional de Pesquisas Espaciais – INPE, São José dos Campos, 2006.

HAO, X.; HALE, B.A.; ORMROD, D.P. & PAPADOPOULOS, A.P. Effects of pre-exposure to ultraviolet-B radiation on responses of tomato (*Lycopersicon esculentum* cv. New Yorker) to ozone in ambient and elevated carbon dioxide. **Environmental Pollution**, n. 110, p. 217-224, 2000.

HARRISON, J. W.; SMITH, R. E. H. Effects of ultraviolet radiation on the productivity and composition of freshwater phytoplankton communities. **Photochem. Photobiol. Sci.**, v. 8, p. 1218-1232, 2009.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 900p.

IQBAL, M. **An introduction to solar radiation**. London: Academic Press., 1983. 390 p.

JACOVIDES, C.P. ; ASSIMAKOPOULOS, V.D.; TYMVIOSA, F.S. ; THEOPHILOUC, K.; ASIMAKOPOULOS, D.N. Solar global UV (280–380 nm) radiation and its relationship with solar global radiation measured on the island of Cyprus. **Energy**, v. 31, p. 2728–2738, 2006.

JIMÉNEZ, A. E.; ESTRADA, C. A.; COTA, A. D.; ROMÁN, A. Photocatalytic degradation of DBSNa using solar energy. **Solar Energy Materials and Solar Cells**, v.60, p.85-95, 2000.

JOSEFSSON, W. LANDELIUS, T. Effect of clouds on UV irradiance: As estimated from cloud amount, cloud type, precipitation, global radiation and sunshine duration. **Journal of Geophysical Research**, v. 105, n. 4, p. 4927-4935, 2000.

KALASHNIKOVA, O. V.; MILL, F. P.; ELDERING, A.; ANDERSON, D. Application of satellite and ground-based data to investigate the UV radiative effects of Australian aerosols. **Remote Sensing of Environment**, v.107, p. 65-80, 2007.

KASTEN F.; YOUNG A. T. Revised optical air mass tables and approximation formula. **Applied Optics**, v. 28, n. 22, p. 4735-4738, 1989.

KORONAKIS, P. S., SFANTOS, G. K., PALIATSOS, A.G., KALDELLIS, J. K., GAROFALAKIS, J. E., KORONAKI, I.P., Interrelations of UV-global/global/diffuse/solar irradiance components and UV-global attenuation on air pollution episode days in Athens, Greece. **Atmos. Environ.**, v.36, p. 3173– 3181, 2002.

KUDISH, A. I.; IANETZ, A. Analysis of daily clearness index, global and beam radiation for Beer Sheva, Israel: Partition according to day type and statistical analysis. **Energy Conversion and Management**, v.37, p.405-416, 1996.

KYLLING, A.; BAIS, A.; BLUMTHALER, M.; SCHEREDER, J.; ZEREFOS, C.; KOSMIDIS, E. Effect of aerosols on solar UV irradiances during the photochemical activity and solar ultraviolet radiation campaign, **J. Geophys. Res.**, v.103, n. 26, p.26051-26060, 1998.

LARA, L.L. ; ARTAXO, P.; MARTINELLI, L.A.; CAMARGO, P.B.; VICTORIA, R.L.; FERRAZ, E.S.B. Properties of aerosols from sugar-cane burning emissions in Southeastern Brazil. **Atmospheric Environment**, v.39, n. 26, p. 4627-4637, 2005.

LATHA, K. M.; BADARINATH, K. V. S.: Impact of chemical fire accident on spectral solar irradiance over urban environment – A case study aerosol and air. **Quality Research**, v.5, n.1, p.14-24, 2005.

LEAL, S. S.; TÍBA, C.; PIACENTINI, R.; Daily UV radiation modeling with the usage of statistical correlations and artificial neural networks. **Renewable Energy**, v.36, p. 3337-3344, 2011.

LI, D. H. W; LAM, J. C. An analysis of climatic parameters and sky condition classification. **Building and Environment**, v.36, p.435-450, 2000.

LÓPEZ, G.; RUBIO, M. A.; MARTÍNEZ, M.; BATLLES, F. J.; Estimation of hourly global photosynthetically active radiation using artificial neural network models. **Agricultural and Forest Meteorology**, v. 107, p.279–291, 2001.

LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machines. **Revista de Informática Teórica e Aplicada**. v. XVI. n. 2, 2007.

MARTINEZ-LOZANO, J. A.; TENA, F.; UTRILLAS, M. P. Ratio of UV to global broad band irradiation in Valencia, Spain. **International Journal of Climatology**. v. 19, p. 903–911, 1999.

NAZARI, E. M.; AMMAR, D.; BEM, A. F.; LATINI, A.; MÜLLER, Y. M.; ALLODI, S. Effects of environmental and artificial UV-B radiation on freshwater prawn *Macrobrachium olfersi* embryos. **Aquat. Toxicol.**, v. 98, p. 25-33, 2010.

OGUNJOBI K. O.; KIM Y.J. Ultraviolet (0.280-0.400 μm) and broadband solar hourly radiation at Kwangju, South Korea: analysis of their correlation with aerosol optical depth and clearness index. **Atmospheric Research**, v. 71, p. 193-214, 2004.

OKOGBUE, E. C.; ADEDOKUN, J. A.; HOLMGREN, B. Hourly and daily clearness index and diffuse fraction at a tropical station, Ile-Ife, Nigéria. **International Journal of Climatology**, v.29, p.1035-1047, 2009.

OLIVEIRA, O. C. **Desacoplamento de um gerador síncrono através de um controle adaptativo pode modelo de referência baseado em funções de base radial**. 2011. 86p. Dissertação (Mestrado em Engenharia Elétrica e de Computação). Universidade Federal do Rio Grande do Norte. 2011.

PARREIRAS, L. P. R. F. **Modelo genético-neural de gestão de carteiras de ações**. 2003. 127p. Monografia (Graduação em Engenharia de Produção). Universidade de São Paulo. 2003.

PIENITZ, R.; VINCENT, W. F. Effect of climate change relative to ozone depletion on UV exposure in subarctic lakes. **Nature**, v. 404, p. 484-487, 2000.

PORFIRIO, A. C. S.; SOUZA, J. L.; LYRA, G. B.; LEMES, M. A. M. An assessment of the global UV solar radiation under various sky conditions in Maceió-Northeastern Brazil. **Energy**. v. 44. p. 584-592, 2012.

POWELL, M. J. D. Radial basis function for multivariable interpolation. **IMA conference on algorithms for the approximation of functions and data**, RCMS, Shrivenham, UK, p. 143-167, 1985.

RIBEIRO, L. N. **Rede neural com retropropagação: uma aplicação na classificação de imagens de satélite**. 2003. 131p. Dissertação (Mestrado em Ciências da Computação e Estatística). Universidade Federal do Mato Grosso do Sul. 2003.

ROOBA, S. M. A study of ultraviolet solar radiation at Cairo urban area, Egypt. **Solar Energy**, v. 77, p. 251-259, 2004.

ROZEMA, J.; BJÖRN, L.O.; BORNMAN, J.F. GABERSCIK, A.; HÄDER, D.-P.; TROST, T.; GERM, M.; KLISCH, M.; GRÖNIGER, A.; SINHA, R.P.; LEBERT, M. HE, Y.Y.; BUFFONI-HALL, R.N.; BAKKER, V.J. VAN DE STAAIJ, J.; MEIJKAMP, B.B. The role of UV-B radiation in aquatic and terrestrial ecosystems – an experimental and functional analysis of the evolution of UV-absorbing compounds, **J. Photochem. Photobiol.**, v. 66, p. 2-12, 2002.

SABBURG, J. M.; PARISI, A. V.; Spectral dependency of cloud enhanced UV irradiance. **Atmospheric Research**, v.81, p. 206-214, 2005.

SANTOS, J. B.; VILÁN, D. M.; CASTRILLO, A. M. Analysis and cloudiness influence on UV total radiation. **Int. J. Climatol.** v. 31, p. 451–460, 2011.

SCHÖLKOPF, B.; SMOLA A. J. Learning with Kernels. Support Vector Machines. Regularization, Optimization, and Beyond, **MIT Press**, Cambridge, MA, 2002.

SEMOLINI, R. **Support vector machines, inferência transdutiva e o problema de classificação**. Dissertação (Mestrado em Engenharia Elétrica e de Computação) Universidade Estadual de Campinas. 2002.

SILVA, P. L. **Modelagem de superfícies seletivas de frequência e antenas de microfita utilizando redes neurais artificiais**. 2006. Dissertação (Mestrado em Energia Elétrica). Universidade Federal do Rio Grande do Norte. 2006.

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 neural network technique. **Applied Energy**, v. 79, n. 2, p. 201-214, 2004.

SOMMER, R.; CABAJ, A.; HAIDER, T. Microbiological effect of reflected UV radiation in devices for water disinfection. **Water Science and Technology**, v.34, n.7-8, p.173-177, 1996.

SOUZA, R. C. T. **Previsão de séries temporais utilizando rede neural treinada por filtro de Kalman e evolução diferencial**. 2008. 85p. Dissertação (Mestrado em Engenharia de Produção e Sistemas). Pontifícia Universidade Católica do Paraná. 2008.

STEINWART, I.; CHRISTMANN, A. **Support Vector Machines**. Springer, New York, 2008. 601p.

TIAGO, C. M.; LEITÃO, V. M. A. Utilização de funções de base radial em problemas unidimensionais de análise estrutural. In. V Congreso de Métodos Numéricos en Ingeniería, Madrid, Anais **Métodos Numéricos em Ingeniería V**, Madrid, Espanha, 2002. CDROM.

TOSSERAMS, M.; ROZEMA, J. Effects of ultraviolet-B radiation (UV-B) on growth and physiology of the dune grassland species *Calamagrostis epigeios*. **Environmental Pollution**, v. 89, n. 2, p. 209-214, 1995.

VAPNIK, V. **Statistical learning theory**. New York: John Wiley and Sons, 1998. 736 p.

VARO, M.; PEDROS, G.; MARTÍNEZ-JIMENEZ, P. Modelling of broad band ultraviolet clearness index distributions for Cordoba, Spain. **Agricultural Forest Meteorology**, v. 135, p. 346-351, 2005.

VILLÁN, M. D.; CASTRILLO, A. M.; SANTOS, J. B. Empirical models of UV total radiation and cloud effect study. **International Journal of Climatology**, v. 30, p. 1407-1415, 2010.

WILLMOTT, C.J. On the validation of models. **Physical Geography**, v.2, p.184-194, 1981.

WORLD HEALTH ORGANIZATION. **Public health and the environment's solar ultraviolet radiation**. Global burden of disease from solar ultraviolet radiation, No.13, 2006.

WU, W.; LIU, H.-B. Assessment of monthly solar radiation estimates using support vector machines and air temperatures. **International Journal of Climatology**, v. 32, p.274-285, 2012.

ZANETTI, S. S.; SOUZA, E. F.; CARVALHO, D. F. BERNARDO, S. Estimacão de evapotranspiração de referência no estado do Rio de Janeiro usando redes neurais artificiais. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.12, n.2, p.174-180, 2008.