



Grupo de Estudo de Geração Eólica, Solar e Armazenamento-GES

UM MODELO DE PREVISÃO DE CURTO PRAZO DE ENERGIA SOLAR FOTOVOLTAICA, USANDO MÁQUINA DE VETORES DE SUPORTE (SVM)

ROBERTO PEREIRA CALDAS(1); ALBERT C GEBER DE MELO(2);
JOSÉ FRANCISCO MOREIRA PESSANHA(2); DJALMA MOSQUEIRA FALCÃO(1);
UFRJ(1);CEPEL - UERJ(2);

RESUMO

A capacidade de previsão da geração solar fotovoltaica (SPV) é um desafio para a sua integração a sistemas elétricos. Fonte intermitente, a SPV apresenta elevada dispersão e múltiplos ciclos sazonais, com variações horárias, mensais e anuais. Este trabalho apresenta uma metodologia de previsão de curto prazo - até um dia à frente, com resolução horária, para uma planta SPV. Com dois módulos executados sequencialmente, filtra medições de geração solar e irradiação solar, e ajusta modelos SVM (Suport Vector Machine) com dados de geração solar e de previsão meteorológica. Aplicada aos dados de um sistema real na Itália, apresentou bons resultados.

PALAVRAS-CHAVE

Energia Solar fotovoltaica, Irradiação Solar, Aprendizado de Máquina (ML), Máquina de Vetores de Suporte (SVM).

1.0 - INTRODUÇÃO

Seguindo a tendência mundial [1], verifica-se no Brasil, um grande impulso na implantação de novas energias renováveis, como a energia eólica, que vem crescendo exponencialmente na matriz elétrica brasileira, e a energia solar fotovoltaica (SPV), que tem apresentado relevante evolução tecnológica, traduzindo-se em maior eficiência na conversão para energia elétrica e custos decrescentes de produção de seus componentes. No entanto, a intermitência da geração SPV, dada a natureza intrínseca da variabilidade da irradiação solar, é um desafio em termos de sua integração a sistemas de energia elétrica [2,3]. Os valores de geração apresentam ampla dispersão e são sujeitos a múltiplos ciclos sazonais com variações horárias, mensais e anuais, e, portanto, podem ser de difícil previsão. Uma capacidade inadequada de prever a produção de energia SPV afeta adversamente a estabilidade, a confiabilidade e a programação da operação do sistema de energia, e também o seu benefício econômico.

A geração SPV depende fundamentalmente da irradiação solar ao nível do solo, incidente em um plano horizontal (GHI), que por sua vez é influenciada por fatores meteorológicos, como a nebulosidade, precipitação pluviométrica, temperatura ambiente, pressão atmosférica, direção do vento e umidade, entre outras. Estas características têm promovido um amplo espectro de estudos e desenvolvimentos de métodos, técnicas e modelos de previsão. Em [4], apresenta-se uma revisão bibliográfica sobre os modelos de previsão de geração SPV, abrangendo o período 2007-2016, os quais foram classificados em três categorias: *físicos* (uso de equações analíticas para modelar o sistema SPV); *estatísticos* (modelos estocásticos para previsão de séries temporais e métodos de aprendizagem de máquina); e *abordagens híbridas*. A Figura 1 ilustra a distribuição do número de trabalhos, classificados pelos principais métodos de previsão. Observa-se que 72% dos trabalhos

correspondem aos métodos estatísticos, dentre os quais se destaca o método SVM (18%), o qual tem ganho impulso entre os principais métodos de aprendizado de máquina – o que motivou a sua escolha neste trabalho.

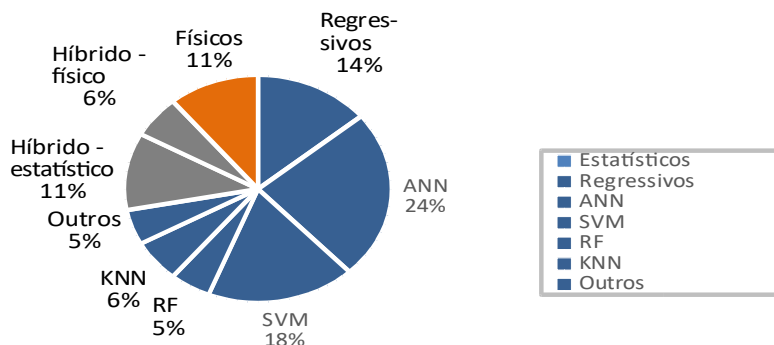


FIGURA 1 - Distribuição do número de publicações entre 2007 e 2016 – previsão de energia SPV [4].

O presente artigo visa apresentar as principais características de uma metodologia desenvolvida para a previsão de curto prazo – até um dia à frente, com resolução horária, de geração SPV de uma planta fotovoltaica, horizonte compatível com a programação diária da operação. A metodologia proposta organiza-se em dois módulos básicos: (i) tratamento de dados; e (ii) modelo de previsão. O modelo de previsão baseia-se em Máquina de Vetores de Suporte (Support Vector Machine – SVM) [5], onde se ajusta, para cada hora, uma regressão não linear, na qual a geração SPV é explicada em função de variáveis selecionadas, e.g., temperatura, cobertura de nuvens e previsão do tempo. O desempenho de um modelo de previsão depende da qualidade dos dados utilizados no seu ajuste; assim, na etapa de tratamento de dados, que precede o modelo de previsão, as medições de geração SPV e irradiação solar global horizontal foram analisadas de forma conjunta, empregando técnicas estatísticas [6-9] e de mineração de dados [10-13] para a correção de dados discrepantes e preenchimento de lacunas. A metodologia desenvolvida foi aplicada a um sistema SPV localizado em Monteroni di Lecce, Puglia, Itália, com com 4.710 m² de área efetiva, composto por 3.000 módulos de 320 WP, totalizando 960 kWp e eficiência nominal de 19,6% (condições STC) [14,15], apresentando bons resultados.

2.0 - MÉTODO SVM

Proposta por Vapnik na primeira metade dos anos 1990 [5], as Máquinas de Vetores de Suporte podem ser usadas em problemas de classificação de padrões e em problemas de regressão. Neste trabalho, objetiva-se modelar a relação entre a geração SPV e suas variáveis explicativas por meio de modelo de regressão não linear. Portanto, trata-se de uma aplicação do algoritmo SVM em um problema de regressão.

Considere uma série temporal com n observações de geração SPV $y_i \forall i=1,n$ acompanhadas das respectivas variáveis explicativas representadas pelo vetor $X_i \forall i=1,n$. A previsão de geração SPV pode ser efetuada pelo ajuste do seguinte modelo de regressão:

$$y = f(X) = w^T X + b \quad (1)$$

Conforme ilustrado na Figura 2a, a equação de regressão em (1) pode ser tratada como um hiperplano no centro de um hipertubo (*hypertube*) de largura 2ε com $\varepsilon > 0$ que envolve todas as observações.

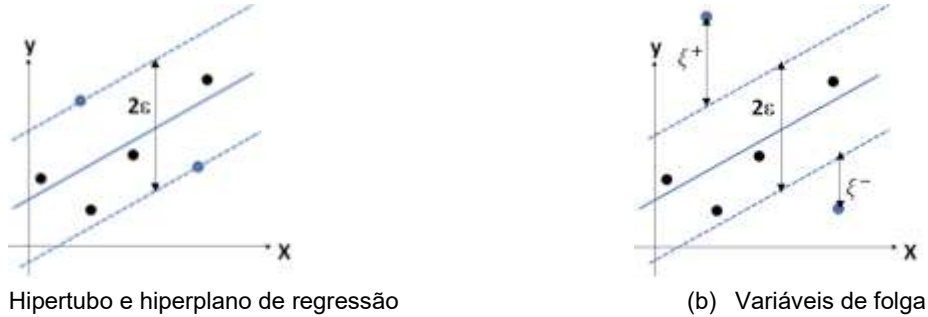


FIGURA 2 - Hipertubo, hiperplano de regressão e variáveis de folga

A melhor forma de alinhar o hipertubo aos dados consiste em posicioná-lo de maneira que as observações fiquem próximas da parede do hipertubo. Apesar do alinhamento ótimo do hipertubo aos dados não ser efetuado pelo método dos mínimos quadrados ordinários, o seu posicionamento é determinado pela solução do seguinte problema de programação quadrática:

$$\begin{aligned} \text{Min}_{w, b} \quad & \frac{1}{2} w^T w \quad \zeta \quad y_i - (w^T X_i + b) \leq \varepsilon \\ \zeta (w^T X_i + b) - y_i \quad & \zeta \quad \zeta \quad \zeta \quad \zeta \end{aligned} \quad (2)$$

Contudo, não se pode garantir que todos os dados estejam dispostos dentro de um hipertubo de largura 2ε . Para acomodar estas situações o modelo em (2) deve ser modificado com a adição de variáveis de folga ξ indicadas no modelo em (3). Geometricamente, as variáveis de folga representam as distâncias entre as paredes do hipertubo e as observações exteriores ao hipertubo, conforme indicado na Fig. 2b.

$$\zeta \quad (3)$$

No problema de otimização em (3), a constante C é positiva e permite controlar o trade-off entre a maximização da margem dentro hipertubo e a minimização das variáveis de folga, um mecanismo de prevenção ao sobreajuste (*overfitting*). Adicionalmente, a solução do problema de otimização em (3) na sua forma dual em (4) é mais simples de ser alcançada.

$$\zeta \quad (4)$$

em que λ denota os multiplicadores de Lagrange associados com as restrições do problema de otimização em (3).

As relações entre as soluções dos problemas primal em (3) e dual em (4) são estabelecidas pelas equações (5) e (6) a seguir:

$$w = \sum_{i=1}^n \lambda_i \zeta_i \quad (5) \quad b = \frac{1}{n} \sum_{i=1}^n w^T X_i - y_i \quad (6)$$

Substituindo a equação (5) em (1) tem-se a equação do predictor em (7) para uma entrada X .

$$f(X) = \sum_{i=1}^n \lambda_i \zeta_i \quad (7)$$

Uma motivação das máquinas de vetores de regressão é a facilidade com que se pode estender o modelo linear no modelo não linear. A transformação das equações (1) e (7) e do problema de programação matemática em (4) para o caso não linear consiste em substituir o produto $X_i^T X_j X_i^T X$ por uma função núcleo (*kernel trick*), que mapeia o espaço de entrada em um espaço de maior dimensão:

$$\Phi(X_i, X_j) = \phi(X_i) \phi(X_j) \quad (8)$$

O modelo de regressão linear resultante é ilustrado em (9). Na sequência, em (10) tem-se a equação de regressão para o caso não linear:

$$\hat{y} \quad (9)$$

$$f(X) = \sum_{i=1}^n \hat{y}_i \quad (10)$$

Os resultados de uma Máquina de Vetores de Suporte para regressão são sensíveis à definição dos seus hiperparâmetros, um conjunto formado pela constante C, pela largura do hipertubo ϵ , pela função núcleo (*kernel*) e seus parâmetros. A identificação da melhor configuração envolve uma busca em grade de possíveis valores para cada hiperparâmetro.

3.0 - METODOLOGIA PROPOSTA PARA PREVISÃO DE CURTO PRAZO DE ENERGIA SOLAR FOTOVOLTAICA

A metodologia proposta é constituída por dois módulos, executados sequencialmente. Inicialmente, as medições de geração SPV e de GHI são tratadas pelo módulo de filtragem de dados. Em seguida, os dados de geração SPV, em conjunto com dados de previsão de variáveis meteorológicas, são utilizados no ajuste, para cada hora, de um modelo SVM possibilitando a previsão de um perfil horário de geração SPV um dia à frente.

3.1 Tratamento dos dados

O modelo de previsão proposto baseia-se no histórico de medições de geração SPV. Para o bom desempenho do modelo de previsão é fundamental que os dados observados tenham a maior qualidade possível. No caso ideal, os dados devem estar livres de erros e perturbações provocados, por exemplo, por falhas nos equipamentos e sensores de medição (medidores de potência, sensores, subsistemas de transmissão e armazenamento de dados) ou eventos no sistema SPV, como defeitos nos painéis e equipamentos auxiliares (inversores, transformadores), e demais causas não naturais que afetam a trajetória da potência de saída. Erros e perturbações levam a valores discrepantes (*outliers*), descontinuidades e lacunas de dados (*missing data*), comprometendo o ajuste do modelo e a qualidade das suas previsões. Como medida de mitigação, é necessário o tratamento prévio dos dados de geração SPV. Dada a forte dependência entre geração SPV e GHI, a estratégia adotada consistiu na filtragem dos dados das medições de GHI e na estimação do seu efeito na geração SPV, por meio de equação de regressão linear.

Mais especificamente, a estratégia adotada no tratamento de dados compreende cinco etapas:

- (i) *tratamento de valores nulos de GHI e geração, no horário de 4h00 a 21h00*: inicialmente, ajusta-se um modelo de regressão linear simples aos dados brutos de geração SPV e GHI. Em seguida, os valores nulos de GHI, foram (a) substituídos por valores de dados de reanálise no horário e dia correspondentes; e (b) os respectivos valores de geração, pelos valores esperados determinados pela reta de regressão. E, no caso de valores medidos nulos apenas de geração, os mesmos foram substituídos conforme (b);
- (ii) *filtragem dos registros superestimados em cada hora H*: inicialmente analisam-se os *boxplots* [6,9] dos registros horários de GHI e de geração SPV para cada hora do dia; portanto, 24 *boxplots* para cada variável. Os *boxplots* permitem identificar os valores discrepantes em cada hora H ($\forall H=1,24$), especialmente as medições superestimadas e que correspondem aos valores acima da cerca superior do *boxplot* da hora H, cujo valor é determinado pela seguinte expressão:

$$\text{cerca superior}_H = \text{terceiro quartil}_H + 1,5 \times (\text{terceiro quartil}_H - \text{primeiro quartil}_H) \quad \forall H=1,24 \quad (11)$$

A filtragem dos registros superestimados em cada hora H consiste em substituí-los pelo respectivo valor horário da cerca superior.

- (iii) *identificação e preenchimento dos perfis diários com lacunas de dados nos registros horários de GHI*: para os dias sem medições e os perfis com lacunas de dados entre 4h:00 e 21h00 que excedam 4 horas de duração realiza-se a substituição por estimativas oriundas de reanálises [16]; já para os demais dias, aplica-se um algoritmo regressão não paramétrica (e.g., LOESS [7,12]) em cada perfil diário de GHI, com a finalidade de corrigir possíveis valores discrepantes e preencher eventuais lacunas com durações

inferiores a 4 horas, que possam estar presentes nos dados. Para um dado perfil diário de GHI, o algoritmo LOESS gera a sua versão suavizada, cujos valores horários substituem as medições discrepantes e as lacunas de dados;

- (iv) *análise de agrupamentos (clusters analysis)* [6,9,10,11] *dos perfis diários de GHI*: classificam-se os perfis por similaridade em um determinado número de grupos (*clusters*), de tal forma que os perfis semelhantes sejam classificados em um mesmo grupo, enquanto perfis distintos sejam classificados em grupos diferentes. Na execução desta tarefa pode-se contar com a ajuda de um Mapa de Kohonen [10], uma rede neural artificial autoorganizável e com treinamento não supervisionado. A elevada similaridade entre os perfis classificados em um mesmo grupo (*cluster*) permite obter um perfil típico para cada grupo, bem como identificar perfis atípicos ou observações discrepantes que ainda possam estar presentes nos dados. A identificação de valores horários discrepantes em cada *cluster* pode ser realizada por meio de *boxplots* horários, como na etapa inicial da filtragem, porém os *boxplots* são construídos apenas com base nos perfis classificados em um mesmo cluster e os valores discrepantes em uma hora H são substituídos pela mediana da GHI na hora H; e
- (v) *ajuste de um novo modelo de regressão linear simples (geração SPV explicada pela GHI)*: o modelo ajustado permite calcular intervalos de predição (i.e., com 95% ou 99% de confiança), da geração SPV condicionada ao valor de GHI. As medições de SPV discrepantes correspondem às observações fora dos limites dos intervalos de predição e são substituídas por valores esperados determinados pela reta de regressão.

3.2 – Previsão do perfil horário da geração SPV um dia à frente

Para previsão da potência gerada pelo sistema PV em foco, a metodologia proposta baseou-se na disponibilidade de previsões meteorológicas, com resolução horária, para um dia à frente. Entre as previsões meteorológicas requisitadas, destacam-se as seguintes previsões horárias:

- previsão da temperatura ambiente
- previsão da precipitação pluviométrica
- previsão da cobertura de nuvens
- previsão do tempo

A previsão do tempo é uma variável categórica, cujo valor fornece um descritivo do tempo em cada hora do dia, por exemplo: parcialmente nublado, céu limpo, sol, possibilidade de chuva irregular, chuvisco irregular, aguaceiros fracos, aguaceiros moderados ou fortes, períodos de chuva forte e chuva forte, entre outros. Naturalmente, algumas categorias podem ser pouco frequentes e, por esta razão, foram agregadas previamente em um menor número de classes. Um histórico das previsões supracitadas, para a localidade de Salento, pode ser obtido em [17].

O predictor proposto admite que nos períodos da madrugada (0h00 – 4h00) e noite (21h00-24h00) a geração fotovoltaica é nula. Já para cada hora H no intervalo entre 5h00 e 20h00 ajuste-se um modelo SVM com os dados para a mesma hora H nos 60 dias anteriores ao dia da previsão. Logo, ao todo são ajustados 16 modelos SVM.

Para cada hora H do horizonte de previsão constrói-se um conjunto de dados para o treinamento do respectivo modelo SVM, responsável pela previsão da geração fotovoltaica na hora H. O treinamento do SVM é do tipo supervisionado, assim o conjunto de dados para o treinamento inclui dois tipos de variáveis:

- uma variável resposta, neste caso, a geração fotovoltaica
- cinco variáveis explicativas, GHI, a temperatura ambiente, a precipitação pluviométrica, a cobertura de nuvens e as condições meteorológicas previstas.

Portanto, para cada hora H do horizonte, a previsão da geração fotovoltaica é determinada pelas previsões de GHI, temperatura ambiente, precipitação pluviométrica, cobertura de nuvens e condições meteorológicas para a hora H.

A formação do conjunto de treinamento para a previsão da geração em uma hora H do horizonte de previsão

10 a 13 de novembro de 2019
Belo Horizonte - MG

tem um olhar retrospectivo, i.e., cada hora T do conjunto de treinamento é formado pelos valores de geração fotovoltaica na mesma hora H em uma janela de W dias anteriores ao dia da previsão, e.g., $W=60$ dias. Adicionalmente, cada observação passada T vem acompanhada das respectivas previsões de tempo. Assim, não necessariamente as W observações participam do treinamento do modelo SVM para a hora H, mas apenas aquelas nas quais as respectivas ocorrências de tempo conferem com a previsão de tempo para a hora. Por exemplo, se para a hora H do horizonte de previsão há uma expectativa de "Sol", então, o conjunto de treinamento é formado pelas observações passadas com "Sol" onde o horário T é igual a H.

4.0 - ESTUDO DE CASO

O estudo baseou-se no sistema PV [14,15], localizado no campus da Universidade de Salento, em Monteroni di Lecce (LE), Puglia, Itália ($40^{\circ}19'32''16N$, $18^{\circ}5'52''44E$) desenvolvido no âmbito do projeto europeu "7th Framework Programme Building Energy Advanced Management Systems (BEAMS)". O sistema PV, com 4.710 m^2 de área efetiva, é composto por 3.000 módulos de 320 W_P . Com potência nominal total de 960 kW_P e eficiência nominal de 19,6% (condições STC), foi montado em estrutura metálica num estacionamento para veículos, segundo duas seções de inclinações diferentes (3° e 15°), e orientado para o sudeste (azimute de -10°). A primeira seção do sistema possui 1.104 módulos e uma área efetiva de $1.733,3 \text{ m}^2$, e a segunda, 1.896 módulos e área efetiva de $2.976,7 \text{ m}^2$. Conta ainda com inversores para injeção da energia gerada na rede, um sensor de temperatura ambiente, dois outros conjuntos de sensores (temperatura dos módulos e de irradiação solar global), um para cada seção do sistema, e com um sistema SCADA para coleta e armazenamento de dados.

4.1 - Dados utilizados

Foram utilizados os seguintes conjuntos de dados:

- medições do sistema PV da Universidade de Salento, de período de 21 meses (de abril de 2012 a dezembro de 2013), abrangendo medições, em base horária média, de irradiação solar global nas duas seções de painéis SPV e da geração total dos painéis PV [14,15];
- dados horários de reanálise de irradiação solar direta e difusa, nas coordenadas geográficas do sistema SPV, produzidos pelo modelo SARAH (Surface Solar Radiation Data Set – Heliosat) [16]; e
- dados de previsões meteorológicas de variáveis meteorológicas, nas coordenadas geográficas da instalação fotovoltaica, em base horária, abrangendo temperatura ambiente, precipitação pluviométrica, cobertura de nuvens e condições meteorológicas [17].

4.2 - Resultados do tratamento de dados

Na Figura 3 são ilustradas situações de tratamentos de dados GHI por meio da metodologia proposta. Na Figura 3a ilustra-se um exemplo de preenchimento de lacuna de dados com valores de reanálise [16] (dados observados em cor azul e dados tratados em vermelho). Já na Figura 3b apresenta-se um exemplo de correção de valores discrepantes, usando o mesmo padrão de cores anterior.

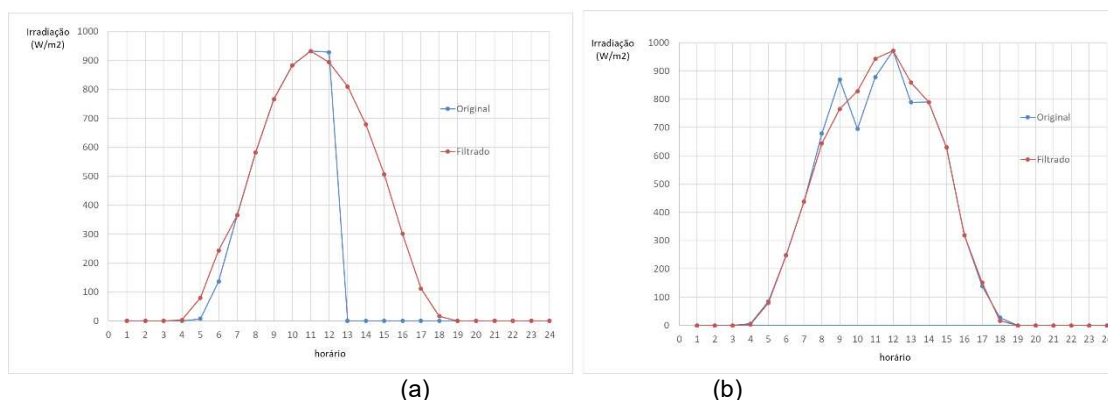


FIGURA 3 – Exemplos de tratamento de dados de GHI.

Já a Figura 4 apresenta, em termos de diagramas de *boxplot*, o resultado final obtido a partir das etapas de tratamento para os valores aberrantes, comparando a situação encontrada com os dados originais, “brutos”, de irradiação solar, e a posterior, com os dados “filtrados”. Cerca de 13% dos dados “brutos” observados, de GHI e geração do sistema SPV foram tratados. Em termos de correlações entre dados brutos e tratados, obteve-se valores de 95,5%, para os de irradiação, e de 95,7%, para os de potência.

4.3 - Resultados da previsão do perfil horário da geração SPV um dia à frente

A partir dos dados de geração SPV obtidos com a aplicação da etapa de filtragem, em conjunto com dados de previsão das variáveis meteorológicas descritas, foi ajustado, para cada hora, um modelo SVM, com base em um *kernel* do tipo radial (*radial basis function* - RBF) [5], para permitir a previsão de um perfil horário de geração SPV um dia à frente. A Figura 5 apresenta as previsões obtidas com a metodologia proposta para um conjunto de 8 dias do ano de 2013, acompanhadas das respectivas médias dos erros absolutos (MAD) em relação aos valores realizados, expressos em kW.

Para avaliar a qualidade das previsões produzidas pela metodologia proposta, adotou-se como *benchmark* as previsões obtidas pelo modelo *naive*, i.e., a previsão do perfil horário da geração SPV é determinada pelo perfil do dia anterior. Os resultados são também apresentados na Figura 5, com os respectivos MADs.

Verifica-se que, na maioria dos casos, o predictor proposto produziu um perfil horário com forma similar ao perfil verificado, apesar dos desvios entre os valores previstos e observados. Por outro lado, quando o perfil do dia da previsão é idêntico ao perfil do dia anterior, o predictor *naive* apresenta um melhor desempenho. Entretanto, estas situações não são frequentes, o que implica na necessidade de utilização de um predictor que considere variáveis explicativas, como o proposto neste trabalho.

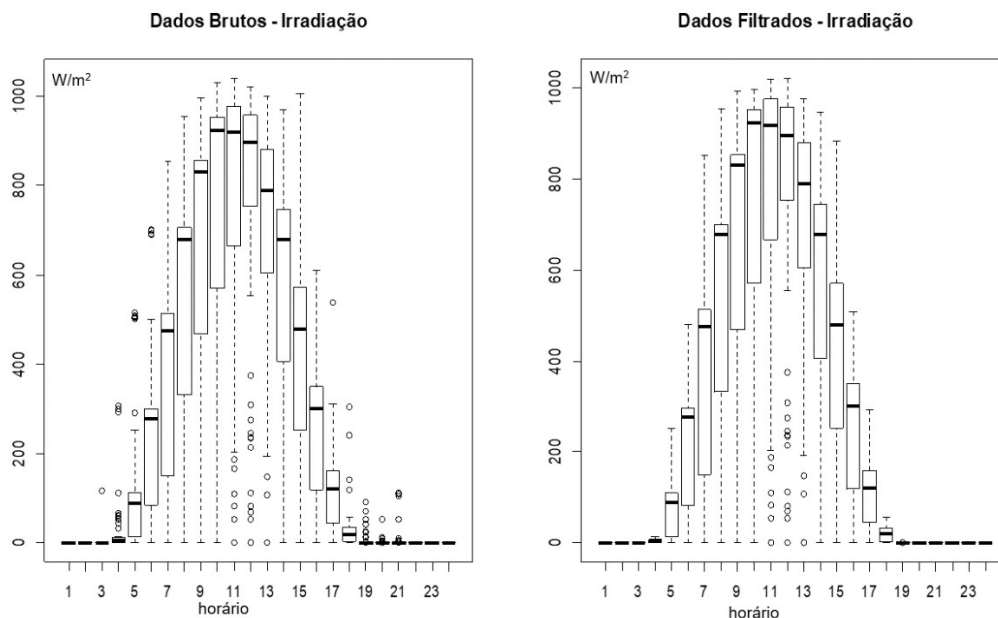
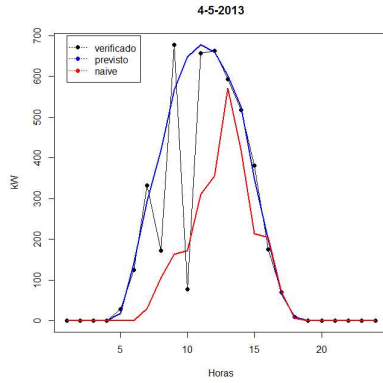
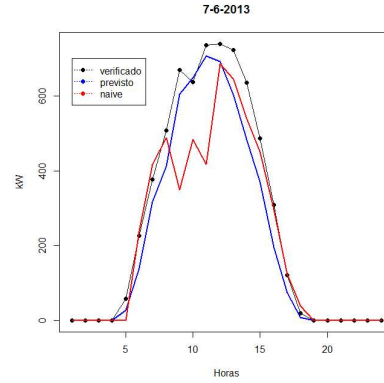


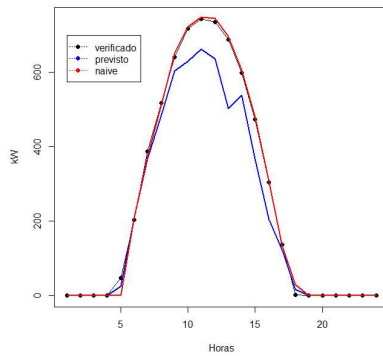
FIGURA 4 – Comparação entre os dados de GHI observados e tratados.



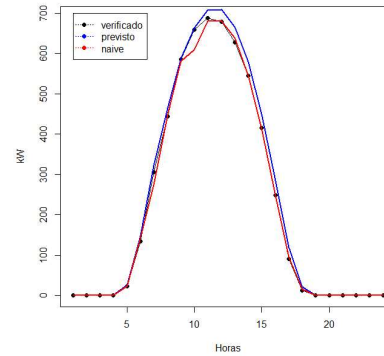
$mad_{proposto} = 73,4$ $mad_{naive} = 140,5$



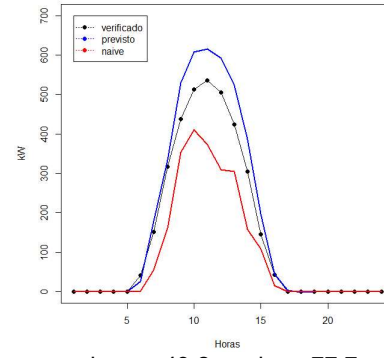
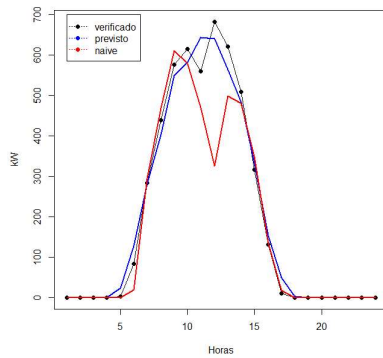
$mad_{proposto} = 65,7$ $mad_{naive} = 81,2$



$mad_{proposto} = 58,0$ $mad_{naive} = 10,5$



$mad_{proposto} = 19,4$ $mad_{naive} = 8,8$



$mad_{proposto} = 43,2$ $mad_{naive} = 77,7$

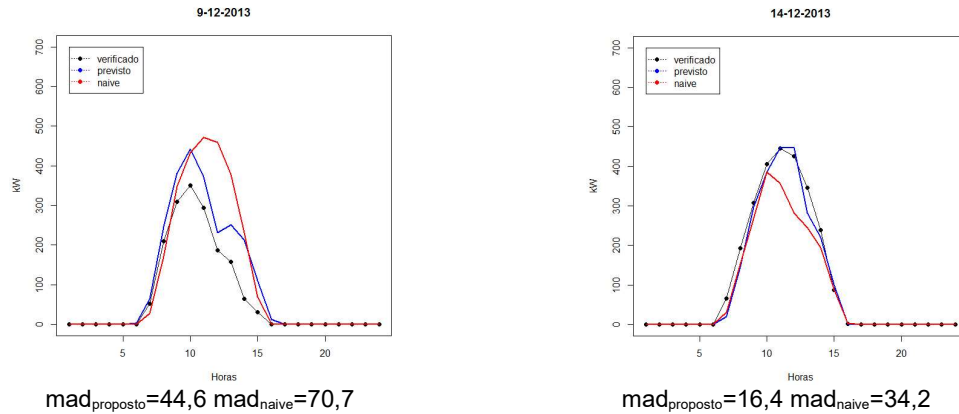


FIGURA 5 – Valores observados e previsões de geração SPV pela metodologia proposta e pelo método *naive*.

5.0 - CONCLUSÕES

Este trabalho apresentou as principais características de uma metodologia desenvolvida para a previsão de curto prazo – até um dia à frente, com resolução horária, de geração solar para uma planta fotovoltaica. A metodologia proposta é constituída por dois módulos, executados sequencialmente. Inicialmente, as medições de geração solar e de irradiação horizontal global são tratadas pelo módulo de filtragem de dados. Em seguida, os dados de geração solar obtidos, em conjunto com dados de previsão de variáveis meteorológicas, são utilizados no ajuste, para cada hora, por meio de um modelo SVM (*Support Vector Machine*) - possibilitando a previsão de um perfil horário de geração solar um dia à frente, compatível, portanto, com o horizonte da programação diária da operação.

A metodologia desenvolvida foi aplicada a um sistema SPV localizado na região mediterrânea da Itália com capacidade total de 960 kW_p, sendo o seu desempenho aferido pelas médias dos erros absolutos (MAD) em relação aos valores realizados. Adicionalmente, a qualidade das previsões produzidas pela metodologia proposta, os seus resultados foram comparados com aqueles obtidos pelo modelo *naive*, i.e., a previsão do perfil horário da geração SPV é determinada pelo perfil do dia anterior. Constatou-se que, na maioria dos casos, o predictor proposto apresentou bons resultados, produzindo um perfil horário com forma similar ao perfil verificado. Assim, os resultados obtidos encorajam a continuidade da pesquisa para o aprimoramento da metodologia desenvolvida e apontam para a necessidade de disponibilização de uma base pública de dados solares, com discretização horária, e.g., a partir dos projetos vencedores nos leilões públicos de compra de energia elétrica.

6.0 - REFERÊNCIAS BIBLIOGRÁFICAS

- (1) INTERNATIONAL ENERGY AGENCY World Energy Outlook 2018; IEA Publications (www.iea.org), 2018.
- (2) FALCAO D.M., TARANTO, G.N., HINCAPIE, C.C.O. Chronological Simulation of the Interaction between Intermittent Generation and Distribution Network. In: 2013 IEEE PES Innovative Smart Grid Technologies LATIN AMERICA (ISGT LA 2013), 2013, São Paulo.
- (3) ROSAS, G.B., LOURENCO, E.M., FALCAO, D.M., FERNANDES, T.S. Superação de Equipamentos, Reserva de Energia e Controle de Tensão em Sistemas com Forte Penetração de Energias Eólica e Solar. In: XIV SEPOPE Symposium of Specialists in Operation Planning and Electrical Expansion, 2018, Recife.
- (4) ANTONANZAS, J., OSORIO, N., ESCOBAR, R., URRACA, R., MARTINEZ-DE-PISON, F.J., ANTONANZAS-TORRES, F. Review of photovoltaic power forecasting; *Solar Energy*, vol. 136, PP. 78–111, 2016..
- (5) HAMEL, L. Knowledge Discovery with Support Vector Machines; Wiley-Interscience, 1 edition, 2009.



- (6) JOHNSON, R.A., WICHERN, D.W Applied Multivariate Analysis, 4th ed., New Jersey: Prentice Hall, 1998.
- (7) PESSANHA, J.F.M., JUSTINO, T.C., MACEIRA, M.E.P. Metodologia para filtragem de registros de carga; XII SEPOPE 17, PP. 1–10, 2012.
- (8) PESSANHA, J. F. M., CASTELLANI, V., JUSTINO, T.C., PENNA, D.D.J., MACEIRA, M.E.P. Uma metodologia para filtragem de medições anemométricas, Learning and Nonlinear Models, v. 10, p. 90-98, 2012.
- (9) PESSANHA, J.F.M., MELO, A C G, JUSTINO, T.C., MACEIRA, M.E.P. Combining Statistical Clustering techniques and Exploratory Data Analysis to Compute Typical Daily Load Profiles - Application to the Expansion and Operational Planning in Brazil. In: Probabilistic Methods Applied to Power Systems, Boise, 2018.
- (10) KOHONEN, T. Self-Organization and Associative Memory; Springer Berlin Heidelberg, vol. 8., 1989.
- (11) PESSANHA, J.F.M., VELASQUEZ, R.M., MELO, A.C.G., CALDAS, R.P. Técnicas de Cluster Analysis na Construção de Tipologias de Curva de Carga; XV Seminário Nacional de Distribuição de Energia Elétrica, Salvador, 2002.
- (12) WILLIAM, S.C. Robust locally weighted regression and smoothing scatterplots; Journal of the American statistical association, 74.368, PP. 829-836, 1979.
- (13) TAN, P.N., STEINBACH, M., KUMAR, V. Introdução ao Data Mining Mineração de Dados; Ciência Moderna, 2009.
- (14) DE GIORGI, M.G., CONGEDO, P.M., MALVONI, M., LAFORGIA, D. Error analysis of hybrid photovoltaic power forecasting models: A case study of mediterranean climate; Energy Conversion Management - Elsevier, vol. 100, pp. 117–130, 2015.
- (15) MALVONI, M., DE GIORGI, M.G., CONGEDO, P.M. Data on photovoltaic power forecasting models for Mediterranean climate; Data in Brief - Elsevier, vol. 7, no. May, pp 1639–1642, 2016.
- (16) PFENNINGER, S., STAFFELL, I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data,” Energy - Elsevier, vol. 114, pp. 1251–1265, 2016
- (17) WORLDWEATHERONLINE Historical or Past Weather API; [https:// developer .worldweatheronline .com/ api/ docs/historical-weather-api.aspx](https://developer.worldweatheronline.com/api/docs/historical-weather-api.aspx), 2019.

7.0 - DADOS BIOGRÁFICOS



Roberto Pereira Caldas (robertopcaldas@gmail.com) é Engenheiro de Eletrônica pelo Instituto Tecnológico de Aeronáutica (1978), Mestre (1990) e Doutorando em Engenharia Elétrica pela Universidade Federal do Rio de Janeiro (2018). No Centro de Pesquisas de Energia Elétrica - CEPEL, atuou como pesquisador no período de 1983 a 2017, ocupando ainda os cargos de Diretor de Pesquisa, Desenvolvimento e Inovação (2008 - 2016); Assistente da Diretoria de Pesquisa e Desenvolvimento (2001 - 2007); Coordenador do Programa de Pesquisa em Estudos Econômico-Financeiros (2000 - 2001); pesquisador da área de Medição e Conservação de Energia Elétrica (1983 - 2000) e engenheiro de ensaios (1982). Tem experiência na área de Engenharia Eletrônica aplicada a Energia Elétrica, em pesquisa e desenvolvimento voltados à industrialização de dispositivos, equipamentos e sistemas de medição de energia elétrica; e, mais recentemente, à introdução de redes elétricas inteligentes (*Smart Grid*) e de novas energias renováveis.



Albert C Geber de Melo (albert@cepel.br) é graduado pela UFPE (1983), mestre (1986) e doutor (1990) pela PUC-Rio, em Engenharia Elétrica. Pesquisador do CEPEL, onde atua desde 1985 na coordenação e no desenvolvimento de projetos nas áreas de confiabilidade; planejamento da expansão e da operação, incluindo novas renováveis; alocação eficiente de custos; tarifação; análise de riscos; análise de investimentos; análise de agrupamentos; comercialização; desenvolvimento sustentável; mudanças climáticas; política tecnológica; e política energética. Exerceu vários cargos gerenciais no CEPEL, incluindo Diretor de P&D (Jan 2005 – Jul 2008) e Diretor-Geral (Ago 2008 – Jan 2017), tendo representado esta instituição em diversos fóruns, como Comitê de Monitoramento do Setor Elétrico – CMSE e a Comissão Permanente de Análise de Metodologias e Programas Computacionais do Setor Elétrico – CPAMP/CNPE. Representou o Brasil em diversos fóruns internacionais, como na Agência Internacional de Energia – IEA, incluindo o *Renewable Energy Working Party* (REWP) e o *Committee on Energy Research Technology* (CERT), e foi Coordenador Técnico do Programa de Trabalho entre o Brasil e a IEA de 2011 a 2016. Participou ainda dos Grupos de Trabalho de Reforma do Setor Elétrico que resultou nos novos marcos regulatórios de 1997 e 2004. É Professor Adjunto da UERJ, *Distinguished Member* do CIGRÉ e *Senior Member* do IEEE, e membro titular da Academia Nacional de Engenharia.



José Francisco Moreira Pessanha (francisc@cepel.br) é Estatístico (ENCE, 1992) e engenheiro eletricista (UERJ, 1994). É mestre (COPPE/UFRJ, 1999) e doutor (PUC-Rio, 2006) em engenharia elétrica. Atualmente é professor adjunto da Universidade do Estado do Rio de Janeiro (UERJ) e pesquisador do Centro de Pesquisas de Energia Elétrica (CEPEL). No CEPEL desde 1995, tem participado de projetos de pesquisas e estudos e no desenvolvimento de modelos computacionais com aplicação de métodos estatísticos/probabilísticos e de otimização em sistemas de potência, em particular, nas áreas de confiabilidade de sistemas de potência, previsão de carga, previsão de longo prazo do mercado de energia elétrica, construção de tipologias de curva de carga, energia eólica, regulação técnico-econômica das distribuidoras de eletricidade e na tarifação do uso dos sistemas de distribuição com base em custos marginais. Em 2016 realizou pós doutorado no Inesc Tec Porto, Portugal, sobre previsão probabilística da geração eólica.



Djalma Mosqueira Falcão (falcao@nacad.ufrj.br) graduado em Engenharia Elétrica pela Universidade Federal do Paraná (1971), mestrado em Engenharia Elétrica pela COPPE / Universidade Federal do Rio de Janeiro (1973), doutorado em Engenharia Elétrica pela University of Manchester Institute of Science and Technology, Reino Unido (1981) e pós-doutorado pela University of California at Berkeley, USA (1993). Atualmente é professor titular da COPPE / Universidade Federal do Rio de Janeiro. De 2003 a 2004 atuou como Assessor do Presidente da Eletrobras. Em 2004 foi elevado ao grau de Fellow do Institute of Electrical and Electronics Engineers (IEEE). Em 2016 foi admitido como Membro Titular da Academia Nacional de Engenharia. Tem experiência na área de Engenharia Elétrica, com ênfase em Sistemas Elétricos de Potência, atuando principalmente nos seguintes temas: Sistemas de Potência, Inteligência Computacional, Processamento Paralelo e Redes Elétricas Inteligentes (*Smart Grid*).