

PREVISÃO DE IRRADIAÇÃO SOLAR PARA UM DIA A POSTERIORI USANDO A TRANSFORMADA WAVELET HAAR PARA OBTENÇÃO DE ATRIBUTOS

Felipe Pinto Marinho (UFC) - fpmarinho@alu.ufc.br

Juliana Silva Brasil (UFC) - julianasbra@gmail.com

Paulo Alexandre Costa Rocha (UFC) - paulo.rocha@ufc.br

Maria Eugenia Vieira da Silva (UFC) - eugenia@ufc.br

Juarez Pompeu de Amorim Neto (UFC) - juarezneto33@hotmail.com

Resumo:

Neste trabalho, realizou-se a previsão de irradiação solar global para um horizonte de um dia a posteriori por meio da utilização de dois modelos de aprendizagem de máquina, a saber: Máquina de Aprendizagem Mínima e Regressão por Vetor Suporte. Estes algoritmos foram aplicados em dois conjuntos de dados, um com a presença de atributos formados por variáveis exógenas (insolação, temperatura do ar, precipitação, etc), variáveis endógenas (dados históricos de irradiação solar) e variáveis temporais (ano, mês e dia da medição). O outro conjunto de dados consiste de preditores caracterizados por serem coeficientes obtidos por meio da aplicação da transformada wavelet Haar à série histórica de irradiação solar, considerando os valores de irradiação solar de 15 dias anteriores em conjunto com o valor de irradiação do dia atual. Para o primeiro caso formou-se um conjunto com 44 atributos e 3254 observações, no segundo se teve um conjunto com 16 atributos e 3254 observações. A média da raiz do erro quadrático médio obtida pela aplicação dos modelos no primeiro conjunto de dados foi de 41,04 W/m² com uma habilidade de previsão média de 9,62 %, já para o segundo caso a média da raiz do erro quadrático médio foi de 44,15 W/m² com uma habilidade de previsão média de 2,78%. Desta forma, percebe-se que utilizando apenas os coeficientes wavelets como preditores para os modelos considerados, chegou-se a uma acurácia de previsão comparável ao caso em que se considera o conjunto com variáveis exógenas, endógenas e temporais. Ressaltando, que o uso desta decomposição traz os benefícios de se ter uma redução no número de preditores, bem como o uso de apenas variáveis endógenas, o que reduz o custo relacionado à aquisição dos demais atributos.

Palavras-chave: *Energia Solar, Aprendizagem de Máquina , Transformada Wavelet Haar*

Área temática: *Radiação Solar*

Subárea temática: *Recursos Solares e Meteorologia da Radiação Solar*

PREVISÃO DE IRRADIAÇÃO SOLAR PARA UM DIA A POSTERIORI USANDO A TRANSFORMADA WAVELET HAAR PARA OBTENÇÃO DE ATRIBUTOS

Felipe Pinto Marinho – fpmarinho@alu.ufc.br
Juliana Silva Brasil – julianasbra@gmail.com
Paulo Alexandre Costa Rocha – paulo.rocha@ufc.br
Maria Eugênia Vieira da Silva – eugenia@ufc.br
Juarez Pompeu de Amorim Neto – juarezneto33@hotmail.com
Universidade Federal do Ceará, Departamento de Engenharia Mecânica

1. Radiação Solar

Resumo. Neste trabalho, realizou-se a previsão de irradiação solar global para um horizonte de um dia a posteriori por meio da utilização de dois modelos de aprendizagem de máquina, a saber: Máquina de Aprendizagem Mínima e Regressão por Vetor Suporte. Estes algoritmos foram aplicados em dois conjuntos de dados. Um deles é constituído de atributos formados por variáveis exógenas (insolação, temperatura do ar, precipitação, etc), variáveis endógenas (dados históricos de irradiação solar) e variáveis temporais (ano, mês e dia da medição). O outro conjunto de dados consiste de preditores que são coeficientes wavelets obtidos por meio da aplicação da transformada wavelet Haar à uma série histórica de irradiação solar formada pelos valores de irradiação solar de 15 dias anteriores em conjunto com o valor de irradiação do dia atual. Para o primeiro caso formou-se um conjunto com 44 atributos e 3254 observações, no segundo se teve um conjunto com 16 atributos e 3254 observações. A média da raiz do erro quadrático médio obtida pela aplicação dos modelos no primeiro conjunto de dados foi de 41,04 W/m² com uma habilidade de previsão média de 9,62 %, já para o segundo caso a média da raiz do erro quadrático médio foi de 44,15 W/m² com uma habilidade de previsão média de 2,78%. Desta forma, percebe-se que utilizando apenas os coeficientes wavelets como preditores para os modelos considerados, chegou-se a uma acurácia de previsão comparável ao caso em que se considera o conjunto com variáveis exógenas, endógenas e temporais. Ressaltando que o uso desta decomposição traz os benefícios de se ter uma redução no número de preditores, bem como o uso de apenas variáveis endógenas o que reduz o custo relacionado à aquisição dos demais atributos.

Palavras-chave: Energia Solar, Aprendizagem de Máquina, Transformada Wavelet Haar

1. INTRODUÇÃO

Dada a necessidade de se progredir a uma matriz energética proveniente em sua grande maioria de fontes de energias renováveis e tendo em vista o grande potencial que a região nordeste apresenta para o aproveitamento de tais fontes, faz-se necessário a realização de pesquisas que proporcionem avanços no sentido de acelerar o processo de transformação da matriz. Neste contexto, a energia solar merece destaque tanto no que diz respeito ao aproveitamento solar fotovoltaico como também na captação térmica concentrada.

Uma das principais dificuldades para o uso da fonte solar na geração de eletricidade para a rede se deve ao seu comportamento intermitente. A natureza estocástica do recurso solar em conjunto com as variações sazonais na geração e nos perfis de carregamento, o alto custo no armazenamento de energia e os requerimentos das indústrias que devem buscar o equilíbrio entre flexibilidade e confiabilidade da rede conduzem a necessidade de se obter previsões confiáveis dos valores de irradiação solar (Inman *et al.*, 2013; Denholm e Margolis, 2007).

Neste sentido, percebe-se a importância da obtenção de previsões de irradiação solar em diversos horizontes temporais. Previsões de curto prazo, que são aquelas realizadas para horizontes de até 6 horas a posteriori, são adequadas para prever quedas de fornecimento, enquanto as de longo prazo, que são realizadas em intervalos de dias, são adequadas na programação de manutenções preventivas.

Dependendo do horizonte de previsão, diferentes tipos de modelos são apropriados (Pedro *et al.*, 2018). Metodologias de previsão que utilizam modelos *Numerical Weather Prediction* (NWP) (Mejia *et al.*, 2018; Nonnemacher *et al.*, 2016) apresentam uma boa acurácia para previsões de longo prazo, já métodos baseados em análise de séries temporais fornecem um bom desempenho para previsões de curto prazo (Trapero *et al.*, 2015; Dong *et al.*, 2013).

Recentemente, têm ganhado muito destaque as abordagens que utilizam algoritmos de aprendizagem de máquina para a realização de previsões de irradiação solar (Koo *et al.*, 2019; Yagli *et al.*, 2019; Benali *et al.*,

2019), tais métodos muitas vezes conseguem desempenhos adequados tanto para previsões de curto prazo como para previsões de longo prazo (1 dia a posteriori).

Utilização de técnicas de processamento de imagens de forma isolada ou em conjunto com modelos de aprendizagem de máquina também são comumente empregadas no contexto de previsões de irradiação solar de curto prazo (Pedro *et al.*, 2018; Pedro *et al.*, 2015; Pawar *et al.*, 2019).

Neste estudo, realizou-se previsões de médias diárias de irradiação solar global para um horizonte de um dia a posteriori por meio da aplicação de algoritmos de aprendizagem de máquina (Máquina de Aprendizagem Mínima e Regressão por Vetor Suporte) em dois conjuntos de dados: Um formado por variáveis exógenas (insolação, temperatura do ar, precipitação, etc.), variáveis endógenas (série temporal de irradiação solar) e variáveis temporais (ano, mês e dia da aquisição) totalizando 44 atributos e 3254 observações. O outro conjunto de dados consiste de preditores que são coeficientes wavelets obtidos por meio da aplicação da Transformada *Wavelet Haar* (Stollnitz *et al.*, 1995; Mallat, 1999) à uma série histórica de irradiação formada pelos valores de irradiação solar de 15 dias anteriores em conjunto com o valor da irradiação do dia atual totalizando 16 atributos e 3254 observações. O desempenho dos modelos é avaliado pelo cálculo das métricas Erro Médio Absoluto (MAE), Erro Médio por Viés (MBE), Raiz do Erro Quadrático Médio (RMSE), Raiz do Erro Quadrático Médio Relativo (rRMSE) e a Habilidade de Previsão (s).

O objetivo é o de avaliar se o uso de informações extraídas dos coeficientes wavelets da série histórica de irradiação solar (aproximação e detalhes) conseguem explicar de forma adequada a variabilidade da resposta quando comparado com o caso em que se considera o conjunto com variáveis exógenas, endógenas e temporais. Assim, o uso desta decomposição traz benefícios, como a redução no número de preditores, bem como o uso de apenas variáveis endógenas o que reduz o custo relacionado à aquisição dos demais atributos.

O restante do artigo é formado pela seção 2 que trata da forma de obtenção dos conjuntos de dados, além de abordar sobre o procedimento aplicado para a realização do treinamento dos algoritmos. Na seção 3 os modelos aplicados são detalhados, além de se ter a apresentação das métricas de erro utilizadas. Posteriormente, na seção 4, tem-se a apresentação dos resultados e, por fim, a seção 5 ilustra as conclusões obtidas.

2. DADOS

Os dados foram coletados na estação meteorológica da Universidade Federal do Ceará, na cidade de Fortaleza-CE, Brasil 3°43'6"S e 38°32'36"O, ao nível do mar, entre os anos de 1969 à 1979 e 1995 à 1997. Dentre as variáveis exógenas foi considerado a intensidade dos fenômenos climáticos do El Niño e da La Niña. Esta intensidade foi convertida em preditor pela atribuição de valores de -4 à 4, no qual os valores extremos representam as intensidades máximas dos fenômenos, sendo 4 para El Niño mais intenso e -4 para La Niña mais intensa. O zero representa uma condição de neutralidade, como indicado na Tab. 1. As intensidades desses fenômenos foram obtidas pelo *Golden Gate Weather Service*, serviço estadunidense, que avalia seus níveis pelo Índice de Niño Oceânico (ONI), obtido com base na diferença entre a temperatura medida e um valor pré-determinado para a região do pacífico com coordenadas 5°N-5°S, 120°E-170°W.

Tabela 1 – Intensidades do El Niño e da La Niña

| El Niño | | | La Niña | | |
|--------------|---------------------------------------|-----------------|--------------|---------------------------------------|-----------------|
| Intensidades | Aumento característico da temperatura | Valor atribuído | Intensidades | Redução característica da temperatura | Valor atribuído |
| Muito Forte | $T \geq 2^\circ$ | 4 | Muito Forte | $T \leq -2^\circ$ | -4 |
| Forte | $1,5^\circ \leq T \leq 1,9^\circ$ | 3 | Forte | $-1,5^\circ \geq T \geq -1,9^\circ$ | -3 |
| Moderado | $1^\circ \leq T \leq 1,4^\circ$ | 2 | Moderado | $-1^\circ \geq T \geq -1,4^\circ$ | -2 |
| Fraco | $0,5^\circ \leq T \leq 0,9^\circ$ | 1 | Fraco | $-0,5^\circ \geq T \geq -0,9^\circ$ | -1 |
| Neutro | $T < 0,5^\circ$ | 0 | Neutro | $T > -0,5^\circ$ | 0 |

A média diária da irradiação solar global foi obtida pelo uso de um actinógrafo R.Fuess Berlin-Steglitz, com numeração G6714. Este instrumento apresenta uma precisão na faixa de 15% a 20% com uma sensibilidade que compreende todo o espectro de radiação, com exceção das regiões do espectro que são absorvidas pela cúpula de vidro presente no aparato (ultravioleta e infravermelho distante, bandas que pouco contribuem para a energia total).

O uso deste sensor, que é considerado de terceira classe, é justificado pela ausência de dados de irradiação solar para a cidade de Fortaleza, principalmente dos anos de 1969 à 1979, que foram coletados por equipamentos de melhor precisão, como um pirânometro.

Para os preditores pressão, temperatura do ar, temperatura de bulbo úmido, umidade relativa, velocidade do vento e nebulosidade foram consideradas medições às 9h e 15h de cada dia. As irradiações dos 20 dias anteriores e a data de aquisição de cada observação (ano, mês e dia) também foram considerados atributos.

Os conjuntos de dados foram divididos aleatoriamente de tal forma que 70% ficou para o treinamento dos algoritmos e o restante constituiu o conjunto de validação. O ajuste dos métodos foi feito pela aplicação de validação cruzada 10-Fold (James *et al.*, 2013) no conjunto de treinamento.

2.1 Transformada *wavelet* haar

As funções *Wavelets* e suas transformadas são ferramentas matemáticas para análise e decomposição hierárquica de funções, ou seja, representação da função em diferentes níveis de resolução, ou escalas (Stollnitz *et al.*, 1995). Pode-se dizer que na escala mais grosseira um sinal pode ser representado por uma curva média (coeficientes de aproximação), e em cada escala mais refinada são acrescentados detalhes complementares (coeficientes de detalhes) que somados ao padrão médio recuperam exatamente a função original.

Similarmente aos senos e cossenos da transformada de Fourier (Brigham e Morrow, 1967), as *Wavelets* servem como base para representação de outras funções (Vidakovic e Mueller, 1991). De fato, as funções *Wavelets*, $\psi_i^j(x)$, de base são formadas pela realização de sucessivas expansões/contrações e translações na função *Wavelet* mãe, ψ , como verificado na Eq. (1).

$$\psi_i^j(x) = 2^{\frac{j}{2}} \psi(2^j x - i) \quad (1)$$

Para o caso das *Wavelets* de Haar a função mãe é dada pela Eq. (2).

$$\psi = \begin{cases} 1, & \text{se } 0 \leq x < \frac{1}{2} \\ -1, & \text{se } \frac{1}{2} \leq x < 1 \\ 0, & \text{se contrário} \end{cases} \quad (2)$$

Apesar de parecer complexo, a obtenção dos coeficientes *Wavelets* (aproximação e detalhes) para o caso em estudo, *Wavelet* de Haar discreta e unidimensional, pode ser realizada pela execução de sucessivos cálculos de médias aritméticas e subtrações. Em geral, para um vetor C_j de dimensão 2^j (representação diádica), tal que cada componente é representado pela notação $C_{j,k}$ com $k = \{0, 1, \dots, 2^j - 1\}$, tem-se que os coeficientes de aproximação e de detalhes resultantes da aplicação da transformada *wavelet* Haar neste vetor são dados pelas Eq. (3) e Eq. (4), respectivamente.

$$A_{j-1,i} = \frac{(C_{j,2i} + C_{j,2i+1})}{2} \quad (3)$$

$$D_{j-1,i} = C_{j,2i} - C_{j-1,i} \quad (4)$$

Para $i = \{0, 1, \dots, 2^{j-1} - 1\}$.

Neste estudo, a transformada foi aplicada em uma série temporal de irradiação solar, considerando os valores de irradiação solar de 15 dias anteriores em conjunto com o valor de irradiação do dia atual. Desta forma, nota-se que esta série trata-se de um vetor com 16 componentes no qual pela aplicação da transformada *Wavelet* de Haar resulta em um novo vetor com 16 coeficientes *Wavelets*, onde foi considerada a maior escala de refinamento obtendo assim um único coeficiente de aproximação com os restantes sendo coeficientes de detalhes.

3. METODOLOGIA

3.1 Modelo de baseline

O modelo de baseline escolhido atribui para uma nova observação a irradiação média do conjunto de treinamento (Hyndman e Athanasopoulos, 2018). Seja $T = \{G_1, G_2, \dots, G_n\}$ os valores de irradiação do conjunto de treinamento. Desta forma, o valor atribuído para uma nova observação é dado pela Eq. (5).

$$G = \frac{\sum_{i=1}^n G_i}{n} \quad (5)$$

Funciona como um modelo que fornece um desempenho mínimo, atuando como uma cota inferior para a acurácia dos demais modelos utilizados. Espera-se que o mesmo forneça os maiores valores para as métricas de erro RMSE e MAE.

3.2 Regressão por vetor suporte

Este modelo por mais que seja uma versão de uma Máquina de Vetor Suporte (SVM), apresenta a diferença de ser utilizado em problemas de regressão, enquanto o SVM é um método comumente aplicado no contexto de classificação. Uma das grandes vantagens em se trabalhar com o SVR se deve à possibilidade de construção de funções não lineares complexas, além de se ter a garantia de otimalidade global via a solução de um problema quadrático convexo (QP), treinamento efetivo e, geralmente, apresentando um desempenho preditivo positivo com um reduzido risco de *overfitting* graças ao uso do princípio de minimização do risco estrutural. Neste estudo, o algoritmo SVR aplicado foi do tipo ε -SVR, onde ocorre a construção de um ε -tubo em que os erros são ignorados no seu interior, mas penalizados fora do mesmo (Carrasco *et al.*, 2019; Drucker *et al.*, 1997).

Dado $x_i \in \mathbb{R}^n$ e as saídas correspondentes $y_i \in \mathbb{R}$, para $i = \{1, 2, \dots, m\}$ o ε -SVR linear busca obter um hiperplano $f(x) = w^T x + b$ pela construção de um ε -tubo como explicado anteriormente; para tanto, faz-se necessário a solução do QP.

$$\begin{aligned} \min_{w,b,\xi,\xi^*} \quad & \frac{1}{2} \|w\|^2 + C e^T (\xi + \xi^*) \\ \text{s. t.} \quad & y - (Aw + be) \leq \varepsilon e + \xi \\ & (Aw + be) - y \leq \varepsilon e + \xi^* \end{aligned}$$

Onde $A = [x_1 \ x_2 \ \dots \ x_m]^T \in \mathbb{R}^{m \times n}$, $y = (y_1, y_2, \dots, y_m) \in \mathbb{R}^m$, $e \in \mathbb{R}^m$ é um vetor de uns, $\xi, \xi^* \in \mathbb{R}^m$ são variáveis de folga que penalizam as observações que estão fora do ε -tubo e $C > 0$ é um fator que regulariza o *trade-off* de maior ou menor flexibilidade do modelo, como pode-se notar tanto a função objetivo como as funções restrições de desigualdade são convexas, portanto o QP é convexo e pode ser resolvido pela aplicação de algum método eficiente para a solução de programas convexas, como o método dos pontos interiores (Mehrotra, 1992).

O ε -SVR pode ser estendido para uma configuração não linear pelo uso de *kernels*. Isto pode ser feito pela aplicação de um artifício no problema dual correspondente (Drucker *et al.*, 1997). A versão de *kernels* para o SVR é dada pelo programa convexo.

$$\begin{aligned} \min_{\alpha, \alpha^*} \quad & \frac{1}{2} (\alpha - \alpha^*)^T K(A, A^T) (\alpha - \alpha^*) - y^T (\alpha - \alpha^*) + \varepsilon e^T (\alpha + \alpha^*) \\ \text{s. t.} \quad & e^T (\alpha - \alpha^*) = 0 \\ & 0 \leq \alpha, \alpha^* \leq C e \end{aligned}$$

Onde $K(A, A^T) \in \mathbb{R}^{m \times m}$ representa a matriz das funções *kernel* da forma $k_{is} = \kappa(x_i, x_s)$, $\kappa: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ e α, α^* são vetores de variáveis duais (multiplicadores de Lagrange do problema primal). O *kernel* utilizado neste estudo foi o radial gaussiano, que apresenta as funções de *kernel* da forma dada pela Eq. (6).

$$\kappa(x_i, x_s) = \exp\left(-\frac{\|x_i - x_s\|^2}{2\sigma^2}\right) \quad (6)$$

Neste trabalho, tomou-se $\varepsilon = 0,1$ e os parâmetros $\sigma = 0,157$ e $C = 1$ foram obtidos na etapa de treinamento. Uma representação esquemática do uso de *kernels* para se trabalhar com funções não lineares é ilustrado na Fig. 1

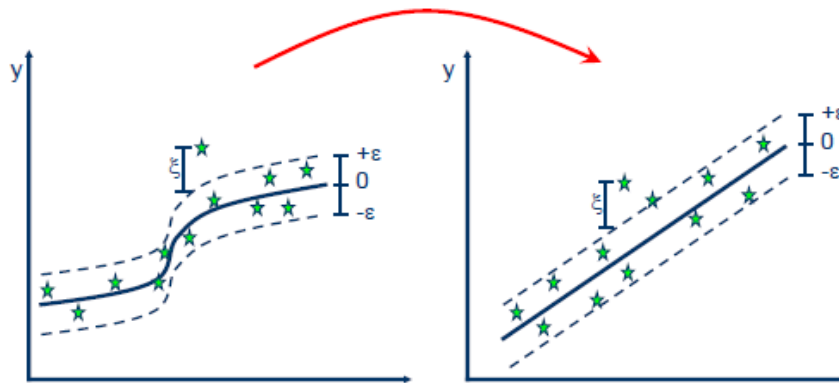


Figura 1 – Representação do uso de *kernels* no SVR

3.3 Máquina de aprendizagem mínima

Este algoritmo foi originalmente apresentado em trabalhos recentes (De Souza Júnior *et al.*, 2015). A metodologia consiste na construção de um mapeamento linear (regressão linear múltipla) entre as matrizes distâncias de entrada e saída. Posteriormente, o mapeamento linear ajustado (ótimo) é utilizado para estimar as

distâncias de K pontos de referência no espaço das saídas em relação à resposta que deseja-se determinar. Por fim, a saída é estimada via a solução de um problema de multilateração. A formulação geral do MLM é perfeitamente capaz de trabalhar com problemas de regressão não linear, bem como no contexto em que o espaço das saídas é multidimensional (De Souza Júnior *et al.*, 2015).

Seja um conjunto com N observações com atributos $X = \{x_i\}$, $i = \{1, 2, \dots, N\}$, $x_i \in \mathbb{R}^D$ e o conjunto com as saídas correspondentes $Y = \{y_i\}$, $i = \{1, 2, \dots, N\}$, $y_i \in \mathbb{R}^S$. Assume-se a existência de um mapeamento contínuo entre o espaço de entrada e o de saída dado pela Eq. (7).

$$y = f(x) + \varepsilon \quad (7)$$

Com y e x sendo vetores colunas e $\varepsilon \in \mathbb{R}^{N \times S}$ é a matriz de resíduos. O MLM opera em duas etapas.

1. Ajuste de uma regressão linear múltipla para obter matrizes distância de saída a partir de matrizes distância de entrada;
2. Estima a resposta desejada por meio da solução de um problema de multilateração.

Regressão linear múltipla entre matrizes distância. Esta etapa inicia-se com a escolha aleatória de algumas observações do conjunto de dados denominadas pontos de referência. Vale ressaltar que considera-se para cada ponto de referência os seus atributos e suas saídas correspondentes de forma separada, formando assim os conjuntos $\mu = \{m_k\}$, $k = \{1, 2, \dots, K\}$, $\mu \subset X$ que é constituído pelos atributos dos pontos de referência e o conjunto $T = \{t_k\}$, $k = \{1, 2, \dots, K\}$, $T \subset Y$ das saídas correspondentes. Define-se a matriz $D_x \in \mathbb{R}^{N \times K}$, tal que sua k -ésima coluna representa as distâncias $d(x_i, m_k)$, $i = \{1, 2, \dots, N\}$ entre os atributos x_i e os preditores do k -ésimo ponto de referência. De forma análoga, tem-se a matriz $\Delta_y \in \mathbb{R}^{N \times K}$, tal que sua k -ésima coluna representa as distâncias $\delta(y_i, t_k)$, $i = \{1, 2, \dots, N\}$ entre a saída y_i e a resposta do k -ésimo ponto de referência. Basicamente, assume-se que há uma relação linear entre as matrizes distância da entrada e matrizes distância da saída, obtendo assim um problema de regressão linear múltipla, como representado pela Eq. (8).

$$\Delta_y = D_x B + E \quad (8)$$

Com $B \in \mathbb{R}^{K \times K}$ sendo a matriz dos coeficientes. O ajuste deste modelo pode ser feito por meio da solução de um problema de mínimos quadrados, ou seja, pela minimização da função custo, dada pela Eq. (9).

$$RSS(B) = \text{tr} \left((\Delta_y - D_x B)^T (\Delta_y - D_x B) \right) \quad (9)$$

Para o trabalho em questão, o número de observações é tal que $N > K$. Desta forma, a solução por mínimos quadrados fornece a Eq.(10).

$$B^* = (D_x^T D_x)^{-1} D_x^T \Delta_y \quad (10)$$

Para o cálculo da inversa na Eq.(10), recomenda-se determinar a pseudo-inversa de Moore-Penrose (Ben-Israel e Greville, 2003; Penrose, 1955). Com isto, para uma observação de teste com atributos $x \in \mathbb{R}^D$, calcula-se as distâncias em relação aos vetores $\{m_k\}$, $k = \{1, 2, \dots, K\}$ obtendo assim o vetor $d(x, \mu) = [d(x, m_1), d(x, m_2), \dots, d(x, m_K)]$, onde por meio do modelo de regressão linear múltipla ajustado pode-se estimar as distâncias entre a resposta desejada y e as saídas conhecidas $\{t_k\}$, $k = \{1, 2, \dots, K\}$ dos pontos de referência, como ilustrado na Eq. (11).

$$\delta^*(y, T) = d(x, \mu) B^* \quad (11)$$

Obtendo assim o vetor $\delta^*(y, T) = [\delta^*(y, t_1), \delta^*(y, t_2), \dots, \delta^*(y, t_K)]$

Estimativa das respostas. Estimar as respostas conhecendo as saídas $\{t_k\}$, $k = \{1, 2, \dots, K\}$ dos pontos de referência e as estimativas das distâncias $\delta^*(y, T) = [\delta^*(y, t_1), \delta^*(y, t_2), \dots, \delta^*(y, t_K)]$ pode ser visto como um problema de multilateração (Niewiadomska-Szynkiewicz e Marks, 2009). Existem muitas estratégias que podem ser adotadas para a solução de tal problema, a utilizada neste estudo busca determinar \hat{y} por meio de uma interpretação geométrica, que gera K equações não lineares relacionadas à consideração de se ter K hiperesferas S -dimensional centradas nos t_k e passando por \hat{y} com raios iguais a $\delta^*(y, t_k)$, como ilustrado na Fig. (2) para o caso bidimensional.

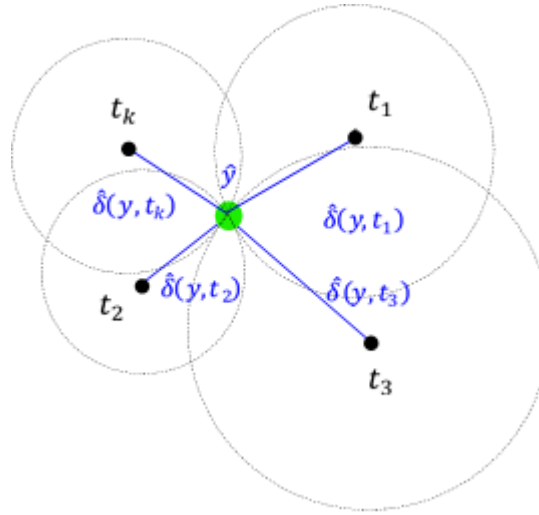


Figura 2 – Interpretação geométrica para a solução do problema de multilateração

Desta forma, tem-se a Eq. (12).

$$(\hat{y} - t_k)^T (\hat{y} - t_k) = (\delta^*(y, t_k))^2 \quad (12)$$

Assim a resposta é estimada pela minimização da função objetivo dada pela Eq. (13).

$$J(y^*) = \sum_{k=1}^K ((\hat{y} - t_k)^T (\hat{y} - t_k) - (\delta^*(y, t_k))^2)^2 \quad (13)$$

Por ser um problema de mínimos quadrados não linear, recomenda-se o uso do algoritmo Levenberg-Marquardt (Marquardt, 1963; Moré, 1978) para a solução de tal problema de otimização. Assim, percebe-se que o único hiperparâmetro presente no modelo é o número de pontos de referência K que pelo processo de treinamento foi obtido $K = 85$.

3.4 Métricas de erro

As métricas de erro utilizadas para realizar a avaliação do desempenho dos modelos de aprendizagem de máquina utilizados são apresentadas nesta seção.

Erro médio absoluto. O MAE calcula a média das diferenças absolutas entre o valor previsto, \hat{y}_i , e o valor real, y_i , isso é, não leva em consideração se o erro é para mais ou para menos e às diferenças absolutas não é atribuído peso, como indica a Eq. (14).

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (14)$$

Erro médio por viés. O MBE se assemelha ao cálculo do MAE, mas se diferencia deste por considerar o sinal do erro, isto é, não calcula o valor absoluto das diferenças. Desse modo, requer prudência em sua análise, uma vez que permite a compensação de erros (erros com sinais distintos). A mesma é dada pela Eq. (15).

$$MBE = \frac{\sum_{i=1}^n y_i - \hat{y}_i}{n} \quad (15)$$

Raiz do erro quadrático médio. O RMSE calcula a magnitude da média do erro pela raiz quadrada da média dos quadrados dos erros. Desse modo atribui um peso maior aos erros de maior magnitude, e peso menor aos erros de menor magnitude. É obtido na mesma unidade da variável em análise, sendo definida pela Eq. (16).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (16)$$

Raiz do erro quadrático médio relativo. O rRMSE é a razão entre o RMSE e a média dos valores da variável em análise, como dado pela Eq. (17).

$$rRMSE = \frac{\sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}}}{\bar{y}} \quad (17)$$

Onde $\bar{y} = \sum_{i=1}^n y_i$. Tal métrica foi utilizada porque fornece faixas de classificação do desempenho das previsões: $rRMSE < 10\%$ excelente, $10\% < rRMSE < 20\%$ bom, $20\% < rRMSE < 30\%$ razoável, $rRMSE > 30\%$ ruim (Li *et al.*, 2013).

Habilidade de previsão. Esta métrica avalia o desempenho do modelo aplicado quando comparado com a acurácia do modelo de baseline (Murphy, 1988), sendo definido pela Eq. (18).

$$s = 1 - \frac{RMSE}{RMSE_B} \quad (18)$$

Onde o $RMSE$ representa a o valor da métrica de erro para o modelo em análise e o $RMSE_B$ representa o valor da métrica de erro para o modelo de baseline.

4. RESULTADOS

Os resultados obtidos pela aplicação dos modelos no conjunto de dados constituídos de variáveis endógenas, exógenas e temporais são indicados na Tab. 2, já aqueles obtidos para o conjunto de dados formados pelos coeficientes *wavelets* estão reportados na Tab. 3.

Tabela 2 – Resultados para o conjunto com variáveis endógenas, exógenas e temporais

| Métricas | Modelos | |
|--------------------------|---------|-------|
| | MLM | SVR |
| MBE (W/m ²) | -1,97 | -8,34 |
| MAE (W/m ²) | 30,56 | 29,43 |
| RMSE (W/m ²) | 40,88 | 41,20 |
| rRMSE (%) | 20,70 | 20,86 |
| s (%) | 9,97 | 9,28 |

Tabela 3 – Resultados para o conjunto formado pelos coeficientes *wavelets*

| Métricas | Modelos | |
|--------------------------|---------|-------|
| | MLM | SVR |
| MBE (W/m ²) | -2,75 | -8,17 |
| MAE (W/m ²) | 32,51 | 31,78 |
| RMSE (W/m ²) | 43,82 | 44,48 |
| rRMSE (%) | 22,13 | 22,47 |
| s (%) | 3,51 | 2,05 |

Observando os resultados, percebe-se que a acurácia de previsão dos modelos para o caso em que se utilizou o conjunto com variáveis endógenas, exógenas e temporais foi melhor do que aquela obtida quando se aplicou os algoritmos no conjunto formado pelos coeficientes *wavelets* da série de irradiação solar. Entretanto, ressalta-se que esta queda de desempenho por parte dos modelos não é tão significativa, de tal maneira que a aplicação dos métodos no conjunto formado pelos coeficientes *wavelets*, que têm um menor número de preditores, fornece uma acurácia comparável ao caso em que se aplica os algoritmos no outro conjunto de dados o que indica que estes coeficientes são estatisticamente significativos para explicar a variância dos valores de irradiação solar para um dia a posteriori.

Além do benefício decorrente da redução de dimensionalidade do conjunto de dados, a utilização dos coeficientes *wavelets* trouxe também uma vantagem econômica, que é o uso de apenas valores de irradiação solar para a construção do conjunto de dados. Assim, elimina-se o custo relacionado à aquisição de outros preditores (precipitação, insolação, umidade, temperatura do ar, etc.).

Os resultados encontrados são condizentes com aqueles obtidos por outros autores que utilizam em suas metodologias de previsão algumas *wavelets* (Sharma *et al.*, 2016); Teixeira Júnior, 2015). No caso, estes trabalhos utilizaram decomposições *wavelets* em conjunto com redes neurais artificiais (Haykin, 2009) para a realização de previsões de irradiação solar. Em Sharma *et al.* (2016), foram obtidos valores variando na faixa de -13,27 W/m² à 14,36 W/m² para a métrica MBE, já em Teixeira Júnior *et al.* (2015), realizou-se previsões de irradiação solar em

diversas localidades (Cuiabá, Petrolina, Natal, Palmas, Joinville, etc.) onde foram obtidos valores variando na faixa de 7,78 W/m² à 54,45 W/m² para a métrica RMSE.

Destaca-se também o desempenho preditivo superior do MLM em relação ao SVR, que é um método clássico e bem posto. O MLM é um modelo relativamente recente que ainda carece de trabalhos voltados para a sua aplicação. Desta forma, tem-se a avaliação de um algoritmo que ainda não foi utilizado para propósitos de previsão de irradiação solar, obtendo uma acurácia superior ao de um método clássico com a vantagem de se ter apenas um único parâmetro de ajuste, K , na etapa de treinamento.

5. CONCLUSÃO

No trabalho em questão realizou-se previsões da média diária de irradiação solar global para um horizonte de previsão de 1 dia a posteriori pela aplicação de modelos de aprendizagem de máquina: Máquina de Aprendizagem Mínima e Regressão por Vetor Suporte, em dois conjuntos de dados, um com a presença de variáveis endógenas (série histórica dos valores de irradiação solar), variáveis exógenas (insolação, temperatura do ar, precipitação, etc.) e variáveis temporais (ano, mês e dia da medição) e outro formado por coeficientes (aproximação e detalhes) obtidos pela aplicação da transformada *wavelet* Haar a uma série temporal formada pelos valores de irradiação solar de 15 dias anteriores em conjunto com o valor de irradiação do instante atual. A acurácia de previsão foi avaliada pelo cálculo das métricas de erro MAE, MBE, RMSE, rRMSE, s. Os resultados indicam que os modelos de aprendizagem apresentaram um desempenho comparável para os dois conjuntos de dados em estudo. Para o conjunto de dados de maior dimensão, os algoritmos apresentaram uma melhor acurácia do que quando foram aplicados no conjunto formado pelos coeficientes *wavelets*, mas esta diferença não foi tão significativa. Desta forma, o uso da decomposição *wavelet* Haar traz as vantagens de se ter uma redução no número de preditores e o uso de apenas valores de irradiação solar para a construção do conjunto de dados ocasionando uma redução nos custos para aquisição de outros atributos.

Agradecimentos

Este estudo foi financiado em parte pela Coordenação para o Aperfeiçoamento de Pessoal de Ensino Superior – Brasil (CAPES) – Código de Financiamento 001, e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), órgãos governamentais brasileiros. O suporte recebido é reconhecido com gratidão.

REFERÊNCIAS

- Benali, L. *et al.*, 2019. Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components. *Renewable energy*, vol. 132, pp. 871-884.
- Brigham E. O., Morrow, R. E., 1967. The fast Fourier transform. *IEEE spectrum*, vol. 4, pp. 63-70.
- Ben-Israel, A., Greville, T. N. E., 2003. *Generalized inverses: theory and applications*, Springer Science & Business Media.
- Carrasco, M., López, J., Maldonado, S., 2019. Epsilon-nonparallel support vector regression. *Applied Intelligence*, pp. 1-14.
- Denholm, P., Margolis, R. M., 2007. Evaluating the limits of solar photovoltaics (PV) in traditional electric power systems. *Energy policy*, vol. 35, pp. 2852-2861.
- De Souza Júnior, A. H. *et al.*, 2015. Minimal learning machine: a novel supervised distance-based approach for regression and classification. *Neurocomputing*, vol. 164, pp. 34-44.
- Dong, Z. *et al.*, 2013. Short-term solar irradiance forecasting using exponential smoothing state space model. *Energy*, vol. 55, pp. 1104-1113.
- Drucker, H. *et al.*, 1997. Support vector regression machines. *Advances in neural information processing systems*. pp. 155-161.
- Hyndman, R. J., Athanasopoulos, G., 2018. *Forecasting: principles and practice*, OTexts.
- Inman, R. H., Pedro, H.T.C., Coimbra, C.F.M., 2013. Solar forecasting methods for renewable energy integration. *Progress in energy and combustion science*, vol. 39, n. 6, pp. 535-576.
- James, G. *et al.*, 2013. *An Introduction to Statistical Learning*, Springer.
- Koo, C. *et al.*, 2019. A novel estimation approach for the solar radiation potential with its complex spatial pattern via machine-learning techniques. *Renewable energy*, vol. 133, pp. 575-592.
- Kuhn, M., Johnson, K., 2013. *Applied predictive modeling*, Springer.
- Li, M. *et al.*, 2013. General models for estimating daily global solar radiation for different solar radiation zones in mainland China. *Energy conversion and management*, vol. 70, pp. 139-148.
- Mallat, S., 1999. *A wavelet tour of signal processing*, Elsevier.
- Marquardt, D. W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, vol. 11, pp. 431-441.
- Mejia, J. F., Giordano, M., Wilcox, E., 2018. Conditional summertime day-ahead solar irradiance

- forecast. *Solar Energy*, vol. 163, pp. 610-622.
- Moré, J. J., 1978. The Levenberg-Marquardt algorithm: implementation and theory. *Numerical analysis*. Springer, pp. 105-116.
- Murphy, A. H., 1988. Skill scores based on the mean square error and their relationships to the correlation coefficient. **Monthly weather review**, vol. 116, pp. 2417-2424.
- Niewiadomska-Szynkiewicz, E., Marks, M., 2009. Optimization schemes for wireless sensor network localization. *International Journal of Applied Mathematics and Computer Science*, vol. 19, pp. 291-302.
- Nonnenmacher, L., Kaur, A., Coimbra, C. F. M., 2016. Day-ahead resource forecasting for concentrated solar power integration. *Renewable energy*, vol. 86, pp. 866-876.
- Pawar, P. *et al.*, 2019. Detecting clear sky images. *Solar Energy*, vol. 183, pp. 50-56.
- Pedro, H. T. C., Coimbra, C. F. M., 2015. Nearest-neighbor methodology for prediction of intra-hour global horizontal and direct normal irradiances. *Renewable Energy*, vol. 80, pp. 770-782.
- Pedro, H. T. C *et al.*, 2018. Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts. *Renewable Energy*, vol. 123, pp. 191-203.
- Penrose, R., 1955. A generalized inverse for matrices. *Mathematical proceedings of the Cambridge philosophical society*. Cambridge University Press, pp. 406-413.
- Sharma, V. *et al.*, 2016. Short term solar irradiance forecasting using a mixed wavelet neural network. *Renewable Energy*, vol. 90, pp. 481-492.
- Stollnitz, E. J., Deroose, A. D., Salesin, D. H., 1995. Wavelets for computer graphics: a primer 1. *IEEE Computer Graphics and Applications*, vol. 15, pp. 76-84.
- Trapero, J. R., Kourentzes, N., Martin, A., 2015. Short-term solar irradiation forecasting based on dynamic harmonic regression. *Energy*, vol. 84, pp. 289-295.
- Teixeira Júnior, L. A. *et al.*, 2015. Artificial neural network and wavelet decomposition in the forecast of global horizontal solar radiation. *Pesquisa Operacional*, vol. 35, pp. 73-90.
- Vidakovic, B., Mueller, P., 1991. Wavelets for kids: A tutorial introduction, Duke university.
- Yagli, G. M., Yang, D., Srinivasan, D., 2019. Automatic hourly solar forecasting using machine learning models. *Renewable and Sustainable Energy Reviews*, vol. 105, pp. 487-498.

SOLAR IRRADIANCE FORECASTING DAY-AHEAD USING TRANSFORM WAVELET HAAR FOR EXTRACTION OF FEATURES

Abstract. *In this paper, perform a global solar irradiation forecasting day-ahead by using two machine learning models: Minimal Learning Machine and Support Vector Regression. These algorithms were applied in two data sets. One of them consists of attributes formed by exogenous variables (insolation, air temperature, temperatures, etc.), endogenous variables (solar irradiation historical data) and temporal variables (year, month and day of acquisition). The other data set consists of features that are coefficients wavelets obtained by applying the Haar Wavelet Transform at one historical series of solar irradiation, considering the solar irradiation values from the previous 15 days together with the irradiation value of the current day. For the first case, formed a set of 44 attributes and 3254 samples, and the second set was formed of 16 attributes and 3254 samples. The average root mean square error obtained by applying the models in the first data set was 41.04 W / m² with an average forecast skill of 9.62%, while for the second case the average root mean square error was 44.15 W / m² with an average forecasting skill of 2.78%. Thus, it is clear that, using only the coefficient wavelets as features for evaluated models, it was possible to reach a forecast accuracy comparable to the case in which it considered a set of exogenous, endogenous and temporal variables. Noting that the use of this decomposition has the benefits of a reduction in the number of predictors, as well as the use of only endogenous variables, which reduces the acquisition cost the others features.*

Key words: *Solar Energy, Machine Learning, Haar Wavelet Transform*