

PREVISÃO DA ENERGIA SOLAR DIÁRIA BASEADA NO RECONHECIMENTO DE PADRÕES DA SÉRIE TEMPORAL DA RADIAÇÃO HORIZONTAL ATRAVÉS DA TRANSFORMADA DE WAVELET E REDES NEURAS ARTIFICIAIS

Manfred Georg Kratzenberg – manfred@labsolar.ufsc.br

Sergio Colle – colle@emc.ufsc.br

Universidade Federal de Santa Catarina, Departamento de Engenharia Mecânica

Resumo. A utilização de previsões de energia solar com baixas incertezas é particularmente importante para desenvolver sistemas inteligentes de aquecimento solar que são integrados em um sistema de gerenciamento de demanda com objetivo de reduzir em grande proporção a carga máxima do sistema de energia elétrica no setor residencial. Essa potência demandada por chuveiros elétricos pode ser reduzida com o pré-aquecimento inteligente do reservatório térmico que é baseada na previsão da radiação solar. Esse pré-aquecimento elétrico ocorre nas horas da madrugada, quando a tarifa de energia elétrica é baixa no contexto de mercado regulado por consequência da menor demanda. Como esta proposta apresenta múltiplos benefícios econômicos, o uso do aquecimento solar na área residencial pode se tornar economicamente viável em grande escala. Um dos benefícios economicamente destacável é representado pela redução do consumo dos chuveiros elétricos no horário da maior demanda no sistema elétrico. No presente trabalho é desenvolvido um novo método para a previsão do total diário da radiação solar incidente em superfície horizontal com horizonte de previsão de 24h. É pressuposto que o comportamento da cobertura de nuvens de dias anteriores possui informações importantes, as quais podem ser utilizadas para a previsão da radiação solar. A Transformada Discreta de Wavelet (TDW) é utilizada para melhorar o reconhecimento de padrões referentes ao comportamento da radiação. Pela decomposição TDW são obtidas para cada horário um valor que corresponde à energia da radiação solar e diferentes valores que correspondem às variações em diferentes bandas de frequência da radiação solar. A Regressão Linear de Múltiplas variáveis (RLM) é utilizada para selecionar aquelas variáveis de entrada do modelo que possuem a maior significância estatística e o melhoramento do desempenho de previsão é verificado com o uso de uma Rede Neural Artificial (RNA).

Palavras-chave: Previsão da Energia Solar, redes neurais artificiais, transformada de wavelet

1. INTRODUÇÃO

O consumo de energia elétrica no horário entre 18 e 20 horas, onde há uma maior demanda na área residencial, possui um alto custo de US\$ 377/kW (Oliva *et al.*, 1994, citado por Salazar, 2004). A maior corrente nesse horário decorre dos chuveiros elétricos que representam uma potência de 4 GW no sistema elétrico nacional, segundo Moreira (1985). Essa potência cresce aproximadamente na mesma proporção do crescimento da população do Brasil. Como foi demonstrado em um experimento com noventa diferentes sistemas de aquecimento, os sistemas compactos de aquecimento solar, utilizados na área residencial em conjunto com os chuveiros elétricos, possibilitam uma redução da potência do sistema elétrico no horário de pico em 60% (Salazar, 2004). Porém, para dias com baixa radiação solar a potência de pico permanece ainda invariável. Com o objetivo de extinguir essa potência em medida maior e obter uma confiabilidade alta nesse processo, torna-se necessário o uso de sistemas inteligentes de aquecimento solar integrados a um sistema de gerenciamento de demanda. Segundo o conceito desse tipo de sistema, o seu reservatório é aquecido exclusivamente pela energia solar em dias que possuem céu claro. Em dias com céu encoberto faz-se necessário um aquecimento elétrico adicional que é realizado durante a madrugada, quando a tarifa efetiva da energia elétrica é baixa. Para determinar a energia deste pré-aquecimento são necessárias as previsões da energia da radiação solar em plano inclinado e também as previsões da temperatura ambiente.

2. REVISÃO BIBLIOGRÁFICA

O intervalo no tempo futuro para o qual uma variável qualquer é prevista é denominado comumente como horizonte de previsão. A energia da radiação solar incidente em superfície horizontal pode ser prevista através de três distintos métodos para diferentes horizontes de previsão, a saber:

(i) Método baseado em simulações utilizando modelos de Previsão Numérica de Tempo – PNT. Esses modelos estimam o estado futuro de diferentes grandezas físicas da atmosfera, a partir de simulações numéricas da circulação da atmosfera em escala global e em escalas regionais. Através de uma estatística da saída do modelo numérico, também conhecido como estatística de saída do modelo (*MOS – Model Output Statistic*), as previsões de uma variável podem ser

obtidas ou corrigidas sob utilização das demais variáveis previstas. Para a MOS pode ser utilizado o modelo de regressão linear de múltiplas variáveis RLM ou modelos do tipo de RNA (Girodo, 2006; Guarnieri, 2006)¹.

(ii) Método baseado no deslocamento de nuvens a partir de imagens sequenciais obtidas por satélites (Lorenz, 2004). Nesse método aparecem duas incertezas distintas, quais sejam: a incerteza da medição indireta da radiação solar através de imagens da superfície terrestre e a incerteza agregada ao deslocamento estimado de nuvens. Com um horizonte de previsão superior a seis horas, as incertezas aumentam exponencialmente, reduzindo consideravelmente o alcance desse método para horizontes de previsão maiores.

(iii) Método baseado em séries temporais de medições de grandezas meteorológicas, sobretudo a radiação solar incidente. Normalmente são utilizados modelos que interpretam o comportamento de séries temporais, tais como a Auto-Regressão (AR) e as Redes Neurais Artificiais (RNA) (Mellit and Pavan, 2010).

O presente trabalho concentra-se no método (iii) cuja série temporal é baseada em dados de radiação solar incidente em superfície horizontal medida com piranômetro de padrão secundário (World Meteorological Organization, 2004). O modelo baseia-se em uma RNA cujos preditores ou variáveis de entrada são obtidas a partir da transformada discreta de wavelet (TDW). Esta transformada é utilizada para decompor a série de medições da radiação mencionada. O modelo faz a previsão do preditando, que consiste na média diária da radiação medida no dia posterior aos preditores. (Cao e Cão, 2004) utilizaram as médias diárias da radiação solar adotando um método de previsão de médias futuras desconhecidas. Os autores fizeram uso de uma série temporal dessas médias, da transformada TDW e de uma RNA, para estimar o valor futuro de uma previsão, para cada escala de tempo ou banda de frequência. Esse método de predição de séries temporais já foi utilizado por vários outros autores nas mais distintas aplicações. Em lugar da referida RNA, alguns autores utilizaram modelos mais simples como o da AR de múltiplas entradas. Pelo método mencionado, a série temporal é decomposta em sub-séries pela TDW. Os valores futuros dessas sub-séries são depois preditos pelas RNA ou pelos modelos AR. Após à predição das sub-séries, os valores preditos da radiação são obtidos pela reconstrução TDW. Na aplicação desse método ocorre um interessante e inesperado efeito, que não é diretamente visível pelas equações da TDW. A decomposição baseada na TDW acessa, através da convolução com as funções wavelet, não apenas valores passados do parâmetro $H(n-1) \dots H(n-k)$ como deveria, mas também os valores futuros da série temporal a ser decomposta. Por exemplo, para se calcular $H(n-1)$ da sub-série da primeira escala de decomposição, sob emprego da função Daubechies de quarta ordem, a TDW utiliza as médias diárias H do tempo $n, n+1, n+2$. Esse acesso aos valores futuros é visível, quando se programa o cálculo da TDW em uma planilha do programa ExcelTM sob análise desse cálculo com o uso da função 'rastrear células precedentes' neste programa. O efeito ocorre nas duas séries temporais, quais sejam, a série utilizada para o treinamento e a série independente utilizada para a validação. Mellit (2006) utiliza funções wavelet da Transformada Contínua de Wavelet - TCW como funções de ativação em uma RNA com recorrência. A característica de recorrência é atingida quando as informações entre as entradas e a saída ou as saídas de uma RNA não se propagam de forma unidirecional na utilização da RNA treinada. Devido à recorrência, esse tipo de RNA possui uma memória curta e assim é habilitada em modificar os seus parâmetros e, portanto, a sua característica de modelo em função do comportamento da série temporal precedente à previsão e também das previsões precedentes. Previsões precedentes não são disponíveis no início de uma série de amostras constituída de preditores e preditandos. Esse início ocorre, também, após cada intervalo no qual o sistema de medição está em estado de falha e não fornece medições. Por possuir esta característica não é propício utilizar essa RNA sob utilização de séries temporais incompletas, onde há um grande número de intervalos de tempo sem medições. Mellit e Pavan (2010) utilizam a série medida da radiação solar, da temperatura ambiente e o número do dia como preditores para treinar uma RNA de múltiplas camadas sob utilização do algoritmo de Levenberg-Marquard. Fizeram uso das médias horárias da radiação, obtidas de uma estação de medição localizada na cidade de Trieste, na Itália. Os autores obtiveram para um horizonte de previsão de um dia incertezas do tipo *Root Mean Square Error* (RMSE) referentes às médias horárias que variam em função do dia empregado entre 32,98 até 70,2 %. Hocaoglu et. al (2008) observou duas distintas correlações das médias horárias da radiação solar na cidade de Eylul na Turquia. Na primeira, os autores compararam as médias horárias I [W/m^2] do dia n com as do dia $(n-1)$ no mesmo horário e, na segunda, compararam as médias da hora h com as da hora $(h-1)$ do mesmo dia. Compararam as predições de médias horárias obtidas por diferentes RNA com as predições obtidas a partir do uso de filtros lineares. Avaliando separadamente os resultados destas duas configurações obtiveram o menor RMSE de 34,57 W/m^2 para uma RNA com oito neurônios na camada oculta e três preditores, $I(n,h-1)$, $I(n,h-2)$ e $I(n-1,h)$ para prever os valores de $\hat{I}_n(n,h)$.

¹ Utilizando as medições da estação de meteorologia da UFSC do laboratório de energia solar, o autor (Guarnieri, 2006) obteve com o modelo de PNT Eta um RMSE de 6,76 MJ/m^2 para as previsões da energia diária com horizonte de um dia para Florianópolis. Este valor corresponde a uma incerteza relativa de $rRMSE = 43,21\%$ em relação à média dos integrais diários de 15,635 MJ/m^2 . Com a MOS, o autor obteve um $rRMSE = 25,8\%$. Como referência o autor utilizou as medições da radiação em Florianópolis de Janeiro 2002 até Outubro 2005. Convertendo estes dois RMSE para as médias diárias da radiação, o autor obteve os valores de $RMSE = 78,20 W/m^2$ e $46,68 W/m^2$ da média diária da radiação de 180,96 W/m^2 .

3. MÉTODO PROPOSTO

A série das médias minutais da radiação solar medida é uma função da posição do sol, da transmitância das nuvens, das suas velocidades de deslocamento e das suas expansões. A radiação contém, portanto, implicitamente informações sobre o estado ou o regime das nuvens que se deslocam acima da estação de meteorologia com a qual a radiação solar é medida. No presente artigo é pressuposto que padrões de comportamento específicos da radiação antecedem as suas relacionadas intensidades. Pelo uso de uma RNA e da Transformada Discreta de Wavelet (TDW), os padrões mencionados são reconhecidos para serem depois utilizados na estimação da radiação prevista (\hat{H}_n). Obtém-se assim, por exemplo, na entrada de uma frente fria, em um local específico, um padrão de comportamento específico da radiação solar e das nuvens nos k dias que antecedem a cobertura completa. Mais especificamente, pressupõe-se que a média diária da radiação prevista (\hat{H}_n) para o dia n é uma função dos padrões de comportamento da série das médias minutais medidas em horário específico de k dias anteriores, os dias $(n-1)$ a $(n-k)$. Foi desenvolvido um método, no qual a TDW é utilizada para configuração dos preditores, decompondo separadamente o número de (h, k) séries de médias minutais. As séries foram obtidas dos $(n-1) \dots (n-k)$ dias anteriores ao dia da previsão n considerando que o sol aparece ao máximo em quatorze diferentes horários $h = 1 \dots 14$. A decomposição pela TDW é utilizada para avaliar as variações em diferentes bandas de frequência em cada uma dessas séries. A partir das médias e desvios padrão das sub-séries obtidas pela decomposição, proporciona-se preditores que caracterizem os diferentes comportamentos da atmosfera e assim da radiação solar. Os preditores são utilizados depois pela RLM e também por uma RNA de múltiplas camadas, que não possua recorrência, para prever a radiação solar da média diária H_n . Modelos de previsão tais como a Auto-Regressão (AR) ou RNA são modelos que podem interpretar o comportamento de uma variável de saída, denominada como preditando, em função das variáveis de entrada, expressos como preditores. A minimização da função de custo Co , que é comumente representada pela soma dos desvios ao quadrado (equação 1) é denominada como regressão no caso do modelo AR ou RLM e como treinamento no caso do modelo de RNA

$$Co = \sum_{i=1}^N (\hat{H}_i\{P\}_i - H_i)^2 \rightarrow \min \quad (1)$$

Na Eq. 1, H_i , é a média diária da radiação medida que determina os preditandos e $\hat{H}_i\{P\}$ é a radiação estimada pelo modelo, que é uma função do vetor dos preditores $\{P\}_i$. Na regressão ou no treinamento do modelo é identificada a relação que existe entre os preditores $\{P\}_i$ e os preditandos H_i , a partir de uma série de amostras, ($\{P\}_1, H_1$), ($\{P\}_2, H_2$) ... ($\{P\}_N, H_N$), elaboradas a partir de medições que pertencem ao intervalo de treinamento. O preditando H_i representa o valor alvo para o qual a saída do modelo $\hat{H}_i\{P\}_i$ deve se ajustar. Na validação do modelo, que pode ocorrer ao longo do treinamento, o seu desempenho de generalização é verificado com amostras provenientes de preditores e preditandos independentes que foram elaborados a partir de medições de um intervalo de tempo distinto. Se a validação ocorre ao longo do treinamento, são adotados aqueles parâmetros do modelo, com as quais foi obtida a menor incerteza com a amostra do intervalo de validação, otimizando desta forma o desempenho de generalização do modelo. Como o treinamento se baseou, neste caso, também nas amostras do intervalo de validação, deve-se utilizar um terceiro intervalo de amostras de teste, que é completamente independente dos primeiros dois, para avaliar o desempenho de generalização da previsão. No presente trabalho foram utilizadas as médias minutais das medições da radiação global incidente na área horizontal. Estas foram obtidas do piranômetro do tipo CM21 da empresa Kipp&Zonen da estação meteorológica da Universidade Federal de Santa Catarina, em Florianópolis (latitude $-27,6^\circ$, longitude $48,52^\circ$). Esta estação é integrada à rede de estações da *BSRN - Baseline Radiation Network (World Meteorological Organization, 2004)*. O período de medição iniciou-se no dia 226 de 2004 e finalizou-se no dia 365 de 2007. Medições ausentes em intervalo inferior a 15 minutos, foram supridas com base em valores de interpolação. Devido à ausência de algumas das medições, os seguintes dias não puderam ser utilizados para configurar os preditores: 2004 (320...336); 2005 (148...150, 210...228, 287...305, 328); 2006 (326); 2007 (151, 171, 186).

4. TRANSFORMADA DE WAVELET

Devido à ocorrência de nuvens de forma aparentemente aleatória, a série temporal das médias minutais da radiação solar é constituída de uma variedade de frequências harmônicas com composição distinta ao longo do dia. Portanto, essa série representa um sinal não estacionário. Para acessar o tempo t no qual uma frequência harmônica ocorre ao longo de um sinal não estacionário, pode-se utilizar um intervalo de tempo limitado desse sinal, ao qual é aplicada a transformada de Fourier rápida (FFT – Fast Fourier Transform). Se esse intervalo é demasiado curto, a FFT aumenta a sua incerteza na análise de baixas frequências e se o intervalo é demasiado longo, este método aumenta a sua incerteza para definir o local onde ocorre uma dada frequência harmônica (Addison, 2002). A transformada de wavelet possui simultaneamente altas resoluções na análise das frequências e na análise do local no qual essas frequências aparecem ao longo de um sinal não estacionário. Por possuir esta característica excelente, essa transformada é também

denominada como microscópio matemático (Nanavati, 2004). As duas resoluções são obtidas através de uma grande quantidade de funções wavelet normalizadas ψ^{ab} , as quais são utilizadas na convolução. As funções ψ^{ab} são originadas a partir de uma função wavelet-mãe ψ , segundo a equação que segue Onde a é o fator de escala, que define a expansão de ψ e b é o fator de translação que define o deslocamento de cada uma das funções expandidas ao longo do sinal e t é o tempo que representa um local de referência em relação ao deslocamento das funções ψ^{ab} . Na decomposição do sinal com a transformada discreta de wavelet (TDW) em múltiplas escalas ($m = 0, 1, 2, \dots, m_x$), o fator de escala é definido por ($a = 2^m$) e o posicionamento das funções wavelet expandidas é obtido com ($b = a \cdot n$), onde $n = 0, 1, 2, \dots, n_x$ ($m, n \in \mathbb{Z}$). A TDW segue o modelo de um banco de filtros digitais, onde os detalhes ou variações do sinal são obtidos com a função wavelet, que representa o filtro passa alta e o comportamento aproximado do sinal é obtido com a função escala, que representa um filtro passa baixa. A largura da função wavelet é análoga ao comprimento de onda na FFT e esta é definida pelo número de coeficientes do filtro digital da TDW ao qual essa função corresponde. Pela convolução do sinal não estacionário com as funções expandidas e trasladadas é obtido para cada uma das funções wavelet e escala um coeficiente de wavelet e um coeficiente de escala. O coeficiente de cada uma das funções wavelet e escala determina a amplitude dessas funções presente no sinal. No caso das funções da TDW da família Daubechies, o primeiro banco de filtros, também denominada como transformada Haar, ou transformada Daubechies de primeira ordem, possui dois coeficientes. Portanto, a função wavelet envolve dois valores de uma série temporal na primeira escala ($m = 1$) e o seu intervalo de cobertura corresponde assim ao intervalo de amostragem T do sinal. O filtro da segunda ordem da série Daubechies possui quatro coeficientes e seu intervalo é $3T$. Devido à expansão das funções wavelet ($a = 2^m$), o número de coeficientes é reduzido pelo fator dois em cada incremento de escala. Mesmo assim, os coeficientes retêm a informação do sinal e com uso dos coeficientes wavelet e escala o sinal pode ser reconstruído por completo pela TDW. No presente artigo são utilizados os sinais decompostos da TDW não dizimada, para configurar os preditores de um modelo de previsão. Pela decomposição da TDW não dizimada o sinal é decomposto em sub-sinais em lugar dos vetores de coeficientes. Os sub-sinais possuem a mesma quantidade de valores que o sinal não decomposto e, portanto, essa TDW é também denominada como TDW redundante. A quantidade de sub-séries ($m + 1$) é igual à quantidade dos vetores de coeficientes wavelet e resulta da decomposição em m_x escalas. Obtêm-se m sub-sinais ou vetores de coeficientes que representam os detalhes e um sub-sinal ou vetor que representa a aproximação do sinal. Os sub-sinais de detalhes interpretam as variações do sinal, em cada escala de tempo. Somente para a última escala m_x é obtido o sub-sinal que aproxima o comportamento do sinal a ser decomposta. Pela TDW obtém-se a expectativa de maior eficácia na interpretação dos efeitos sistemáticos, devido à habilidade de discernir entre os efeitos que aparecem em distintas bandas de frequência ou escalas de tempo. Devido à ortogonalidade das funções wavelet da família Daubechies, as quais são utilizadas para a convolução no presente artigo, os sub-sinais obtidos da decomposição não possuem redundância de informação (Daubechies, 1992). Desta forma, os sub-sinais obtidos possuem entre si um coeficiente de correlação muito próximo a zero. Esta característica é vantajosa porque variáveis de entrada que possuem informações redundantes em modelos como a RLM ou a RNA normalmente contribuem para o aumento das suas incertezas. Primeiramente, porque se uma nova variável não possui informações adicionais, em decorrência da sua redundância, esta função não pode melhorar o desempenho do modelo. Em segundo lugar, a variável redundante pode piorar o desempenho do modelo, trazendo ruído ou efeitos aleatórios adicionais na entrada do modelo. Com a sub-série dos detalhes da primeira escala ($m = 1$) obtém-se a banda com as maiores frequências do sinal e com a sub-série de detalhes da última escala ($m = m_x$) obtém-se a banda com as menores frequências. Portanto, a sub-série de detalhes na primeira escala visualiza efeitos de variação da radiação provenientes de nuvens de menor extensão ou altas velocidades de seu deslocamento e a sub-série de detalhes na última escala visualiza os efeitos provenientes de nuvens de maior extensão ou baixas velocidades de seu deslocamento. A velocidade e a extensão das nuvens não podem ser reconhecidas separadamente na estrutura do presente modelo. Sob exclusão dos detalhes nas escalas ($1 \dots m_x$) a partir dos filtros passa alta, retém-se a tendência ou a aproximação do sinal na escala m_x . A curva superior da Fig. 1 mostra a série das médias minutais da radiação global de um dia subdividida em quatorze séries, aparecendo em diferentes cores para $1 \dots 14$ intervalos em horários distintos. Cada uma dessas séries é utilizada para decomposição wavelet. A curva inferior da Fig. 1 mostra exemplarmente as quatorze sub-séries obtidas da decomposição wavelet, utilizando os wavelets da família Daubechies da segunda ordem na escala de decomposição $m = 3$. Sob utilização dessas funções wavelet foi obtido o melhor desempenho de previsão. A TDW foi realizada em intervalo horário e foi utilizada para reconhecer melhor o comportamento das séries da radiação solar de $(n-1) \dots (n-k)$ dias anteriores ao dia de previsão n . Pela decomposição em $m_x = 5$ escalas de tempo, foram obtidas para cada intervalo cinco sub-séries de detalhes, cada uma com um distinto espectro de frequências e uma sub-série que aproxima o valor da radiação. As médias da última sub-série foram utilizadas para quantificar o comportamento aproximado da radiação solar, conforme ilustrado na figura 2. Os seus valores são semelhantes aos valores das médias horárias da radiação solar $I(n-1,1) \dots I(n-k,h)$. Como as médias das sub-séries de detalhes são iguais a zero, foram utilizados os seus desvios padrão em intervalo horário para quantificar as variações das $m_x[(n-1,h) \dots (n-k,h)]$ séries que interpretam os detalhes d do comportamento da radiação solar em cada uma das $m_x = 5$ escalas de tempo.

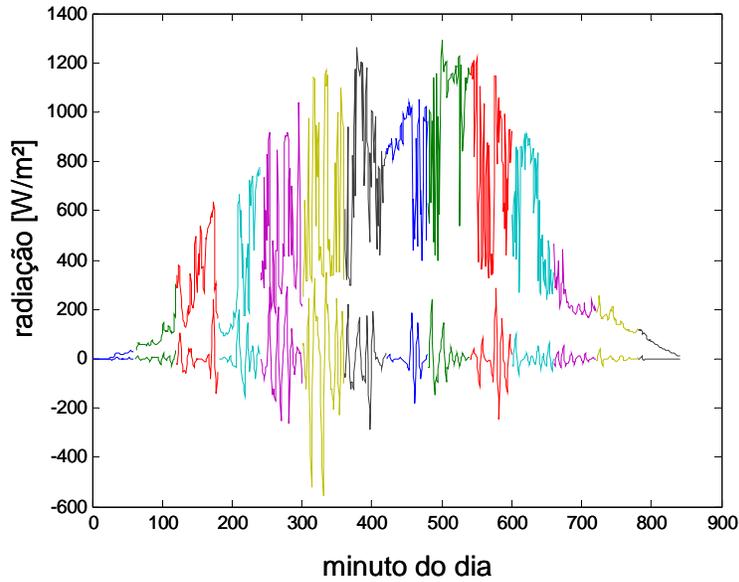


Figura 1 – Curva superior: quatorze séries das médias minutais da radiação global horizontal; Curva inferior: quatorze sub-séries obtidas da decomposição wavelet Daubechies da segunda ordem para a banda de frequência da escala de tempo ($m = 3$).

Utilizando-se $h = 1 \dots 14$ intervalos horários da radiação solar em uma decomposição sob $mx = 5$ escalas para $k = 7$ dias foram assim obtidos $p = 588$ possíveis preditores pela equação que segue

$$p = h (mx + 1) k \quad (3)$$

onde p é número de preditores obtidos.

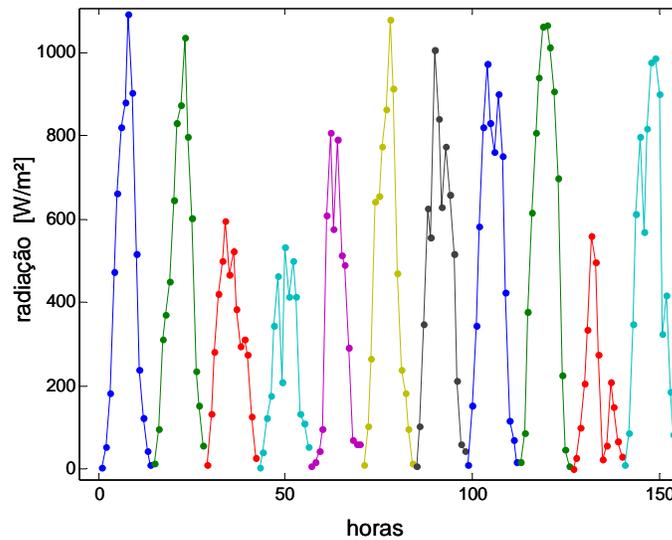


Figura 2 - $h = 1 \dots 14$ médias horárias de onze dias, obtidas das sub-séries de aproximação da escala cinco da decomposição TDW sob utilização da família Daubechies de segunda ordem.

5. REDE NEURAL ARTIFICIAL (RNA) E SEU ALGORITMO DE TREINAMENTO

O modelo de RNA corresponde de forma aproximada ao paradigma das redes neurais do cérebro. São constituídas de diferentes neurônios interligadas e organizadas em camadas. Cada neurônio possui diferentes sinapses ou entradas as quais são ponderadas com pesos e uma função de ativação que pode ser linear ou possuir não-linearidade. A utilização de múltiplos neurônios não lineares em diferentes camadas compôs um modelo cuja não linearidade é obtida a partir da distribuição e combinação das diferentes funções não lineares dos seus neurônios (Haykin, 1999). Outra vantagem nesta

comparação é que o modelo de RNA também pode interpretar interações entre as variáveis de entrada (Donaldson & Kamstra, 1999). Pelo algoritmo de treinamento sob uso de um conjunto de amostras do intervalo de treinamento é reduzida à função de custo (Eq. 1) em cada época do treinamento. Um item importante nos RNA é a escolha das variáveis de entrada e do número de neurônios em cada uma das suas camadas. Sob uso de um alto número de neurônios e variáveis de entrada os RNA podem ficar sujeitos a um sobre-ajuste, perdendo assim o seu desempenho de generalização nos intervalos independentes do treinamento. Por outro lado, se esse número não é suficiente, os modelos RNA podem ter altas incertezas nos três intervalos: no treinamento, na validação e no teste. Com um número adequado de preditores e neurônios nas diferentes camadas das RNA, pode-se obter baixas incertezas no intervalo de treinamento e nos intervalos independentes. Para evitar o sobre-ajuste, podem ser utilizados diferentes métodos (Özesmi et.al, 2006); entre estes o método da regularização e o método denominado como treinamento com parada antecipada (early stopping). Neste segundo método são envolvidos dois conjuntos de amostras no treinamento da RNA, as amostras do intervalo de treinamento e as do intervalo independente de validação. Enquanto os pesos da RNA são ajustados com as amostras do intervalo de treinamento, as amostras do intervalo de validação são utilizadas para verificar o desempenho de generalização do modelo. Quando, durante o treinamento em uma quantidade definida de épocas sucessivas, o valor da função de custo para as amostras do intervalo de validação Co_{val} não diminui, o treinamento é finalizado e são utilizados aqueles valores de pesos da RNA com os quais obtém-se o menor valor de Co_{val} . Ao final do treinamento, o desempenho do modelo é novamente validado com preditores e preditandos de um terceiro intervalo, complementar independentes dos primeiros dois intervalos, denominado como intervalo de teste. No presente artigo, utilizou-se o algoritmo de Regularização Bayesiana (MacKay, 1992) para o treinamento da RNA. Este método restringe a magnitude dos pesos da RNA com objetivo de melhorar o desempenho de generalização. Em comparação com uma RNA que é treinada pelo algoritmo de retro-propagação (*Back Propagation - BP*), também se observou com esse método a vantagem de se obter uma maior reprodutibilidade de seu desempenho, quando repetidos treinamentos são comparados sob utilização da mesma configuração de neurônios da RNA e das mesmas amostras. Outra vantagem é que o algoritmo proporciona uma sugestão para definir a quantidade de parâmetros que devem ser utilizados em uma RNA. Esta é definida pela soma do número dos pesos e das variáveis de entrada da RNA. O maior custo computacional que decorre da Regularização Bayesiana é compensado a partir do algoritmo de Levenberg-Marquardt (LM). O algoritmo de LM é utilizado nas redes neurais artificiais para obter velocidades superiores de treinamento (Plump et. Al, 2005).

6. INTERVALOS DE TREINAMENTO DE VALIDAÇÃO E DE TESTE

O desenvolvimento dos modelos estatísticos deve incluir uma discussão no que diz respeito ao uso da quantidade de amostras, que devem ser utilizados para o treinamento, a validação e o teste do modelo. Cada amostra consiste normalmente de diferentes preditores e pode ter um (RLM, RNA) ou mais preditandos (RNA). Embora muitos autores como, por exemplo, Guarnieri (2006) simplesmente subdividam o total das amostras para usar os primeiros 50% das amostras para o intervalo de treinamento, que é seguido por um intervalo de validação (25%), e um intervalo de teste (25%), outros autores, por exemplo, Von Bremen *et. al* (2006) atualizaram o treinamento da RNA utilizada a cada 15 dias sob uso de um intervalo curto de medições de apenas 120 dias. Von Bremen *et. al* utilizaram uma RNA para predição da potência fornecida por aerogeradores e empregaram como preditores algumas das variáveis previstas por um modelo PNT. O uso de longos intervalos proporciona ao modelo uma amostra mais completa, na qual possa aparecer os mais diferentes efeitos que possam ser identificados pelo modelo e utilizados para a previsão, tornando assim o modelo mais inteligente e eficiente. O uso de intervalos curtos evita o reconhecimento de efeitos que decorrem a longo prazo e pode-se obter assim maiores incertezas em decorrência de mudanças climáticas ou efeitos como *El niño* ou *La niña*. Com o objetivo de identificar o comprimento dos três intervalos com as quais se possa obter o maior desempenho de generalização do modelo, os seus comprimentos foram variados. Obteve-se a menor incerteza com 500 dias ou amostras para o intervalo de treinamento e 10 dias para os intervalos de validação e o de teste. Nesta configuração foram selecionados 30 diferentes conjuntos de preditores e preditandos a partir das medições da radiação. Obteve-se assim 30 atualizações do modelo em intervalo de 10 dias.

7. AVALIAÇÃO E COMPARAÇÃO DO DESEMPENHO DE MODELOS

A comparação dos desempenhos de previsão obtidos por modelos desenvolvidos por distintos autores é importante na seleção do melhor modelo para a previsão operacional da radiação solar, sendo esta por um sistema de gerenciamento de demanda da energia elétrica. Uma dificuldade neste particular é a comparação dos diversos modelos, sendo cada um validado para um local com condições climáticas específicas. Considera-se que para locais que possuem altos índices de cobertura e clima sub-tropical o RMSE não é um bom avaliador para ser utilizado na comparação do desempenho. No comportamento da cobertura ocorrem efeitos aleatórios e não previsíveis e também efeitos que se repetem e possam ser reconhecidos e corrigidos por modelos estatísticos. A radiação de céu claro pode ser estimada com um modelo à base de equações determinísticas e, portanto, em um local onde por pura idealização não aparece cobertura de nuvens, poder-se-ia obter a previsão da radiação com uma incerteza próxima a zero, utilizando este modelo. Este local possui a maior correlação entre H_n e H_{n-1} e proporcionaria assim também a menor incerteza do modelo de persistência (equação 4). Partindo deste pensamento, podem ser atingidas as menores incertezas de previsão em locais que possuem os menores índices de cobertura. O modelo de persistência é o modelo mais simples que pode

ser utilizado para a previsão da radiação solar. Nesse modelo, utiliza-se a média diária do dia anterior à previsão (H_{n-1}) para duas finalidades, primeiro como preditor e segundo como preditando. Pressupondo que $H_n = H_{n-1}$, calcula-se o RMSE deste modelo com a equação que segue

$$RMSE_{perst} = \sqrt{1/N \sum_{i=1}^N (\hat{H}_{n,i} - H_{n,i})^2} = \sqrt{1/am \sum_{i=1}^N (H_{n-1,i} - H_{n,i})^2} \quad (4)$$

onde $\hat{H}_{n,i}$ é a média diária estimada do dia n em uma série de $i = 1 \dots N$ previsões obtidas pelo modelo da persistência.

Para construir uma amostra, constituída de N conjuntos de preditores e preditados, necessitam-se de N+1 valores de médias diárias obtidos das medições da radiação solar, para este modelo.

Considerando-se o uso de um modelo perfeito, que interpreta e prevê todos os efeitos sistemáticos que são reconhecíveis a partir de um conjunto de preditores, obtém-se uma incerteza que é aqui considerada como a incerteza minimamente atingível $RMSE_{min}$, que decorre apenas de incertezas aleatórias. Devido a distintos índices de cobertura em diferentes locais, presume-se que cada local ou região climática possui um $RMSE_{min}$ individual. Como o $RMSE_{min}$ é distinto em locais de diferentes condições climáticas, o fator de aperfeiçoamento *skill* (equação 5) é o melhor indicador para avaliar e comparar o desempenho dos modelos desenvolvidos. A avaliação pelo fator *skill* pôs os resultados obtidos com um modelo desenvolvido em contexto com um segundo modelo mais simples que possui um desempenho inferior, e é denominado como modelo de referência. No presente artigo, o modelo da persistência é utilizado como referência onde $RMSE_{ref} = RMSE_{perst}$. Comparando-se um índice I qualquer de um modelo desenvolvido com o índice de um modelo de referência I_{ref} , obtém-se o fator de aperfeiçoamento, que é também denominado como fator *skill*, como segue

$$skill(I, I_{ref}) = (I - I_{ref}) / (I_{perf} - I_{ref}) \quad (5)$$

onde I pode ser representado pelo RMSE, pelo módulo do RMBE ou pelo fator de correlação R. I_{perf} é o valor perfeito desse índice que é a unidade para o R e é zero para o RMSE e o RMBE. O fator *skill* define apenas o melhoramento obtido por um modelo em relação a um modelo de referência e obtém-se assim, para locais com distintos RMSE e $RMSE_{ref}$, o mesmo fator *skill* = 0,5, por exemplo, quando o RMSE é reduzido para $\frac{1}{2} RMSE_{ref}$. Em regiões climáticas que possuem condições subtropicais, ocorrem no verão regimes de nuvens semelhantes aos regimes de regiões tropicais e no inverno ocorrem regimes semelhantes aos que ocorrem em regiões não tropicais. Comparando-se distintos locais que possuem um $RMSE_{perst}$ igual, supôs-se que o $RMSE_{min}$ de um local com regimes subtropicais é maior que o de locais com regimes puramente tropicais ou não tropicais. A razão disto é que em locais subtropicais o modelo deve levar em conta a transição de regimes tropicais para não tropicais. Como a latitude possui forte influência no clima de um local, pressupõe-se que o $RMSE_{min}$ e o $skill_{max}$, que é considerado como o fator de melhoramento maximamente atingível, são principalmente funções da latitude. Desta forma, o RMSE e os fatores *skill*, obtidos de distintos modelos e validados em distintos locais, poderiam apenas ser comparados se os referidos locais possuísem latitudes semelhantes. No entanto, obtém-se a melhor comparação entre os modelos se estes são aplicados para um ou mais locais de referência sob utilização das mesmas amostras.

8. DISCRIMINAÇÃO ESTATÍSTICA DE VARIÁVEIS UTILIZADA PARA O MELHORAMENTO DO DESEMPENHO DO MODELO

A discriminação estatística considera que algumas das variáveis ou preditores de um modelo não possuem significância estatística e, portanto, a inclusão dessas variáveis aumenta a incerteza do modelo. Pela discriminação estatística estes preditores são excluídos do modelo, obtendo assim um melhor desempenho do modelo.

8.1 Discriminação pelo método de regressão passo a passo

Um dos métodos mais simples, que é utilizado para a discriminação de variáveis, é a RLM do tipo passo a passo que faz parte das ferramentas do programa MatlabTM. Neste método é realizado, a partir de uma primeira regressão, o teste estatístico F com o qual é verificado a significância estatística que cada um dos preditores possui na predição de uma série de preditados. No teste F são assumidas duas hipóteses. Na primeira, denominada como hipótese zero, presume-se que o coeficiente de regressão relacionado a um preditor é igual a zero, sendo assim o valor $p > p_{limite}$. Se a hipótese zero é rejeitada, presume-se que o coeficiente de regressão não seja zero e o preditor *pr* possa ser incluído na regressão sob condição $p_{pr} < p_{limite}$. Partindo da utilização de apenas um preditor, sendo este aquele que possui a maior significância estatística por possuir o menor valor de p, são incluídos os demais preditores em um processo iterativo na regressão passo a passo. Em cada iteração é realizada uma nova regressão sob inclusão daquele preditor, do conjunto de preditores da primeira regressão, que possui a maior significância estatística. No presente método são utilizados preditores elaborados de uma série temporal e medições da radiação solar. Foram utilizados intervalos horários, os quais foram sujeitos ao mesmo tratamento matemático. Como existe correlação entre médias da hora n e n-1, pode haver informações redundantes neste conjunto de preditores. Principalmente sob inclusão de preditores que possuem informações redundantes é possível que a significância estatística de preditores diminua. Portanto, é verificada na regressão passo a passo a significância estatística de todos os preditores após a inclusão de

cada novo preditor. Se um ou mais dos preditores perde a sua significância estatística ($p > p_{limite}$), é excluído em um próximo passo aquele preditor que possui o maior valor de p . Para diminuir a quantidade de iterações oriundas das adesões e rejeições de preditores foram utilizados $p_{limite1} = 0,05$ para a adesão e $p_{limite2} = 0,1$ para a rejeição e adesão de preditores. Estes são os valores adotados na pré-configuração (*default*) da ferramenta de regressão passo a passo do programa MatlabTM. Observou-se nesta configuração que o modelo foi sobre-ajustado, obtendo baixos RMSE para as amostras do intervalo de treinamento e altos RMSE para as do intervalo de validação e de teste, interpretando assim os ruídos aleatórios como se fossem efeitos sistemáticos (figura 3, pontos para $pr = 17$).

8.2 Discriminação a partir da validação do desempenho de generalização do modelo

Presume-se que um modelo qualquer possui o maior desempenho de generalização, quando a relação $RMSE_{tr}/RMSE_{val}$ é mais próxima à unidade, sendo o primeiro obtido com as amostras do intervalo de treinamento e o segundo obtido com as amostras do intervalo de validação. No entanto, obteve-se valores muito pequenos para essa relação com o modelo de regressão passo a passo. Pressupõe-se que se possa obter uma relação $RMSE_{tr}/RMSE_{val}$ próxima à unidade sob exclusão daqueles preditores que possuem a menor significância estatística. A figura 3 foi elaborada sob essa hipótese e partiu-se da inclusão de um número maior de preditores obtidos da regressão passo a passo. Para obter os pontos no lado esquerdo da figura 3 foram retirados iterativamente aqueles preditores que possuíam os maiores valores de p . Após cada iteração foi realizada uma nova regressão sob utilização dos preditores restantes e foram avaliados novamente os valores p . Sob utilização dos coeficientes obtidos pela regressão foi verificado o RMSE e RMBE, tanto para o intervalo de treinamento como para o intervalo de validação. Pode-se observar na Fig. 3 que os RMSE obtidos para os intervalos independentes de validação e de teste são menores sob utilização de um menor número de preditores e conclui-se que o modelo possui um maior desempenho de generalização sob essa condição. Em aproximadamente $66 W/m^2$ ocorre aproximadamente a intersecção entre os $RMSE_{tr}$ e $RMSE_{val}$. Sob a utilização de $pr = 17$ preditores ocorre o maior desempenho de generalização do modelo elaborado que é aqui definido com o valor mínimo de $\max(RMSE_{tr}, RMSE_{val})$. Considera-se, portanto, que o maior desempenho de generalização do modelo possa ser obtido com o modelo de regressão que possui 17 preditores. Como esperado obteve-se para todas as regressões um *Root Mean Bias Error* para as amostras do intervalo de treinamento ($RMBE_{tr}$) igual a zero. Observou-se em algumas das discriminações a tendência do modelo em estimar valores para uma faixa muito limitada ($\hat{H}_i = \pm 40 W/m^2$) sob a seleção de um baixo número de preditores ($pr \sim 1...5$). Por esta razão foi necessário parar esse processo de discriminação com o critério da equação (6)

$$\max(\min(\hat{H}_i) - \min(H_i)) \wedge \max(H_i) - \max(\hat{H}_i) \leq 80 W / m^2 \quad (6)$$

onde $i = 1...N$ representa o conjunto de 500 amostras utilizadas em cada uma das regressões

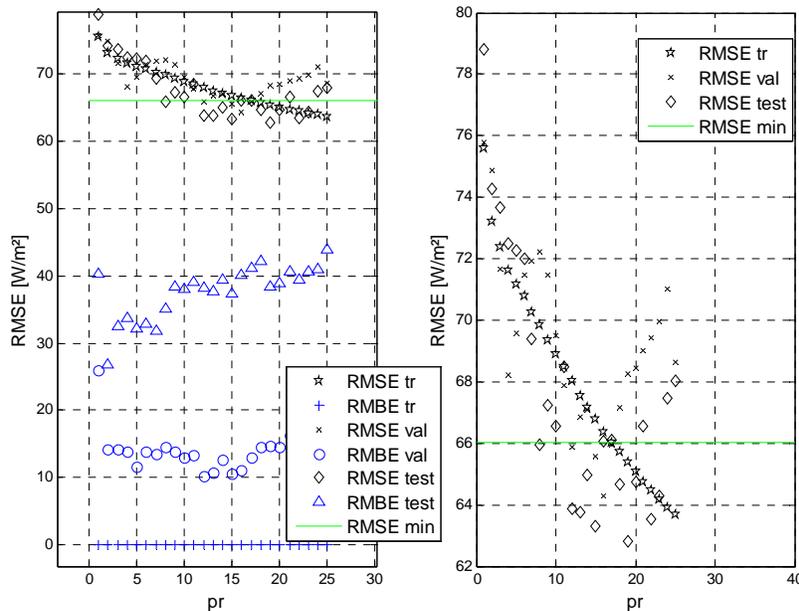


Figura 3-a: RMSE e RMBE obtidos sob avaliação do desempenho para os intervalos de treinamento, de validação e de teste sob utilização de distintos números de preditores, Figura 3-b: ampliação de 3-a

9. RESULTADOS OBTIDOS COM OS MODELOS DE REDES NEURAIS ARTIFICIAIS E DA REGRESSÃO LINEAR DE MÚLTIPLAS ENTRADAS

Como o modelo de RLM verifica apenas a significância estatística de preditores, cuja função é linear entre a entrada e a saída do modelo, presume-se que o modelo RNA possui um maior desempenho de generalização que o de RLM. Neste sentido, alguns dos efeitos que apenas puderam ser identificados como efeitos aleatórios, e não corrigíveis com o modelo RLM, são identificados como efeitos sistemáticos pela RNA sob melhoramento do desempenho do modelo. Assumindo que a maior proporção de efeitos entre entrada e saída ocorre de forma linear, compara-se os desempenhos dos modelos RNA e RLM (figura 6 e tabela 1), sob utilização dos preditores obtidos a partir da discriminação dos preditores com a RLM (intersecção da figura 4). Se, sob a utilização dos mesmos preditores, os RMSE para os intervalos de teste do modelo RNA são inferiores aos RMSE obtidos com o modelo RLM, conclui-se que ocorrem efeitos que não podem ser interpretados pelo modelo simples de RLM, sendo desta forma o uso da RNA justificado. Nos trinta treinamentos da RNA e regressões da RLM na Fig. 6, selecionou-se pela discriminação das variáveis (item 8) vetores que possuem entre 4 e 25 preditores. As figuras 4 e 5 mostram os valores medidos H_m e previstos H_e com a RNA treinada para as amostras dos três intervalos, o de treinamento, o de validação e o de teste.

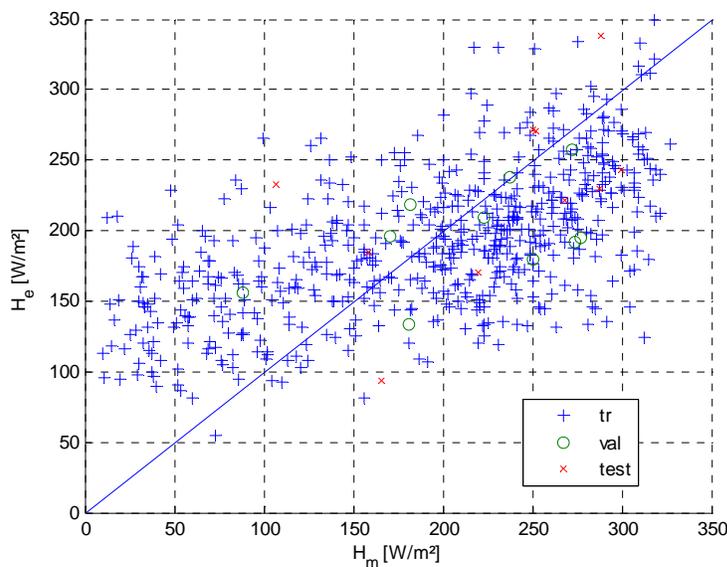


Figura 4 - Média diária medida H_m e média diária estimada pelo modelo da rede neural artificial H_e para os intervalos de medição, de validação e de teste

Os resultados obtidos com os diferentes modelos, na figura 6 e na tabela 1, foram obtidos a partir da atualização pelo deslocamento desses três intervalos ao longo do tempo em passos de 10 dias. Selecionou-se assim, por exemplo, as médias dos dias 8...507 do ano para definir os primeiros 500 preditandos. Para esta mesma configuração são utilizadas as médias dos dias 508...517 como preditandos do intervalo de validação e aquelas dos dias 518...527 como preditandos do intervalo de teste. Cada preditando do dia n antecedem as amostras obtidas com a TDW de $k = 7$ dias anteriores $(n-1)...(n-k)$ que configuram os preditores. Na atualização do modelo são respectivamente utilizados os dias 18...517, 528...537 e 538...547 como preditandos do segundo treinamento. Em cada uma das atualizações, apresentadas na figura 6, obteve-se resultados semelhantes aos apresentados nas figuras 3 a 5. A tabela 1 apresenta as médias obtidas com as diferentes atualizações dos diferentes modelos, também expostas na figura 6. Sob utilização dos mesmos preditores obtidos pela discriminação são comparados na tabela 1 o modelo RLM com diferentes configurações de neurônios da RNA de regularização Bayesiana as quais foram treinadas com o algoritmo de LM. Nos neurônios de entrada e de saída da RNA foram utilizadas funções lineares de ativação e nas camadas escondidas foram utilizadas funções de ativação do tipo tangente hiperbólica sigmoideal (*tansig*). As entradas e saídas da RNA foram normalizadas com a equação que segue:

$$X_{\text{norm},i} = Y_{\text{min}} + (X_i - X_{\text{min}}) (Y_{\text{max}} - Y_{\text{min}}) / (X_{\text{max}} - X_{\text{min}}) \quad (7)$$

X_i é o valor de um dos preditores da amostra i . X_{min} e X_{max} são os valores mínimo e máximo obtidos do conjunto de amostras de preditores da TDW. Os valores de $X_{\text{norm},i}$ encontram-se no intervalo de $[Y_{\text{min}}, Y_{\text{max}}]$, onde $Y_{\text{max}} = 1$ e $Y_{\text{min}} = -1$. Na normalização dos preditandos estimados, as variáveis $X_{\text{norm},i}$, X_i , X_{min} e X_{max} da equação (7) são substituídas,

respectivamente, pelas variáveis $\hat{H}_{norm,i}$, \hat{H}_i , \hat{H}_{min} e \hat{H}_{max} para se calcular a média estimada, resolvendo a equação obtida na variável \hat{H}_i . Também é analisado o RMSE do modelo simples de persistência na tabela 1.

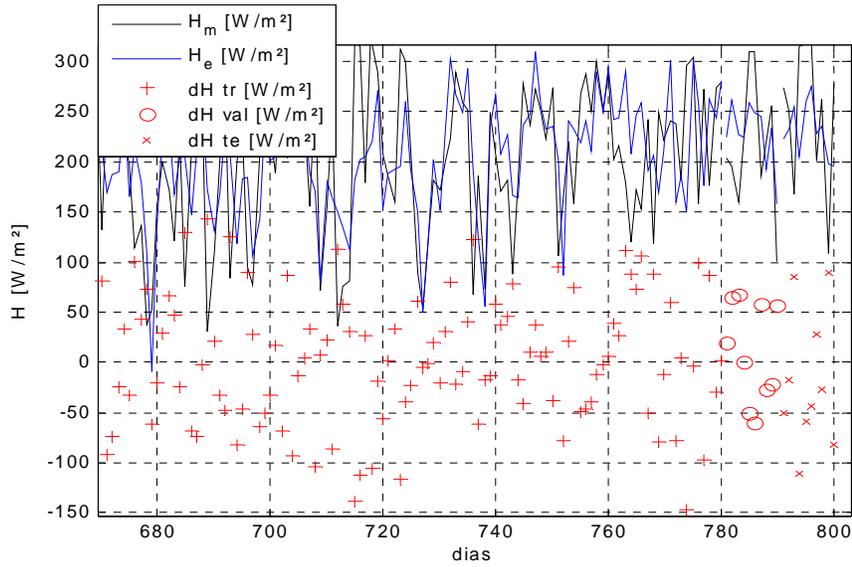


Figura 5 - Média diária medida H_m , média diária estimada pelo modelo da rede neural artificial H_e e seus desvios $dH = (H_m - H_e)$ para um intervalo reduzido do treinamento e para os intervalos de validação e de teste

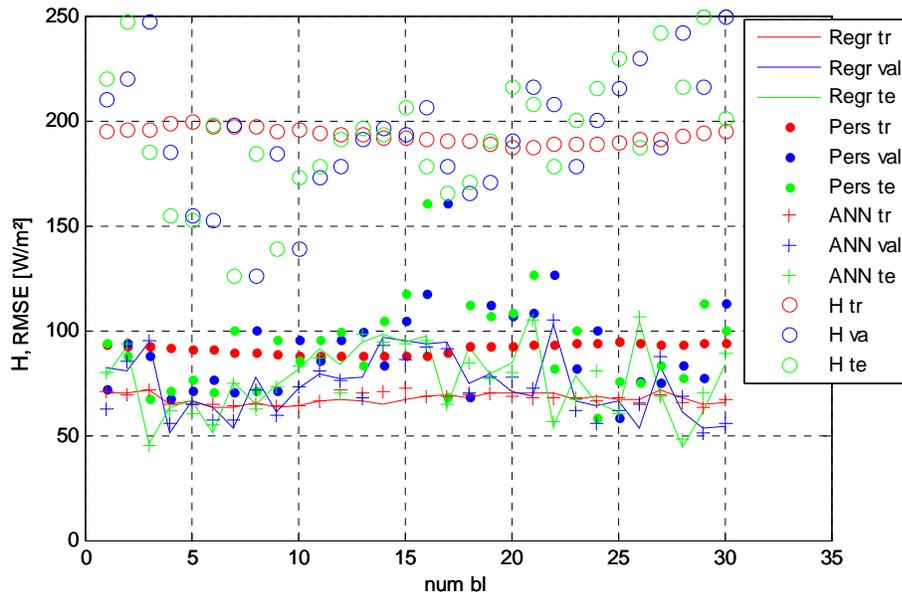


Figura 6 – Média diária medida da radiação (H) e RMSE obtidos dos modelos RLM (Regr) e RNA com um neurônio na camada escondida (ANN's) sob utilização de preditores da TDW em comparação com o modelo da persistência (Perst)

Tabela1: RMSE médios para trinta conjuntos de amostras obtidas das atualizações do modelo

dias	intervalo	DWT			regress.	persist.
		neurônios escondidos				
		5	1	10-3		
500	train	67.23	67.76	71.3	67.33	91.65
10	val	73.88	73.51	74.49	73.85	91.67
10	test	76.56	76.19	76.43	76.79	92.60

10. CONCLUSÕES

No presente artigo é desenvolvido um novo método de previsão da radiação solar global com base nas medições dessa variável de dias que antecedem o dia da previsão e a transformada discreta de wavelet. Com uma RNA de apenas um neurônio na camada escondida foi obtido o menor RMSE médio de 76,19 W/m² para os intervalos independentes de teste. Este modelo melhorou o modelo simples de persistência em $skill = 17,72\%$ e possui um rRMSE de 39,41%. O uso da RNA não proporcionou uma redução significativa em comparação com os resultados obtidos pela RLM, sendo utilizados nos dois modelos os mesmos preditores obtidos da DWT. O desempenho da previsão do modelo apresentado é superior ao desempenho obtido como o modelo ETA, mas é inferior quando a previsão com o modelo ETA é melhorada pela estatística de saída de modelo (MOS). A combinação do modelo de RLM passo a passo com a decomposição da TDW viabiliza o reconhecimento de padrões de comportamento da série não estacionária da radiação solar. Este método possui a tendência de identificar um grande número de preditores que possa proporcionar, em consequência, um modelo com excessiva especificidade. Neste caso, o método tende a identificar os dois, os efeitos sistemáticos e os aleatórios da série da radiação solar como efeitos sistemáticos. Como há uma parcela grande de efeitos aleatórios o modelo perde o seu desempenho de generalização, que se torna visível quando este é aplicado para amostras de preditores e preditandos obtidos de medições independentes. Pela sucessiva exclusão daqueles preditores que possuíram a menor significância estatística, pode-se obter um melhoramento no desempenho de generalização do modelo. Em trabalhos futuros seria interessante associar preditores obtidos de fontes independentes, por exemplo, utilizar a combinação dos preditores obtidos pelo presente método com os obtidos a partir de um modelo PNT-MOS para realizar previsões para locais próximos à estação de medição da radiação. Como esses dois métodos de previsão possuem preditores distintos, e provavelmente não correlacionados, é provável de proporcionar assim uma incerteza de previsão inferior.

REFERÊNCIAS

- Addison P.S, 2002. The illustrated wavelet transform handbook - Introductory theory and applications in science, engineering, medicine, and finance, Institute of physics publishing Bristol and Philadelphia, 362 p.
- Cao S., Cao J., 2004. Forecast of solar irradiance using recurrent neural networks combined with wavelet analysis, Journal of Applied Thermal Engineering pp. 161-172,
- Donaldson K.G., Kamstra M., 1999. Neural network forecast combining with interaction effects, Journal of the Franklin Institute - Pergamon, 10p.
- Daubechies, I., 1992., Ten lectures on wavelets. In: CBMS – NSF Regional Conferences Series in Applied Mathematics SIAM, Philadelphia, 1992, 343 p.
- Girodo M., 2006. Solarstrahlungsvorhersage auf der Basis numerischer Wettermodelle, Ph. D Thesis, University of Oldenburg Germany, Faculty of Mathematics and Natural Sciences, Energy and Semiconductor Research Laboratory, 2006, 159 p.
- Guarnieri, R. A., 2006. Emprego de Redes Neurais Artificiais e Regressão Linear Múltipla no Refinamento das Previsões de Radiação Solar do Modelo Eta, Dissertação de Mestrado, CPTEC, INPE, São José dos Campos (SP).
- Haykin S., 1999. Neural networks – A comprehensive foundation, Prentice Hall, Second Edition, Mc. Master University, Hamilton, Ontario – Canada, , 823 p.
- Hocaoglu F.O, Gerek. Ö. N., Kurban M. 2008. Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural networks, Solar Energy 82, pp.714–726
- Kalogirou S. A. 2001. Artificial neural networks in renewable energy systems applications: a review, Renewable and Sustainable Energy Reviews 5 , pp. 373–401

Lorenz, E. 2004. Methoden zur Beschreibung der Wolkenentwicklung in Satellitenbildern und ihre Anwendung zur Solarstrahlungsvorhersage, Ph.D. thesis, Carl von Ossietzky University, Faculty of Mathematics and Natural Sciences: Oldenburg, Germany, 111 p.

Foresee, F.D., Hagan, M.T., 1997 Gauss-Newton approximation to Bayesian regularization, Proceedings of International Joint Conference on Neural Networks, Nagoya, Japan

Mellit A., Kalogirou S.A. 2008. Artificial intelligence techniques for photovoltaic applications: A review, J., Progress in Energy and Combustion Science, 59 p.

Mellit A., Pavan A.M. 2010. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy, Solar Energy 84, pp. 807–821

Moreira J.G.S., 1985. Energia solar no Brasil – O uso de coletores planos para o aquecimento de água, Rio de Janeiro, UFRJ, Dissertação de mestrado

Nanavati S.P. e Panigrahi P.K., 2004. Wavelet transform: A new mathematical microscope Editora Springer India, Journal Resonance, Volume 9, Nr. 3 /,15p. <http://www.springerlink.com/content/3937325u2mt1kk50/>

Oliva G.A., Borges T.P.F., Martins G., Desenvolvimento de tecnologia de pré-aquecedor solar de água para chuveiros elétrica reduzida. Anais do XII Seminário Nacional da Distribuição de Energia Elétrica. Recife, 16 a 21 de Outubro de 1994.

Kratzenberg M.G. 2008. Correção estatística a base da transformada wavelet para previsão de energia solar através de modelo numérico meteorológico, Exame de qualificação, Universidade federal de Santa Catarina

MacKay DJC. A practical Bayesian framework for backpropagation networks. Neural Comput. 1992; 4: 448–472.

Özesmi S.L., Tan C.O., Özesmi U., 2006. Methodological issues in building, training, and testing artificial neural networks in ecological applications, Ecological Modelling 195, pp. 83–93

Plump A.P., Rowe R.C., York P., Brown M., 2005. Optimisation of the predictive ability of artificial neural network (ANN) models: A comparison of three ANN programs and four classes of training algorithm, European Journal of Pharmaceutical Sciences 25, pp. 395–405

Salazar J. P.L.C., 2004. Economia de energia e redução do pico da curva de demanda para consumidores de baixa renda por agregação de energia solar térmica, Dissertação, Universidade Federal de Santa Catarina, Departamento de Engenharia Mecânica,

Von Bremen, L., Lisboa, L. R., Haas, R.; Araujo, F. F. V.; Maciel, C.; Abreu, S. L.; Colle, S., 2006. Local short-term forecasting for wind power plants in Brazil. In: 8th German Wind Energy Conference - DEWEK, Bremen.

World Meteorological Organization (2004), World Climate Research Programme - Baseline Surface Radiation Network - Operations Manual, Version 2.1, 188p.

PREDICTION OF THE DAILY SOLAR ENERGY, BASED ON THE PATTERN RECOGNITION OF THE RADIATION TIME SERIES, BY THE UTILIZATION OF THE WAVELET TRANSFORMATION AND ARTIFICIAL NEURAL NETWORKS

***Abstract** The utilization of the solar energy prediction with low uncertainties is particularly important for the development of intelligent solar heating systems, which are integrated in a demand management system with the objective to suppress to a great extent the peak power of the electric power system. This power demanded by electric showers heads can be reduced with an intelligent preheating strategy of the heating storage, which is accomplished within the early morning hours. As this technologic proposal presents multiples functionalities and economic benefits, the utilization of solar heating systems in residences may become economically feasible in large-scale. The noticeable economic benefit consists in the power reduction of the electric power system within the hours when the main power consumption appears due to the electric shower heads. In the present work is developed a new solar radiation prediction method. It predicts daily solar energy irradiated on horizontal surface, with a prediction horizon of 24 h. The method considers that the behavior of the sky cloud cover of previous days deliver important information for the daily*

solar radiation prediction. The wavelet transformation is applied to improve the pattern recognition of the time series behavior, separating information of the solar energy and its frequency contents within different frequency bands and day times. With the utilization of the stepwise Multiple Linear Regression (MLR), those model input variables are selected which own statistical significance for the predictions. Furthermore the improvement of the prediction performance is verified by the utilization of an Artificial Neural Network (ANN).

Keywords: *Solar Energy Prediction, Artificial Neural Network, Discrete Wavelet Transformation*