

REGRESSÃO SIMBÓLICA APLICADA NA PREVISÃO DE IRRADIÂNCIA SOLAR INTRA-DIÁRIA NA CIDADE DE GOIÂNIA (BRASIL)

Gabriel Mendonça de Paiva – eegabrielmp@gmail.com

Sérgio Pires Pimentel – sergio_pimentel@ufg.br

Bernardo Pinheiro de Alvarenga – bernardo_alvarenga@ufg.br

Enes Gonçalves Marra – enes.gm@gmail.com

Universidade Federal de Goiás, Escola de Engenharia Elétrica Mecânica e de Computação

Resumo. Este trabalho apresenta uma implementação e validação de um modelo de previsão intra-diária de Irradiância Global Horizontal na cidade de Goiânia, Brasil. A previsão solar intra-diária é especialmente importante para subsidiar a operação de um setor elétrico que vislumbra uma maior participação de geradores intermitentes de energia solar fotovoltaica. Foram utilizados dois anos de registros de dados climáticos coletados na estação meteorológica da Escola de Engenharia Elétrica, Mecânica e de Computação da Universidade Federal de Goiás para a análise de um modelo de Regressão Simbólica (RS) em MATLAB para previsão de irradiância de 0 até 120 minutos a frente. Os resultados foram avaliados por meio do parâmetro Erro Médio Absoluto ou Mean Absolute Error (MAE) para cada intervalo futuro de previsão. A validação do modelo foi feita a partir de comparação dos resultados de MAE com os obtidos por 3 tipos de Persistência. O modelo de RS obteve desempenho similar à persistência do índice de claridade em uma previsão 15 minutos a frente, entretanto o mesmo se mostrou cada vez mais eficaz com o aumento do horizonte de previsão até 120 minutos futuros. Os resultados também foram comparados a dados apresentados em outras publicações.

Palavras-chave: Regressão Simbólica, Previsão de Curto Prazo, Previsão Intra-Diária.

1. INTRODUÇÃO

O desafio de se aumentar a participação de fontes renováveis de energia, dentre elas a solar, na matriz energética global é um dos mais importantes do século XXI. Isto por que o recurso solar é abundante e, teoricamente, capaz de suprir muitas vezes o próprio consumo de eletricidade mundial (Perez, 2015).

A barreira econômica para a disseminação dos sistemas fotovoltaicos tem sido cada vez mais reduzida, e hoje, estes sistemas já apresentam custo de energia produzida inferior ao carvão e à energia nuclear, por exemplo (Schmela, 2016). Entretanto, o aumento da penetração de sistemas fotovoltaicos conectados à rede (SFCR) nas matrizes de eletricidade convencionais resulta, também, no aumento de desafios técnicos. Dentre esses desafios, está a capacidade de operar e planejar o sistema elétrico baseando-se em fontes intermitentes de energia, como solar e eólica, inseridas também por meio de Geração Distribuída (GD).

Neste contexto, a capacidade de previsão futura de potência fotovoltaica gerada ou de irradiância tem sido objeto de estudo em várias partes do mundo. As previsões dos respectivos níveis de geração solar e eólica se tornaram fundamentais na implementação das *smart grids* (Inman *et al.*, 2013).

A Agência Internacional de Energia (IEA) classifica, por meio de um relatório, as previsões da disponibilidade do recurso solar quanto ao horizonte de previsão futuro em: previsão intra-diária, quando a previsão é feita de 0 até 6 horas à frente, ou previsão dia-a-frente, quando é desenvolvida a previsão do próximo dia (Pelland *et al.*, 2013). Normalmente técnicas baseadas em modelos estatísticos apresentam bom desempenho para previsões intra-diárias e técnicas baseadas em modelos físicos como *Numerical Weather Prediction (NWP)* apresentam o melhor desempenho na previsão dia-a-frente.

Alguns estudos que enfatizaram em previsões solares intra-diárias podem ser encontrados na literatura utilizando diferentes técnicas e metodologias. Rana *et al.* (2016) utilizaram Máquinas de Vetores de Suporte e NN Ensemble para realizar previsões solares de 5 até 60 minutos a frente na Austrália, com dados organizados em médias de 5 minutos. Eles avaliaram as duas técnicas e também avaliaram previsões com múltiplas variáveis e previsões baseadas apenas no histórico de potência gerada.

Pedro e Coimbra (2015) avaliaram também algumas técnicas para previsão de irradiância nos Estados Unidos de 15 até 120 minutos futuros em microclimas diversos e constatando que desempenhos diferentes são obtidos pelos métodos para cada microclima específico. Os dados foram organizados em médias de 15 minutos e foram testados por meio dos algoritmos de persistência, k-nearest-neighbors e redes neurais artificiais sem o uso de variáveis exógenas, ou seja, utilizando apenas o histórico de irradiância.

Russo *et al.* (2014) implementaram um modelo de previsão de Potência gerada na Itália para horizontes variando de 15 a 60 minutos a frente com uso de Programação Genética e histórico de potência e das variáveis climáticas como entradas.

A partir dessas motivações e, neste contexto, este trabalho propõe e analisa uma técnica de inteligência artificial para previsão de irradiância global horizontal (G) em Goiânia no contexto intra-diário, de 0 a 120 minutos a frente. É testada uma implementação de regressão simbólica multigene para a previsão de irradiância contendo múltiplas variáveis de entrada. As previsões intra-diárias são principalmente utilizadas para seguimentos de carga e despacho da geração convencional. Os resultados obtidos são discutidos e, também, comparados com aqueles obtidos por técnicas de persistência apresentadas por outros autores. Essa comparação baseia-se na quantificação de parâmetros de desempenho e foi adotada no intuito de quantificar o potencial de aplicação da metodologia proposta neste trabalho.

2. METODOLOGIA

2.1 Dados para Análise

Os dados utilizados neste trabalho correspondem a dois anos de registros de dados climáticos coletados pela Estação Meteorológica da Escola de Engenharia Elétrica, Mecânica e de Computação (EMC) da Universidade Federal de Goiás (UFG). A estação está localizada em uma latitude de $-16,67^\circ$ e uma longitude de $-49,24^\circ$. A estação registra os seguintes dados em intervalos de integração de 1 minuto: Irradiância Global Horizontal (G) em W/m^2 , Velocidade do Vento (W_s) em m/s, Direção do Vento (W_d) em Graus, Temperatura Ambiente (T_a) em $^\circ C$, Umidade Relativa (H_r) em %, Precipitação em mm e Pressão Atmosférica (p_a) em mBar. A Fig. 1 apresenta a Estação Solarimétrica da EMC / UFG e a curva de irradiância do período em análise: de Agosto de 2015 a Julho de 2017. O período de Agosto 2016 a Julho 2017 foi utilizado como treinamento do algoritmo e o período de Agosto de 2015 a Julho de 2016 foi utilizado como período de teste. Esta escolha foi feita pois verificou-se uma pequena perda de dados no primeiro ano de operação da estação (8338 dados de um total de 527040), desta forma o treinamento foi realizado no ano que não teve perda de dados. Entretanto, é importante ressaltar que os dados coletados durante o primeiro ano de operação poderiam ser utilizados no treinamento do algoritmo desde que, previamente, os registros do período desconhecido fossem estimados a partir de um método de interpolação. Tal condição seria aceitável uma vez que o percentual de dados desconhecidos é relativamente pequeno.

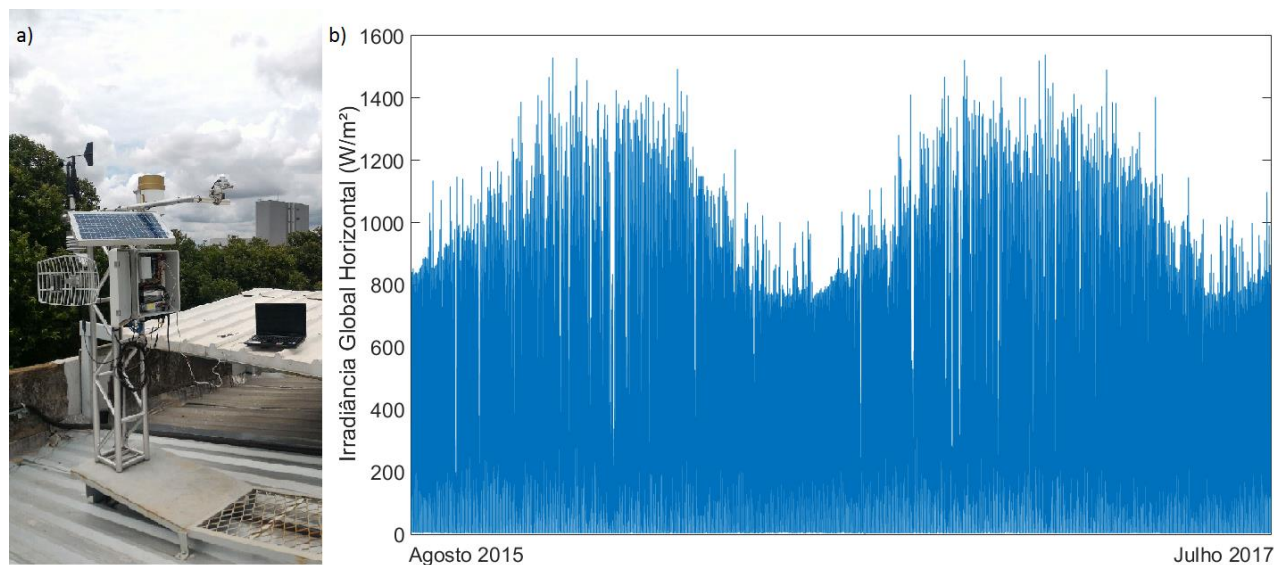


Figura 1 – a) Estação Meteorológica da EMC / UFG e b) curva de irradiância do período em análise.

2.2 Regressão Simbólica

A Programação Genética (PG) é uma técnica de Inteligência Artificial (IA) que pertence ao campo da Computação Evolucionária e é na prática uma extensão do Algoritmo Genético (AG) inicialmente proposta por Koza (1992). A PG é inspirada na genética e na evolução em nível populacional. Ela usa operadores evolucionários como *crossover* e mutação em uma população inicial para aprimorar o encaixe entre a população de programas candidatos e uma função objetivo, de maximização ou minimização (Brownlee, 2011).

Quando a PG é utilizada para construir um modelo matemático a partir de dados adquiridos por um processo ou sistema para prever variáveis ela é chamada Regressão Simbólica (RS).

Uma primeira vantagem do uso da RS é que não é necessário um conhecimento profundo sobre o sistema em análise para se desenvolver modelos precisos, uma vez que a técnica desenvolve tanto a estrutura do modelo de regressão quanto os parâmetros ou coeficientes do modelo matemático em uma estrutura de árvores.

Os modelos de RS podem ser classificados em três tipos quanto à complexidade do modelo matemático: *naive Symbolic Regression*: quando o modelo utiliza apenas um gene (árvore) para relacionar dados de entrada e saída; *scaled*

Symbolic Regression: quando o modelo utiliza um termo de *offset* (constante) adicionado a um gene para relacionar entrada e saída; e *Multigene Symbolic Regression*: quando o modelo utiliza um termo de *offset* adicionado a mais de um gene para relacionar entrada e saída, como descrito pela Eq. (1) (Searson *et al.*, 2010).

$$\hat{y} = d_0 + d_1 \left[\frac{41}{100} x_1 + \tanh(x_2 x_3) \right] + d_2 \left[\frac{45}{100} x_3 + \sqrt{x_2} \right] \quad (1)$$

A Fig. 2 apresenta um exemplo de representação em árvores de uma RS Multigenética equivalente à Eq. (1). A RS desenvolve tanto a estrutura do modelo, no que se refere ao número de genes e funções (nós), quanto os seus parâmetros (constantes d_0 , d_1 , d_2 , $\frac{41}{100}$ e $\frac{45}{100}$).

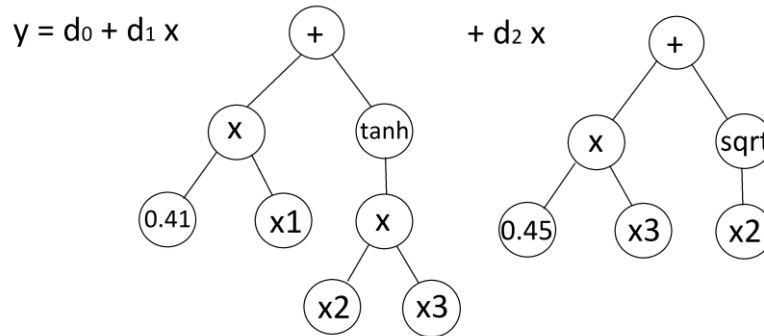


Figura 2 - Representação em árvores do modelo de regressão da Eq. (1).

Uma outra vantagem da RS é que os modelos de regressão se tratam de simples equações intuitivas, que podem ser implementadas *a posteriori* em qualquer linguagem de programação ou dispositivo. Diferentemente de outras técnicas como Redes Neurais Artificiais (RNA) *feed forward* ou Máquinas de Vetores de Suporte (MVS), não são necessários ambientes de software específicos para a aplicação de modelos treinados de RS.

A Fig. 3 apresenta de forma esquemática a relação entre as variáveis para construção dos modelos de RS em MATLAB. Os dados foram reorganizados de períodos de integração de 1 minuto para 15 minutos por meio de média aritmética simples. Foi construído um modelo de regressão para cada intervalo futuro de previsão, variando de 15 minutos à frente até 120 minutos à frente. São utilizadas 62 variáveis de entrada para cada modelo, que são: os registros de 3 horas anteriores de pressão atmosférica (p_a), velocidade do vento (W_s), umidade relativa do ar (H_r), temperatura ambiente (T_a) e irradiância (G) e também uma variável que representa o dia do ano (Dia) e uma que representa o horário do dia (t_s). A contagem de dia do ano se inicia no solstício de inverno (21 de Junho) e finaliza no dia anterior ao solstício de inverno do outro ano. A variável t_s contabiliza o tempo em minutos após o nascer do Sol em cada dia.

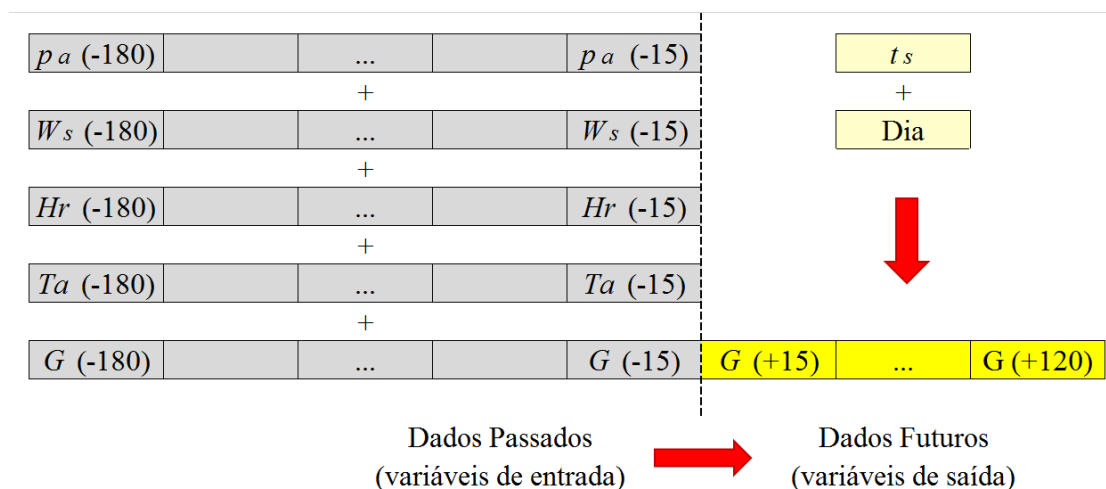


Figura 3 - Representação esquemática das variáveis utilizadas para construção dos modelos de RS.

2.3 Persistência da irradiância

A persistência é o método de previsão mais simples de ser realizado e é normalmente utilizado como base de referência para validação de modelos de previsão mais sofisticados. Serão consideradas três formas distintas de persistência neste estudo, todas utilizadas na literatura. No primeiro caso, uma forma de se prever a irradiância futura é com base nos últimos registros, ou seja, o valor futuro da irradiância sendo igual ao último valor medido, que se resume pela Eq. (2), sendo $G(t+\Delta T)$ o valor futuro da irradiância, $G(t)$ o valor presente e ΔT o horizonte de previsão (de 15 a 120 minutos futuros).

$$G(t + \Delta T) = G(t) \quad (2)$$

2.4 Persistência do índice de claridade

A segunda forma de persistência considerada é algumas vezes denominada na literatura como *smart persistence*, que se trata de considerar que o valor futuro do índice de claridade k_t , dado pela Eq. (3), será igual ao valor atual. Nesta equação, G se trata da irradiância global horizontal e G_o a irradiância extraterrestre. A irradiância extraterrestre foi determinada pela Eq. (4), como descrito por Duffie e Beckman (1991), sendo G_{sc} a constante solar, dada por 1367 W/m², n o dia do ano e as demais variáveis são ângulos que representam a geometria solar, detalhados em (Duffie e Beckman, 1991). A irradiância extraterrestre foi calculada em períodos de 1 minuto e agrupada em equivalentes de 15 minutos por meio de média aritmética.

$$k_t = \frac{G}{G_o} \quad (3)$$

$$G_o = G_{sc} \left(1 + 0.033 \cos \frac{360n}{365} \right) (\cos \phi \cdot \cos \delta \cdot \cos \omega + \sin \phi \cdot \sin \delta) \quad (4)$$

Esta previsão é chamada *smart persistence* devido ao fato de ser uma forma mais inteligente de se prever o valor futuro e que resulta em menores erros. Neste caso o índice de claridade representa o fator de atenuação da irradiância do plano extraterrestre para o plano terrestre pela atmosfera. A previsão *smart persistence* é resumida pela Eq. (5).

$$G(t + \Delta T) = k_t(t) \cdot G_o(t + \Delta T) \quad (5)$$

A terceira forma de persistência considerada se trata de igualar a curva de irradiância do dia presente ao dia anterior, ou seja, a irradiância em cada horário do presente se equivale à irradiância nos respectivos horários no dia anterior.

3. RESULTADOS E DISCUSSÕES

3.1 Análise de desempenho dos modelos de previsão

O erro de previsão foi medido por meio do Erro Médio Absoluto ou *Mean Absolute Error (MAE)*, dado pela Eq. (6), onde \hat{y}_i é o valor previsto e y_i é o valor real. O *MAE* é uma variável frequentemente utilizada na área de previsão solar por ser intuitiva. Foi considerado neste estudo também o *Mean Relative Error (MRE)*, dado pela Eq. (7), que se trata de uma extensão do *MAE* e que pode ser utilizado para comparação com outras localidades, uma vez que o erro é normalizado com relação à faixa de valores que a irradiância G alcança. Para a base de dados considerada, na localidade de Goiânia e em valores médios de 15 minutos, esta faixa R vai de 0 a 1289,91 W/m² no período de teste escolhido.

$$MAE = \frac{\sum_{i=1}^n |(y_i - \hat{y}_i)|}{n} \quad (6)$$

$$MRE = \frac{\sum_{i=1}^n |(y_i - \hat{y}_i)|}{n \cdot R} \cdot 100\% \quad (7)$$

A Fig. 4 apresenta as curvas de *MAE* variando-se o horizonte de previsão para cada um dos modelos de previsão considerados e a Tab. 1 apresenta os valores específicos de *MAE* e *MRE* obtidos por cada modelo nos diferentes horizontes de previsão. É possível observar que a persistência pelo índice de claridade é competitivamente eficaz em

uma previsão muito próxima, uma vez que, normalmente, a condição de atenuação da Irradiância pela atmosfera em um futuro próximo tende a ser semelhante. Entretanto, aumentando-se o horizonte de previsão, os modelos de Persistência resultam em erros maiores do que modelos estatísticos mais robustos como a RS, que a partir de 30 minutos a frente obteve a melhor eficiência de previsão. Os valores de *MRE* representam um desvio médio percentual em relação ao pico de irradiância registrado variando de 4,64% para o horizonte de 15 minutos a frente a 8,61% para o horizonte 120 minutos a frente.

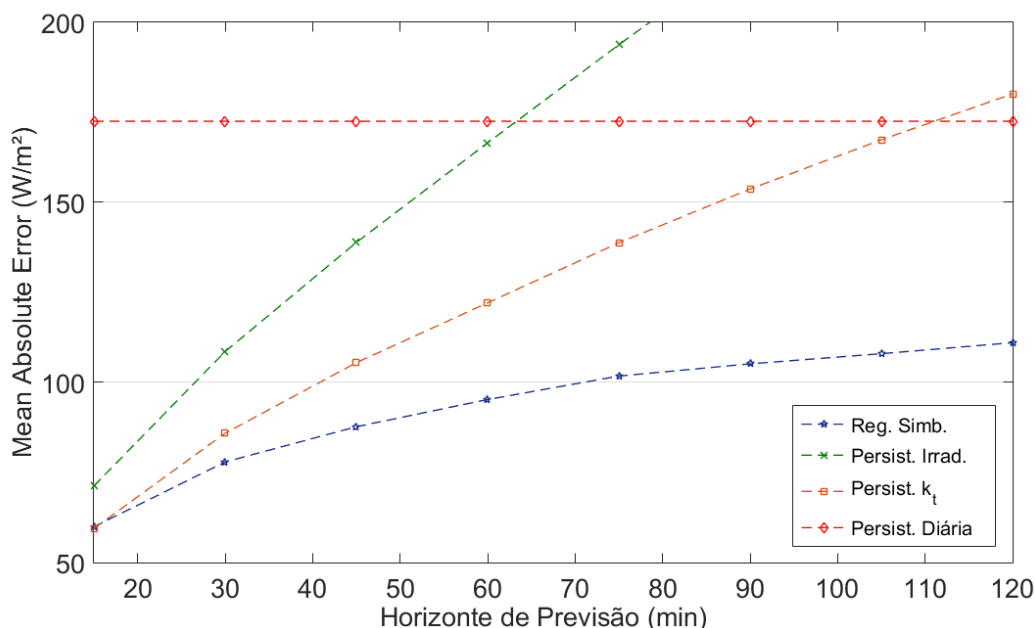


Figura 4 – Curvas de MAE em função do horizonte futuro de previsão para os modelos considerados.

Tabela 1 – Valores de MAE e MRE obtidos pelos modelos em análise.

Horiz. de Previsão (min)	Regressão Simbólica		Persistência Irrad.		Persistência k_t		Persist. dia a frente	
	MAE	MRE	MAE	MRE	MAE	MRE	MAE	MRE
<i>G</i> (+15)	59,91	4,64	71,25	5,52	59,43	4,61	172,53	13,38
<i>G</i> (+30)	77,92	6,04	108,57	8,42	85,91	6,66	172,53	13,38
<i>G</i> (+45)	87,71	6,80	138,88	10,77	105,49	8,18	172,53	13,38
<i>G</i> (+60)	95,26	7,39	166,54	12,91	122,12	9,47	172,53	13,38
<i>G</i> (+75)	101,78	7,89	193,79	15,02	138,79	10,76	172,53	13,38
<i>G</i> (+90)	105,17	8,15	220,20	17,07	153,72	11,92	172,53	13,38
<i>G</i> (+105)	107,99	8,37	245,69	19,05	167,38	12,98	172,53	13,38
<i>G</i> (+120)	111,05	8,61	270,43	20,96	180,13	13,96	172,53	13,38

A Fig. 5 apresenta as curvas de irradiância previstas e medidas para diferentes horizontes de previsão: 30, 60, 90 e 120 minutos a frente. Nela são exibidos os comportamentos dos modelos para um dia de céu limpo (acima), um dia de céu parcialmente nublado (meio) e um dia nublado (abaixo). É possível constatar dois comportamentos: primeiramente, quanto maior o horizonte de previsão, maior o distanciamento entre previsão e medição. Em seguida é possível observar que mesmo em horizontes mais distantes de previsão, os modelos são capazes de representar a característica genérica de um dia de céu limpo, onde os picos de Irradiância são maiores, ou de um dia nublado, onde os picos de Irradiância são menores.

É possível observar também que as previsões apresentam alto desvio em relação aos valores reais com o aumento do horizonte de previsão, entretanto os resultados ainda são melhores que os obtidos pelos modelos de Persistência, conforme é apresentado na Tab. 1.

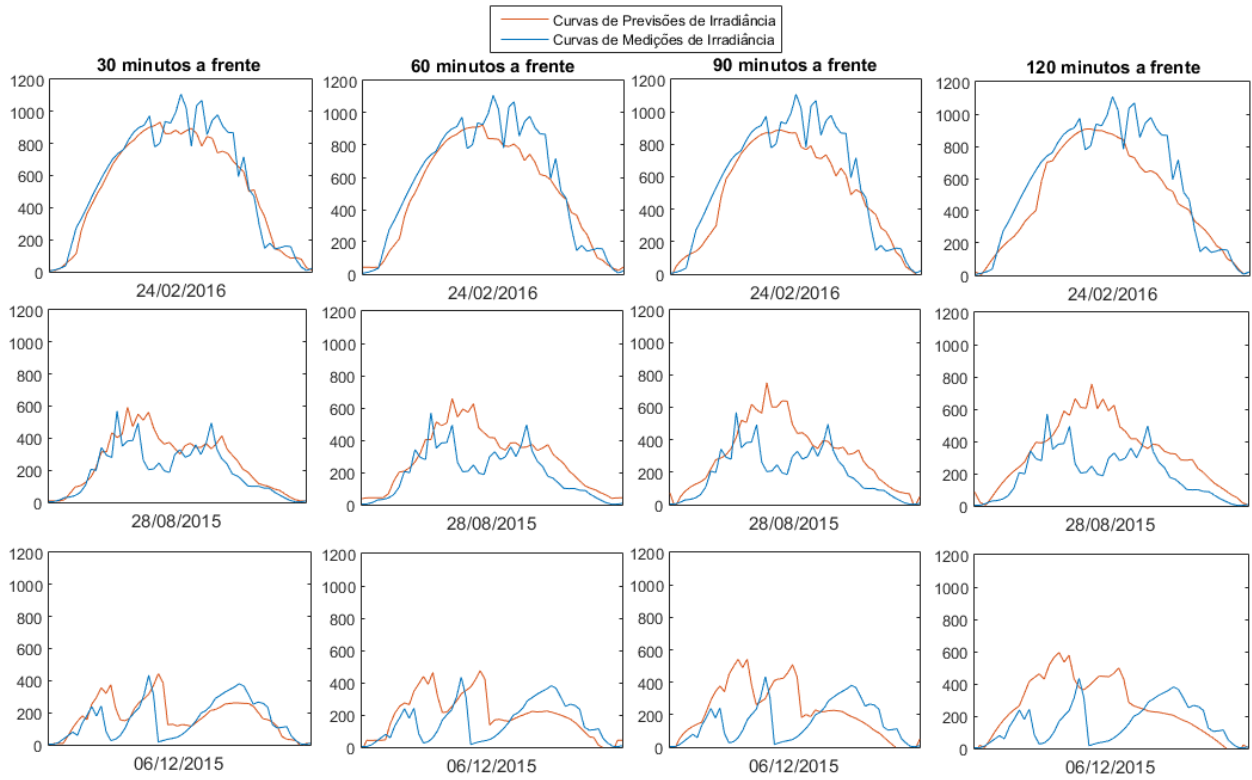


Figura 5 – Curvas de Previsões e Medições para os horizontes 30, 60, 90 e 120 minutos a frente, em um dia de céu limpo (acima), um dia parcialmente nublado (meio) e um dia nublado (abaixo).

3.2 Complexidade dos modelos de regressão simbólica

Os modelos de RS são equações intuitivas (indivíduos) que em uma população obtiveram o melhor desempenho relativo a uma função objetiva que minimiza o erro, neste caso o *RMSE* ou *Root Mean Squared Error*. Os modelos dados pela Eq. (8) e Eq. (9) são exemplos dos modelos de regressão obtidos, neste caso, para a previsão da irradiância 30 minutos e 90 minutos futuros, respectivamente.

$$\begin{aligned} \hat{G}_{(+30)} = & -33,3 + 0,268 \left[G_{(-45)} + G_{(-90)} \right] - 4,16 \cdot 10^{-4} t_s \left[G_{(-45)} + G_{(-90)} \right] + 23,4 \sqrt{G_{(-15)}} \\ & + 5,27 \cdot 10^{-5} G_{(-60)} \left[G_{(-30)} + G_{(-75)} \right] + 1,29 \cdot 10^{-5} \left[G_{(-135)}^2 - t_s^2 \log(G_{(-15)})^2 \right] \end{aligned} \quad (8)$$

$$\begin{aligned} \hat{G}_{(+90)} = & -289 + 0,296 G_{(-60)} + 87 \log(t_s) - 0,1314 \sqrt{G_{(-45)}} (t_s - T_{a(-30)}) \\ & - 0,0672 \sqrt{G_{(-15)}} (t_s - T_{a(-30)} - W_{s(-135)} + \sqrt{Dia}) + 7,63 \sqrt{G_{(-15)}} \log(t_s) \end{aligned} \quad (9)$$

Na análise de uma população de modelos de RS é possível se optar por modelos de menor desempenho com menor complexidade por meio da análise da frente de pareto, que indica os modelos de maior eficiência. A Fig. 6 apresenta os gráficos das populações de indivíduos para a RS a) 30 minutos e b) 90 minutos a frente. Sendo R^2 o coeficiente de correlação entre o modelo de previsão e os dados reais, dado pela Eq. (10), e m_y a média dos valores observados de y , no caso a irradiância G . A frente de pareto se trata dos indivíduos em verde, ou seja, as soluções com melhor ajuste para os diferentes níveis de complexidade. Modelos de menor complexidade representam modelos com menores custos computacionais em sua futura implementação.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - m_y)^2} \quad (10)$$

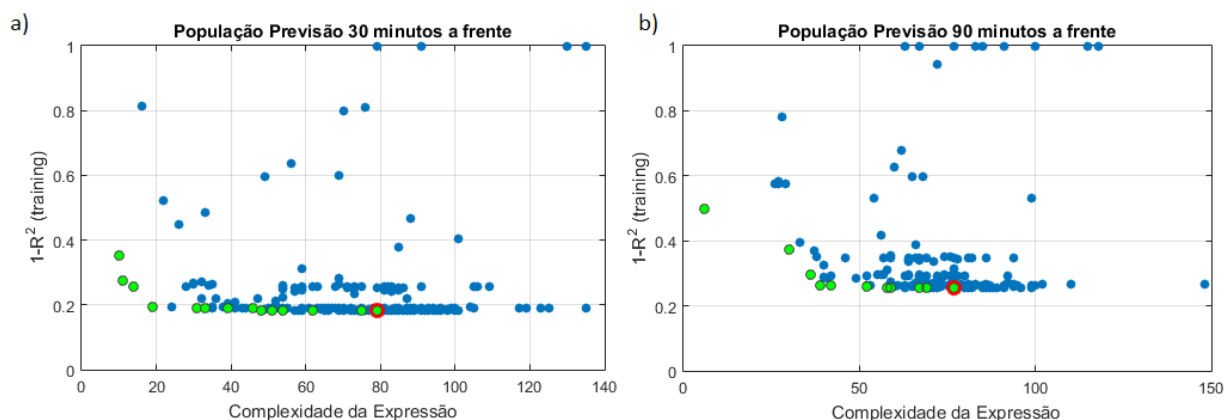


Figura 6 – População de possíveis soluções e frente de pareto para a previsão a) 30 minutos e b) 90 minutos a frente.

Como exemplo, no caso da RS para 30 minutos a frente dado pela Fig. 6 a), o modelo da Eq. (11) apresenta uma redução de 64,8% na complexidade do modelo com uma redução de 1,1% no coeficiente de correlação na fase de treinamento. No caso da RS para 90 minutos futuros dado pela Fig. 6 b), o modelo da Eq. (12) apresenta uma redução de 43,5% na complexidade com uma redução de 1,07% no coeficiente de correlação entre as curvas de previsão e medição, também na fase de treinamento.

$$\hat{G}_{(+30)} = 0,506(G_{(-45)} + G_{(-90)}) - 8,63 \cdot 10^{-4} t_s (G_{(-45)} + G_{(-90)}) + 19,6\sqrt{G_{(-15)}} - 42,7 \quad (11)$$

$$\hat{G}_{(+90)} = 0,137G_{(-60)} + 9,35\sqrt{G_{(-15)}} \log t_s - 0,0894\sqrt{G_{(-15)}}(t_s - T_{a(-30)} + \sqrt{Dia}) + 44,7 \quad (12)$$

3.3 Comparação com publicações recentes

A Tab. 2 apresenta uma comparação direta de valores de MAE obtidos pelo presente estudo e outros presentes na literatura. É importante destacar que são estudos em diferentes localidades e também diferentes variáveis previstas e sistemas fotovoltaicos distintos. Os trabalhos de Rana *et al.* (2016) e Russo *et al.* (2014) são análises de previsão de potência gerada, enquanto que este trabalho e o de Pedro e Coimbra (2015) são de previsão de Irradiância. Além disso, o trabalho de Rana *et al.* (2016) desenvolveu um modelo de previsão para uma base de dados em intervalos médios de 5 minutos, o que eleva a intermitência da base de dados, já os demais são organizados em 15 minutos. Outro fator que provavelmente aumenta o erro de previsão é a localidade e a consequente faixa de valores possíveis: em Goiânia (latitude -16,7°) e em Brisbane (-27,3°) tratam-se de localidades no hemisfério Sul, fator que aumenta a faixa de valores e os níveis de irradiância em geral observados. A faixa de potência do estudo na Austrália vai até 1150 W e no presente estudo até 1289,91 W/m². É possível observar que o presente estudo apresenta resultados satisfatórios se comparado principalmente ao estudo de Rana *et al.* (2016) que se trata de uma localidade, *a priori*, mais semelhante com a analisada aqui.

Tabela 2 – Valores de MAE obtidos pelo presente estudo e outros presentes na literatura.

Horiz. de Previsão (min)	Método Proposto	Rana et al.		Pedro e Coimbra				
	Goiânia	Brisbane	Catania	Merced	Davis	San Diego	Bellingham	Ewa Beach
G (+15)	59,91	67,11	25,46	17,60	17,00	27,90	26,10	69,70
G (+30)	77,92	86,42	35,94	19,40	18,90	33,60	30,00	66,70
G (+45)	87,71	99,62	44,09	20,40	20,00	37,80	32,20	64,70
G (+60)	95,26	110,27	47,13	21,50	21,00	41,10	33,50	63,90
G (+75)	101,78	-	-	-	-	-	-	-
G (+90)	105,17	-	-	22,80	23,10	46,70	37,50	64,10
G (+105)	107,99	-	-	-	-	-	-	-
G (+120)	111,05	-	-	24,10	24,30	52,80	38,60	63,20

4. CONCLUSÕES

É possível concluir por meio da análise do modelo proposto de previsão intra-diária que a técnica e o modelo proposto são uma potencial solução para problemas de previsão solar. Isto porque o modelo proposto se sobressaiu na comparação com as bases de referências de persistência. O modelo também foi avaliado em comparação com outros trabalhos e apresentou resultados semelhantes a outro em uma localidade também no hemisfério Sul.

A RS também apresenta importantes vantagens que potencializam sua aplicabilidade: o fato de o modelo de regressão se tratar de simples equações intuitivas compreensíveis por seres humanos e o fato de que estes modelos podem ser implementados, *a posteriori*, em qualquer dispositivo ou linguagem de programação. Os modelos propostos são ainda independentes do tempo, ou seja, uma vez treinados os modelos de RS, não é necessário processamento futuro do modelo de previsão.

O modelo proposto de previsão pontual de irradiância pode ser avaliado para localidades próximas e em previsões regionais para verificar a potencial aplicabilidade dos modelos em casos reais e no contexto de *microgrids*, onde sistemas fotovoltaicos dispersos em pequenas regiões podem ser controlados em nível regional.

Agradecimentos

Os autores agradecem à Universidade Federal de Goiás (UFG) por meio da Escola de Engenharia Elétrica Mecânica e de Computação (EMC) pelos dados da estação obtidos através do portal da estação meteorológica da EMC/UFG (Pimentel, 2017).

REFERÊNCIAS

- Brownlee, J., 2011. Nature-Inspired Programming Recipes.
- Duffie, J. A., Beckman, W. A., 1991. Solar Engineering of Thermal Processes, John Wiley & Sons.
- 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. 42 p.
- Koza, J. R., 1992. Genetic Programming: On the Programming of Computers by Means of Natural Selection, MIT Press, Cambridge, MA, USA.
- Pelland, S., Remund, J., Kleissl, J., Oozeki, T., De Brabandere, K., 2013. Photovoltaic and Solar Forecasting: State of the Art. International Energy Agency. 40 p.
- Rana, M., Koprinska, I., Agelidis, V. G., 2016. Univariate and multivariate methods for very short-term solar photovoltaic power forecasting. Energy Conversion and Management. 11 p.
- Pedro, H. T. C., Coimbra, C. F. M., 2015. Short-term irradiance forecastability for various solar micro-climates. Solar Energy. 16 p.
- Russo, M., Leotta, G., Pugliatti, P. M., Gigliucci, G., 2014. Genetic Programming for photovoltaic plant output forecasting. Solar Energy. 10 p.
- Perez, R., 2015. A fundamental look at supply side energy reserves for the planet. International Energy Agency. URL: <http://asrc.albany.edu/people/faculty/perez/2015/IEA.pdf>
- Pimentel, S. P., 2017. Weather Station – Sistema Fotovoltaico EMC/UFG. URL (acessada em 19/11/2017): <https://sites.google.com/site/sfvemcufg/weather-station>
- Schmela, M., 2016. Global Market Outlook For Solar Power / 2017 – 2012. Solar Power Europe. 60 p.
- Searson, D. P., Leahy, D. E., Willis, M. J., 2010. Gptips: an open source genetic programming toolbox for multigene symbolic regression. Proceedings of the International multiconference of engineers and computer scientists. 4 p.

SYMBOLIC REGRESSION APPLIED TO INTRA-DAY IRRADIANCE FORECASTING IN THE CITY OF GOIANIA (BRAZIL)

Resumo. *This work presents an implementation and validation of an intra-day forecasting model to predict global horizontal irradiance in the city of Goiania, Brazil. Intra-day solar forecasting is important specially to support the operation of an electric sector that aims to grow the connection of intermittent photovoltaic generators to the conventional electric grid. Two years of registers of climate data were used in this study. They were collected from the meteorologic station of the Electrical Mechanic and Computer Engineering School (EMC) from Federal University of Goiás (UFG). The database was used to analyze Symbolic Regression (SR) models applied to irradiance forecasts from 0 to 120 minutes ahead. The results were evaluated using the Mean Absolute Error (MAE) parameter for each future interval of forecast. The validation was done by comparison of MAE values from 3 types of persistence forecasting. The SR models obtained similar performance to the clearness index persistence 15 minutes ahead, however the SR models became more effective with the increase in forecast horizon up to 120 minutes ahead. The results were also compared to data presented in other publications.*

Key words: *Symbolic Regression, Short-term forecasting, Intra-day forecasting.*