

Modelling and Forecasting for Solar Irradiance from Solarimetric Station

Bruna Kariny Fontes Rodrigues, Matheus Pereira Gomes, Ângelo Márcio Oliveira Santanna, Daniel Barbosa *Senior IEEE*, and Luciana Martinez

Abstract—This paper proposes two approaches for modeling solar irradiance time series. The first is an exploratory analysis and the second is a periodic forecast model for solar irradiance data from solarimetric station in a country city of Bahia, Brazil.

Two normality hypothesis tests were applied, Anderson-Darling and Shapiro-Wilk. The shape and symmetry of the data were also analyzed using boxplot and histograms to investigate the extreme points, as well as great asymmetry in the distribution of data at different times. Through adherence tests, the normal distribution of solar irradiance in most hours within the annual periodic is rejected. A study by season was carried out, which showed a different behavior in relation to the symmetry of the data, with autumn being the most uniform and spring being the most stochastic. Spring and summer presented better conditions for installation of photovoltaic plates, due to the high solar irradiance rates, and autumn and winter presented satisfactory solar irradiance to maintain this form of generation throughout the year. The second approach is a forecast to solar irradiance in photovoltaic generation systems. The importance of this forecast type is to favor the routine of planning, operation and maintenance of these types of systems, in addition to serving as a basis for feasibility studies and expansion of solar generation. Given a stochastic resource and the periodic behavior of solar irradiance data, a periodic autoregressive model was considered; the statistics were used by the maximum likelihood method, based on hourly measurements of irradiance over a period of one year.

Index Terms—Photovoltaic systems; Solar Energy; Solar irradiance; Statistical Analysis; Time Series.

I. INTRODUÇÃO

O Brasil apresenta um abundante potencial de geração de energia renovável. Grande parte dessa capacidade está relacionada à sua dimensão continental e a sua localização geográfica estratégica em relação aos paralelos do globo terrestre. Devido a grande parte da região brasileira estar situada em uma latitude próxima a linha do Equador, não há variação excessiva nos níveis de temperatura, que permanecem altas ao longo de todo o ano [1]. Especialmente a porção que se encontra entre a linha do Equador e o Trópico de Capricórnio, que atravessa os estados de São Paulo, Mato Grosso do Sul e Paraná e marca o fim da zona tropical do país.

As características geográficas e climatológicas conferem ao Brasil um potencial de geração de energia renovável expressivo, sendo o potencial hidrelétrico com capacidade de até 260

GWh [2], abundância em ventos, com cerca de duas vezes maior do que a média mundial [3] e privilegiado em termos de incidência da radiação solar, cujas médias diárias variam de 8 a 22 MJ/m^2 [4].

Para suplantiar possíveis crises de suprimento de energia causadas pelo aumento da demanda ou as mudanças climáticas nos regimes hidrológicos, é habitual recorrer a alternativas à hidrelétrica. A principal destas são as termoeletricas, que são em sua maioria de origem não-renovável, encarecem o custo de produção e estão sujeitas à intervenção da economia internacional, uma vez que o preço do petróleo varia de acordo com mercado internacional [5].

Por conta de importantes fatores como a necessidade de aliar a preservação ambiental e as mudanças climáticas [6] aos assuntos de caráter técnico e econômico [7], o processo de geração de energia elétrica diversificou-se ao longo dos anos. As fontes de geração mais utilizadas no Sistema Elétrico Brasileiro, de natureza hidráulica e térmica, passaram a compartilhar esse cenário com outros tipos de fontes de geração, provenientes de diferentes tipos de recursos presentes na matriz energética brasileira [8], [9].

No Brasil, há destaque para a matriz fotovoltaica e eólica, que tem sido impulsionadas pelo crescente número de regulamentações promovidas pela Agência Nacional de Energia Elétrica (ANEEL), com destaque para as Resoluções Normativas n° 391/2009, n° 482/2012 e n° 687/2015.

Na investigação sobre como a energia solar pode promover a geração de energia elétrica de um parâmetro conhecido como irradiância solar é crucial para esse entendimento. A irradiância solar corresponde ao nível de radiação solar capturado na superfície horizontal. Existem diversas metodologias de estudo (estimação, caracterização e previsão) da irradiância solar, sendo os mais comuns os modelos analíticos, estocásticos, empíricos, estatísticos e as redes neurais [10], [11]. Os modelos analíticos são baseados no conhecimento das leis físicas que regem o fenômeno, como a interação entre radiação solar e a atmosfera terrestre [12]. Esses modelos costumam ser usados para reconstruir a irradiância solar a fim de utilizar como parâmetro de entrada em modelos climáticos [13].

Os modelos empíricos estimam a radiação solar a partir de outros parâmetros climáticos, como duração da luz solar, temperatura e nebulosidade [14]. Esses modelos são mais utilizados em locais onde não há a possibilidade da medição direta da irradiância, sendo [15] o pioneiro desse método.

Outra abordagem é obtida por meio dos modelos estocásticos, a exemplo dos modelos com componentes auto-

Bruna Kariny Fontes Rodrigues, Matheus Pereira Gomes, Daniel Barbosa, Luciana Martinez, Departamento de Engenharia Elétrica e de Computação, Universidade Federal da Bahia, BA, (e-mail: bkfrdrigues@gmail.com, elemat.eng@gmail.com, dbarbosa@ufba.br, lucianam@ufba.br).

Ângelo Márcio Oliveira Santanna, Departamento de Engenharia Mecânica, Universidade Federal da Bahia, BA, (e-mail: angelo.santanna@ufba.br)

regressivos (AR) e média-móvel (MA, do inglês *Moving Average*), que presumem processos lineares [16] e têm como objetivo a previsão de pontos futuros [17], [18]. A não-linearidade dos dados e fenômenos dinâmicos pode afetar a precisão desses modelos. Ainda no contexto de previsão, o uso das redes neurais artificiais é crescente no estudo da irradiância solar. Esse modelo usa técnicas de *Machine Learning* para obter as previsões [19]. Ambas técnicas são amplamente utilizadas, seja de forma única, como feito por [20] (ARIMA), [21] (Redes Neurais), ou em conjunto com operadores lineares, a exemplo das transformadas, para filtrar variações não lineares [22]. Observa-se nestes estudos citados que os modelos não capturam a característica estocástica da irradiância solar e a alta variabilidade das medições.

A análise estatística utiliza banco de dados de medição de irradiância (ou estimações baseadas em modelos empíricos) e buscam caracterizar o fenômeno através de estatísticas descritivas e da frequência de distribuição, como feito por [23] e que também será adotado neste artigo.

Além disso, grande parte dos recursos da matriz energética utilizados para geração de energia elétrica possuem características sazonais e/ou periódicas, associadas a elementos e fenômenos da natureza [9].

Além dos inúmeros modelos de análise de irradiância, há diferentes intervalos de tempo em que as coletas dos dados de irradiância são executados, as quais estão em conformidade com os tipos de sazonalidade apresentado por esse fenômeno. As mais utilizadas são a média mensal dos dados diários de irradiação, como usado por [12], a média diária [22] e a média por hora [16]. Os pontos da média por hora tradicionalmente são organizados sequencialmente, com todas as horas do dia dentro da mesma série. Nesse tipo de série, há uma grande variância, uma vez que a noite a série atinge medições nulas e há picos de irradiância que chegam a mais de 1.000 W/m^2 . A nova abordagem proposta nesta pesquisa é minimizar a variabilidade das medições diárias separando os dados em 24 séries horárias correspondentes às 24 horas no dia.

Inicialmente, este trabalho implementa uma abordagem de estatística exploratória que analisa as medições de irradiância (ou estimações baseadas em modelos empíricos) e caracteriza o fenômeno estocástico da irradiância solar baseado na distribuição de probabilidade da frequência horária. Em seguida, aplicou-se a abordagem de análise periódica aos modelos de séries temporais para modelar as medições históricas e realizar a previsão da irradiância solar em sistemas de geração fotovoltaica. Com base no aspecto periódico das medidas de irradiância solar, o modelo Periódico Auto-Regressivo (PAR) é proposto como contribuição deste trabalho para a previsão de dados horários, sendo considerada uma informação de grande valia para os processos de planejamento, operação e manutenção de sistemas solares, em particular se esses dados de previsão forem utilizados para a obtenção das cartas de controle. Os resultados desse artigo serão obtidos por meio de ferramentas computacionais, como o *Mathworks Matlab R2018b* e a linguagem R, com auxílio das bibliotecas *tidyverse* e *nortest*

II. MATERIAIS E MÉTODOS

A. Características das Medições

As análises apresentadas nesse trabalho contaram com um banco de dados que contém medições de irradiância solar efetuadas entre os anos de 2013 e 2014. Essas medidas foram efetuadas por uma estação solarimétrica, situada no interior da Bahia, que se encontra na zona climática tropical, ou ainda usando a classificação de [24], na Zona Climática Tórrida, o que lhe confere uma alta incidência de radiação solar durante todo o ano.

Os dados originais correspondem a medições de irradiância em intervalos de 10 min (dez minutos) durante o período anual (365 dias). A preparação mais adequada dos dados de forma mais concisa consistiu em efetuar a média a cada seis medidas de irradiância solar, obtendo as medidas de irradiância em intervalos de 1h (uma hora). Assim, obteve-se um banco de dados de 24 estratos, referentes as horas, onde cada um contém 365 registros de irradiância solar.

Essa abordagem permite diminuir a variabilidade diária de amplitudes inerente do fenômeno, que começa com baixa irradiância no amanhecer, cresce até um pico (por volta de 1.000 W/m^2) e decresce ao anoitecer. Ao particionar a série por hora, deixa-se de levar em conta esse processo e passa a ser analisado apenas as diferenças de amplitudes causados pelas estações do ano e por fenômenos climáticos.

É importante ressaltar a diferença entre as abordagens, uma vez que a escolha do intervalo de dados e a forma de considerá-los permite visualizar ou ignorar variação diária, sazonalidade e tendências não-sazonais. Por exemplo, [23] analisa a média diária dentro de cada mês ao longo de três anos, o que permite visualizar as amplitudes sem a interferência das estações do ano.

B. Análise Estatística

Foram utilizados estatísticas descritivas e inferenciais para caracterizar os dados. As medidas estatísticas utilizadas foram mínimo, máxima, média, mediana e o desvio padrão da irradiância solar a cada hora, ao longo de todo o intervalo de medições.

A verificação do ajuste dos dados com a distribuição normal é uma importante informação, uma vez que é recorrente em fenômenos naturais e muitas análises estatísticas sustentam-se nesse pressuposto. Assim, foi realizado os testes de aderência a normalidade de Anderson-Darling e o de Shapiro-Wilk, que são testes paramétricos. Os testes presumem a normalidade dos dados como a hipótese H_0 , e um $p - \text{valor} > 0,05$ caso os resultados confirmem a normalidade para os dados amostrais.

C. Previsão de Séries Temporais

1) *Análise de séries temporais*: Uma série temporal consiste em uma variável aleatória indexada no tempo, denotada por $\{x_t, t \in \mathcal{R}_+\}$, sendo \mathcal{R}_+ o conjunto dos números reais não negativos [25]. O que diferencia as séries temporais de outros conjuntos de dados é a dependência entre as observações vizinhas.

As séries temporais são decompostas em 4 componentes: aleatoriedade, ciclicidade, sazonalidade e tendência. A

não estacionariedade da série temporal pode ser identificada através da análise da dependência do tempo da distribuição de probabilidade de uma amostra de k observações da série. Se a distribuição de probabilidades não depender do tempo, a série é estacionária. Tal verificação pode ser feita através da observação da dependência do tempo das distribuições unidimensionais, pois caso a série seja estacionária, a média e o desvio padrão são invariantes às translações no tempo [26]. A série temporal de irradiância solar apresenta o caráter não estacionário [27]. Geralmente, as séries temporais não estacionárias podem ser representadas pela combinação das componentes sazonal, s_t , e estacionária, \tilde{z}_t :

$$x_t = s_t + \tilde{z}_t + a_t \quad (1)$$

na qual a_t é o ruído branco gaussiano.

A remoção do caráter não estacionário é realizada com a mudança de variável representada por:

$$z_t = \frac{x_t - \mu_m}{\sigma_m} \quad (2)$$

Em (2), μ_m e σ_m são a média e o desvio padrão da série temporal, respectivamente. A série padronizada obtida apresenta, aproximadamente, média zero e variância unitária. Os dados de irradiância apresentam uma componente de sazonalidade, pois, assim como a provisão de luz solar, é possível observar um comportamento periódico nos dados medidos [17].

2) *Modelo Periódico Auto-Regressivo (PAR)*: As séries temporais que apresentam comportamento periódico podem ser analisados por modelos PAR. Seja x_1, x_2, \dots, x_{ns} uma série temporal com período s , número de dias n , índice de tempo definido em função do número de dias $r = 1, 2, \dots, n$ e horas $m = 1, 2, \dots, s$. Assim, $x_{t(r,m)}$ representa a observação da série x_t no dia r e na hora m . Os modelos PAR de ordem p_m , $m = 1, 2, \dots, s$, PAR(p_m), podem ser representados da seguinte forma [26], [28]:

$$z_{t(r,m)} = \sum_{i=1}^{p_m} \phi_{i,m} z_{t(r,m)-i} + a_{t(r,m)} \quad (3)$$

sendo: $\phi_{i,m}$ é uma função periódica com período s , $a_{t(r,m)}$ é um ruído branco com distribuição normal $N(0, \tau_m^{-1})$, $\tau_m^{-1} = \sigma_m^2$ e $\tau_m > 0$. Por fim, $z_{t(r,m)}$ é a série padronizada, que possui, aproximadamente, média nula e desvio padrão unitário.

Neste trabalho, o método da Máxima Verossimilhança [29] foi considerado, em que a obtenção dos parâmetros $\phi_{i,m}$ é feita através da função de verossimilhança condicional. A seleção do modelo para o caso do PAR(p_m) é feita através de uma técnica que utiliza densidades preditivas ordenadas [26]. Particularmente, é utilizado o método de Monte Carlo para estimar a densidade preditiva $c_{t(r,m)} = \hat{\Pi}_m(z_{t(n+1,m)} | \mathbf{Z}_m)$, sendo \mathbf{Z}_m o vetor das observações anteriores. Nesse caso, o modelo que entregar o maior valor de $c_{t(r,m)}$ é classificado como a melhor escolha. Portanto, o critério escolhe o modelo que maximiza o produto:

$$c(p) = \prod_{r=1}^n c_{t(r,m)}(p) \quad (4)$$

Sendo p o p -ésimo modelo autorregressivo periódico. À título de exemplo, $p = 3$ indica que a densidade preditiva $c_{t(r,m)}$ está sendo calculada para o modelo PAR(3).

III. RESULTADOS E DISCUSSÃO

A. Análise Estatística

Fig. 1 representa o gráfico da irradiância global medida às 11h, durante o intervalo de 365 dias, com início no primeiro dia do outono de 2013 e seu fim no último dia do verão de 2014. A Fig. 2 ilustra o comportamento das medições de irradiância solar às 16h, de forma similar às 11h (Fig. 1).

Esses dois horários, 11h e 16h, foram escolhidos por se destacarem na análise estatística apresentada na Tabela I. Pode-se perceber que o horário das 11h apresenta o maior pico de irradiância, assim como, a maior média e mediana. O horário das 16h foi colocado em destaque devido ao resultado nos testes de normalidade. Apresentando $p\text{-valor} > 0,05$ nos testes de Anderson-Darling e Shapiro-Wilk, foi o único dos 24 bancos de dados que não rejeitou a hipótese H_0 , apresentando aderência à distribuição de probabilidade normal.

Os horários cuja irradiância assume valores desejados no tocante ao aproveitamento térmico ou elétrico, ou seja, entre as 8h e 15h, apresentam uma mediana maior que a média, resultado esperado de variáveis aleatórias cuja distribuição não é normal, uma vez que a média é influenciada por valores extremos enquanto a mediana representa o valor central, sendo uma medida melhor representativa para variáveis não-normais.

Os *boxplot* foram utilizados para avaliar se um conjunto de dados obedece a distribuição normal (Fig. 3). Pode-se notar que a mediana, de cada *boxplot* está deslocada do centro, assim como a caixa está deslocada em relação aos máximos e mínimos. Essas assimetrias demonstram que os dados não estão distribuídos de maneira uniforme em torno da média por hora, fato que corrobora os testes anteriores e confirmam a não aderência a normalidade.

Fig. 3 apresentam *outliers*, pontos fora dos limites máximos e mínimos, cuja sobreposição de pontos próximos dá origem a pontos com tonalidade mais escura. Ademais, a Fig. 3 ilustra o perfil da irradiância solar anual, percebe-se *outliers* acima do máximo, nas primeiras horas da manhã e nas duas últimas horas da tarde. Esses pontos correspondem à medições no verão, época em que o sol nasce mais cedo e se põe mais tarde. Os *outliers* abaixo dos valores mínimos podem indicar medidas do outono/inverno, ou dias nublados de qualquer estação.

B. Previsão de Séries Temporais

Neste trabalho as previsões dos dados de irradiância solar utilizando o modelo PAR são realizadas em frequência horária. Duas situações foram consideradas em relação a ordem do modelo PAR: o modelo PAR de ordem 1, onde o parâmetro de previsão depende apenas da hora anterior à previsão, e o modelo PAR(p_m), onde a ordem do modelo varia com a hora da previsão. Em ambos modelos foram considerados ajustes dos parâmetros de previsão referentes a cada hora de um dia.

O ajuste dos parâmetros de previsão do modelo PAR foi realizado utilizando-se os dados históricos de irradiância solar,

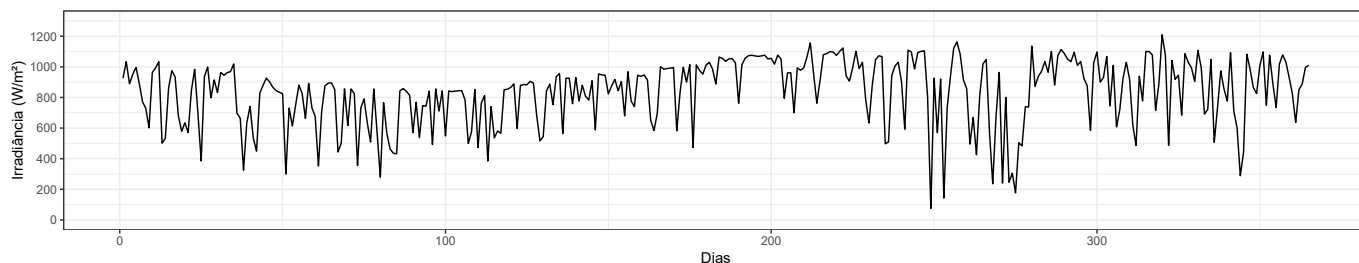


Fig. 1. Irradiância solar anual avaliada às 11h.

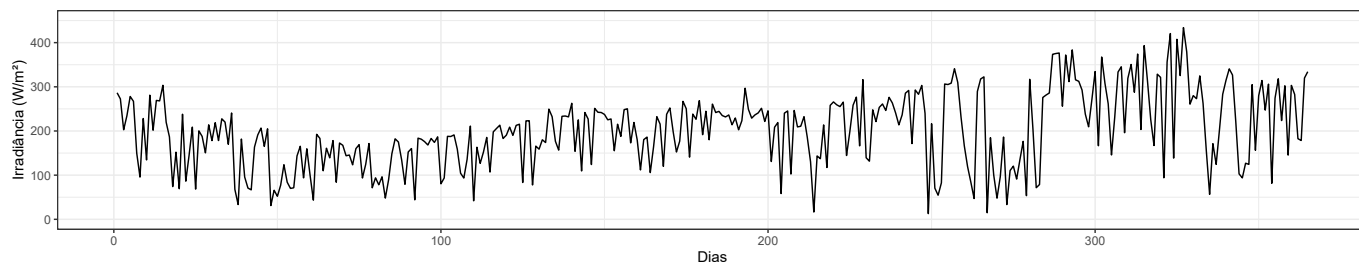


Fig. 2. Irradiância solar anual avaliada às 16h.

TABELA I
RESULTADOS DAS ESTATÍSTICAS DESCRITIVAS E TESTES DE ADERÊNCIA.

| Hora | Mínimo | Máximo | Média | Mediana | Desvio Padrão | Anderson-Darling | AD p-value | Shapiro-Wilk | S-W p-value |
|------|--------|---------|--------|---------|---------------|------------------|------------|--------------|-------------|
| 0 | 0,00 | 0,33 | 0,02 | 0,00 | 0,04 | 65,42 | 0,000 | 0,531 | 0,000 |
| 1 | 0,00 | 0,35 | 0,02 | 0,00 | 0,04 | 53,90 | 0,000 | 0,583 | 0,000 |
| 2 | 0,00 | 0,22 | 0,03 | 0,00 | 0,04 | 50,81 | 0,000 | 0,657 | 0,000 |
| 3 | 0,00 | 0,35 | 0,03 | 0,00 | 0,06 | 50,77 | 0,000 | 0,621 | 0,000 |
| 4 | 0,00 | 0,50 | 0,03 | 0,01 | 0,06 | 45,95 | 0,000 | 0,642 | 0,000 |
| 5 | 0,00 | 50,52 | 7,33 | 1,70 | 11,09 | 44,70 | 0,000 | 0,690 | 0,000 |
| 6 | 16,82 | 273,82 | 104,59 | 87,46 | 63,91 | 10,20 | 0,000 | 0,910 | 0,000 |
| 7 | 39,41 | 533,32 | 295,39 | 288,50 | 113,49 | 0,97 | 0,014 | 0,985 | 0,001 |
| 8 | 70,41 | 773,34 | 493,98 | 506,74 | 153,95 | 1,74 | 0,000 | 0,978 | 0,000 |
| 9 | 145,86 | 953,68 | 674,76 | 700,05 | 180,70 | 5,48 | 0,000 | 0,947 | 0,000 |
| 10 | 146,36 | 1072,26 | 780,54 | 814,62 | 209,02 | 6,29 | 0,000 | 0,936 | 0,000 |
| 11 | 75,05 | 1209,73 | 834,12 | 880,94 | 214,28 | 6,96 | 0,000 | 0,934 | 0,000 |
| 12 | 26,18 | 1151,17 | 808,50 | 848,87 | 216,08 | 5,77 | 0,000 | 0,941 | 0,000 |
| 13 | 14,82 | 1121,51 | 722,96 | 756,95 | 206,57 | 5,65 | 0,000 | 0,945 | 0,000 |
| 14 | 10,44 | 881,74 | 579,80 | 612,93 | 179,53 | 6,36 | 0,000 | 0,942 | 0,000 |
| 15 | 10,87 | 678,04 | 397,04 | 420,86 | 134,12 | 4,85 | 0,000 | 0,964 | 0,000 |
| 16 | 12,90 | 434,01 | 199,74 | 202,22 | 84,31 | 0,54 | 0,163 | 0,992 | 0,057 |
| 17 | 2,23 | 172,78 | 36,40 | 28,16 | 30,36 | 19,30 | 0,000 | 0,824 | 0,000 |
| 18 | 0,00 | 4,87 | 0,29 | 0,00 | 0,72 | 81,38 | 0,000 | 0,468 | 0,000 |
| 19 | 0,00 | 0,32 | 0,01 | 0,00 | 0,03 | 91,99 | 0,000 | 0,290 | 0,000 |
| 20 | 0,00 | 0,47 | 0,01 | 0,00 | 0,04 | 95,79 | 0,000 | 0,273 | 0,000 |
| 21 | 0,00 | 0,53 | 0,01 | 0,00 | 0,04 | 86,45 | 0,000 | 0,343 | 0,000 |
| 22 | 0,00 | 0,29 | 0,02 | 0,00 | 0,04 | 78,38 | 0,000 | 0,442 | 0,000 |
| 23 | 0,00 | 0,47 | 0,02 | 0,00 | 0,04 | 72,23 | 0,000 | 0,426 | 0,000 |

separando os dados das séries entre treino e teste para previsão. Os testes foram realizados considerando-se a previsão de 03 dias consecutivos, sendo selecionados os dias do histórico de irradiância que apresentavam características diferentes em relação à luminosidade: dias de irradiância solar baixa, mediana e altas. Em períodos de pouca irradiância solar, a geração fotovoltaica está mais propensa a sofrer falhas e aumentam-se os riscos de instabilidade no sistema, tornando ainda mais relevante se ter uma boa previsão em tais períodos.

Para avaliar a qualidade da previsão, foi considerado o cálculo da raiz do erro médio quadrado (*Root Mean Square*

Error (RMSE)), definido como:

$$RMSE = \sqrt{\frac{\sum_{i=1}^k (X_{obs,i} - X_{prev,i})^2}{k}} \tag{5}$$

sendo: $X_{obs,i}$, o valor associado ao dado observado, $X_{prev,i}$, o resultado da previsão e k , o horizonte de previsão. O uso do RMSE permite obter uma intuição geométrica acerca da diferença entre o valor predito e o valor observado, pois é possível relacioná-lo a uma distância euclidiana a menos de uma constante multiplicativa. O RMSE é popular por conta

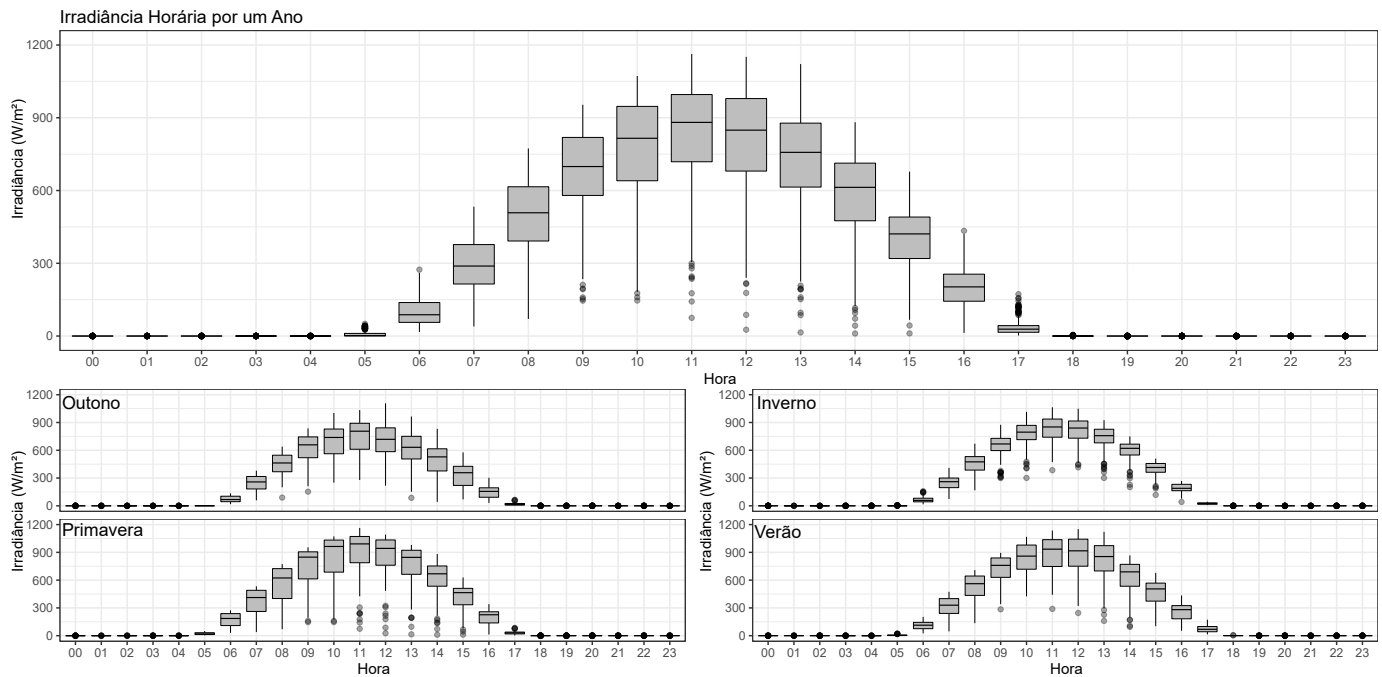


Fig. 3. *Boxplot* da irradiância solar ao longo do ano e por estações do ano.

de algumas particularidades, como: sensibilidade a *outliers* e possuir as mesmas dimensões da variável estudada, sendo características importantes para a análise da acurácia de modelos de previsão [30], [31].

1) *Modelo Periódico Auto-Regressivo de ordem 1 (PAR(1))*: O modelo PAR(1) considera que a irradiância solar na hora corrente depende apenas da irradiância solar da hora imediatamente anterior.

Tabela II mostra os parâmetros $\phi_{1,m}$ para o modelo PAR(1) obtidos considerando-se os dados históricos de irradiância solar, retirando-se os 3 últimos dias, período considerado de irradiações medianas. Foram considerados 24 parâmetros, sendo um parâmetro para cada hora do dia, e um horizonte de previsão de 72 horas. Fig.4 apresenta os dados reais e a previsão no período considerado. Neste caso o RMSE obtido foi de $38,46 \text{ W/m}^2$

TABELA II
PARÂMETROS ϕ_1 PARA O PAR(1) - IRRADIÂNCIA SOLAR
MEDIANA

| Hora | $\phi_{1,m}$ | Hora | $\phi_{1,m}$ |
|------|--------------|------|--------------|
| 1 | -0,0093 | 13 | 0,8740 |
| 2 | 0,1789 | 14 | 0,8787 |
| 3 | 0,0815 | 15 | 0,8404 |
| 4 | 0,1828 | 16 | 0,8351 |
| 5 | 0,2181 | 17 | 0,7727 |
| 6 | 0,0236 | 18 | 0,6703 |
| 7 | 0,7080 | 19 | 0,7340 |
| 8 | 0,7375 | 20 | 0,1396 |
| 9 | 0,8254 | 21 | 0,2050 |
| 10 | 0,8087 | 22 | 0,2898 |
| 11 | 0,8266 | 23 | 0,1539 |
| 12 | 0,8484 | 24 | 0,0936 |

Pode-se observar, de acordo com a Fig. 4, que o modelo PAR(1) forneceu a previsão que segue o perfil do série obser-

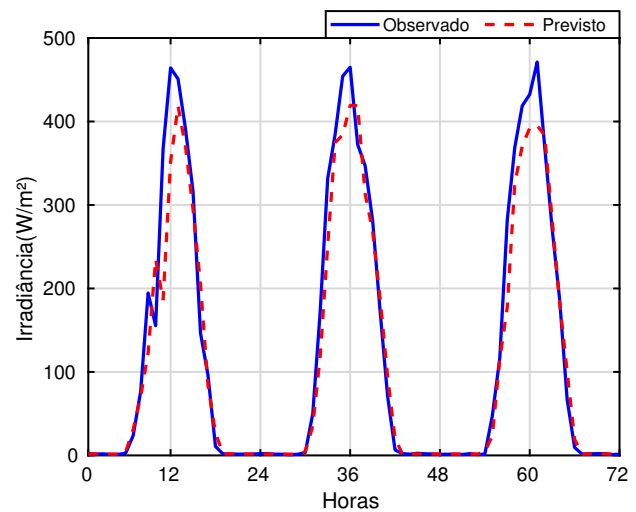


Fig. 4. Previsão do modelo PAR(1), horizonte de 3 dias e período de irradiância solar medianas.

vada de forma adequada. As maiores diferenças são verificadas nos horários de pico da série.

A fim de analisar o comportamento do PAR(1) em períodos com diferentes intensidades de irradiância solar, a Tabela III mostra os parâmetros $\phi_{1,m}$ para o modelo PAR(1) obtidos considerando-se os dados históricos de irradiância solar, retirando-se os 3 últimos dias de menores registros de irradiância solar, período considerado de irradiações baixas. Foram considerados 24 parâmetros, referentes a cada hora do dia, e um horizonte de previsão de 72 horas. Fig. 5 apresenta os dados reais e a previsão no período neste caso, onde o RMSE obtido foi igual a $16,90 \text{ W/m}^2$

TABELA III
PARÂMETROS ϕ_1 PARA O PAR(1) - IRRADIÂNCIA SOLAR BAIXA

| Hora | $\phi_{1,m}$ | Hora | $\phi_{1,m}$ |
|------|--------------|------|--------------|
| 1 | 0,0068 | 13 | 0,8763 |
| 2 | 0,1790 | 14 | 0,8805 |
| 3 | 0,0793 | 15 | 0,8489 |
| 4 | 0,1806 | 16 | 0,8324 |
| 5 | 0,2161 | 17 | 0,7816 |
| 6 | 0,0269 | 18 | 0,6739 |
| 7 | 0,7116 | 19 | 0,7340 |
| 8 | 0,7316 | 20 | 0,1441 |
| 9 | 0,8251 | 21 | 0,1924 |
| 10 | 0,8054 | 22 | 0,2871 |
| 11 | 0,8234 | 23 | 0,1650 |
| 12 | 0,8498 | 24 | 0,1015 |

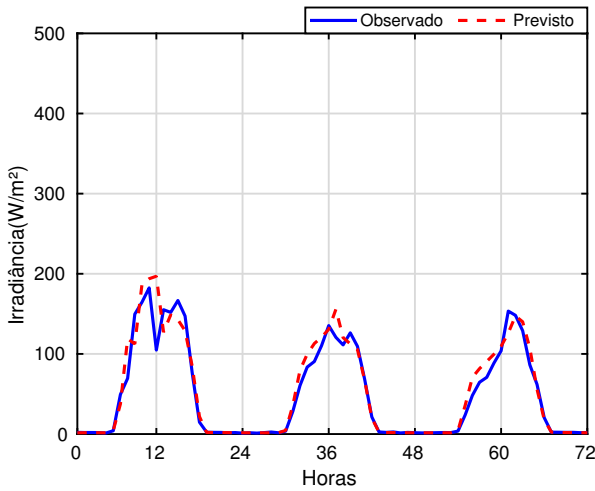


Fig. 5. Previsão do modelo PAR(1), horizonte de 3 dias e período de irradiância solar baixas.

2) *Modelo Periódico Auto-Regressivo de ordem p_m* ($PAR(p_m)$): O $PAR(p_m)$ considera que a irradiância solar na hora corrente depende de dados de irradiância solar de m horas imediatamente anteriores. Neste caso a ordem do modelo varia para cada uma das 24 horas consideradas.

Tabela IV mostra os parâmetros $\phi_{p,m}$ para cada hora do dia, obtidos considerando-se os dados históricos de irradiância solar, retirando-se os 3 últimos dias, período considerado de irradiações medianas. Observa-se que o modelo $PAR(p_m)$ apresentou ordem um ou dois para a maior parte das horas do dia, com exceção das horas 8 e 10, que apresentam ordem 5 e 4, respectivamente. Foi considerado um horizonte de previsão de 72 horas. Fig. 6 apresenta os dados reais e a previsão no período considerado, sendo o RMSE igual a $36,59 W/m^2$.

De acordo com a Fig. 6, as maiores diferenças ocorreram no pico das séries, normalmente por volta do meio dia, de forma semelhante ao modelo PAR(1), (Fig. 4). Os valores de RMSE no período considerado foram bastante próximos no caso do PAR(1) e $PAR(p_m)$, $38,46 W/m^2$ e $36,59 W/m^2$ respectivamente. Vale observar que no modelo $PAR(p_m)$, das 24 horas do dia, 13 horas apresentaram modelo PAR de ordem superior a um, sendo que destas 12 horas, apenas 3 horas apresentaram ordem superior a dois, o que pode justificar a similaridade de comportamento entre os modelos PAR(1) e

TABELA IV
PARÂMETROS $\phi_{n,m}$ PARA O PAR(p_m) - IRRADIÂNCIA MEDIANA.

| Hora | $\phi_{1,m}$ | $\phi_{2,m}$ | $\phi_{3,m}$ | $\phi_{4,m}$ | $\phi_{5,m}$ |
|------|--------------|--------------|--------------|--------------|--------------|
| 1 | 0,0030 | -0,1311 | 0 | 0 | 0 |
| 2 | 0,1789 | 0 | 0 | 0 | 0 |
| 3 | 0,0815 | 0 | 0 | 0 | 0 |
| 4 | 0,1982 | -0,1869 | 0 | 0 | 0 |
| 5 | 0,2161 | 0 | 0 | 0 | 0 |
| 6 | 0,0269 | 0 | 0 | 0 | 0 |
| 7 | 0,7116 | 0 | 0 | 0 | 0 |
| 8 | 0,9410 | -0,3183 | -0,0727 | -0,0876 | -0,0822 |
| 9 | 0,8948 | -0,0315 | -0,1307 | 0 | 0 |
| 10 | 0,9402 | -0,2927 | 0,2410 | -0,1370 | 0 |
| 11 | 0,8234 | 0 | 0 | 0 | 0 |
| 12 | 0,9797 | -0,1579 | 0 | 0 | 0 |
| 13 | 1,0386 | -0,1906 | 0 | 0 | 0 |
| 14 | 0,9987 | -0,1355 | 0 | 0 | 0 |
| 15 | 1,0338 | -0,2102 | 0 | 0 | 0 |
| 16 | 0,8324 | 0 | 0 | 0 | 0 |
| 17 | 0,9244 | -0,1714 | 0 | 0 | 0 |
| 18 | 0,8724 | -0,2547 | 0 | 0 | 0 |
| 19 | 0,7340 | 0 | 0 | 0 | 0 |
| 20 | 0,1441 | 0 | 0 | 0 | 0 |
| 21 | 0,1692 | 0,1599 | 0 | 0 | 0 |
| 22 | 0,2871 | 0 | 0 | 0 | 0 |
| 23 | 0,1650 | 0 | 0 | 0 | 0 |
| 24 | 0,1283 | -0,1629 | 0 | 0 | 0 |

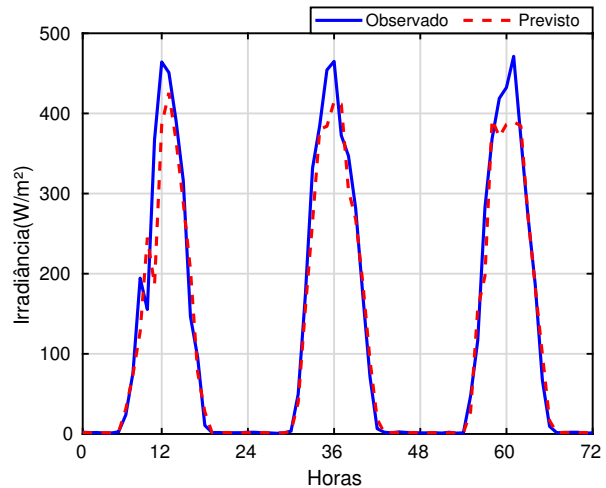


Fig. 6. Previsão do modelo $PAR(p_m)$, horizonte de 3 dias e período de irradiância solar medianas.

$PAR(p_m)$ no horizonte considerado.

A Tabela V mostra os parâmetros $\phi_{p,m}$ para a o modelo $PAR(p_m)$ obtidos considerando-se os dados históricos de irradiância solar. Foram retirados os 3 últimos dias de menores registros de irradiância solar por considerar um período de irradiações baixas. Fig. 7 apresenta os dados reais e a previsão no período neste caso, em que o RMSE encontrado é $18,32 W/m^2$.

Pode-se observar que o modelo $PAR(p_m)$ apresentou discrepâncias maiores no primeiro dia de previsão para o caso de baixos valores de irradiância solar (ver Fig. 7), quando comparado com o resultado do modelo PAR(1) no mesmo período (ver Fig. 5). Como consequência, o RMSE neste caso foi maior do que o obtido com o modelo PAR(1) no período ($18,32 W/m^2$ no caso do $PAR(p_m)$) e $16,90 W/m^2$ no caso

TABELA V
PARÂMETROS $\phi_{n,m}$ PARA O PAR(p_m) - IRRADIÂNCIA
BAIXA.

| Hora | $\phi_{1,m}$ | $\phi_{2,m}$ | $\phi_{3,m}$ | $\phi_{4,m}$ | $\phi_{5,m}$ |
|------|--------------|--------------|--------------|--------------|--------------|
| 1 | 0,0211 | -0,1361 | 0 | 0 | 0 |
| 2 | 0,1790 | 0 | 0 | 0 | 0 |
| 3 | 0,0793 | 0 | 0 | 0 | 0 |
| 4 | 0,1953 | -0,1869 | 0 | 0 | 0 |
| 5 | 0,2181 | 0 | 0 | 0 | 0 |
| 6 | 0,0236 | 0 | 0 | 0 | 0 |
| 7 | 0,7138 | -0,0893 | -0,0251 | -0,1215 | 0 |
| 8 | 0,9300 | -0,3062 | -0,0718 | -0,0887 | -0,0893 |
| 9 | 0,8899 | -0,0239 | -0,1331 | 0 | 0 |
| 10 | 0,9441 | -0,2924 | 0,2391 | -0,1361 | 0 |
| 11 | 0,8226 | 0 | 0 | 0 | 0 |
| 12 | 0,9780 | -0,1571 | 0 | 0 | 0 |
| 13 | 1,0296 | -0,1830 | 0 | 0 | 0 |
| 14 | 0,9855 | -0,1257 | 0 | 0 | 0 |
| 15 | 1,0529 | -0,2421 | 0 | 0 | 0 |
| 16 | 0,8531 | 0 | 0 | 0 | 0 |
| 17 | 0,9343 | -0,1945 | 0 | 0 | 0 |
| 18 | 0,8479 | -0,2300 | 0 | 0 | 0 |
| 19 | 0,7340 | 0 | 0 | 0 | 0 |
| 20 | 0,1789 | 0 | 0 | 0 | 0 |
| 21 | 0,1847 | 0,1457 | 0 | 0 | 0 |
| 22 | 0,2892 | 0 | 0 | 0 | 0 |
| 23 | 0,1539 | 0 | 0 | 0 | 0 |
| 24 | 0,1138 | -0,1299 | 0 | 0 | 0 |

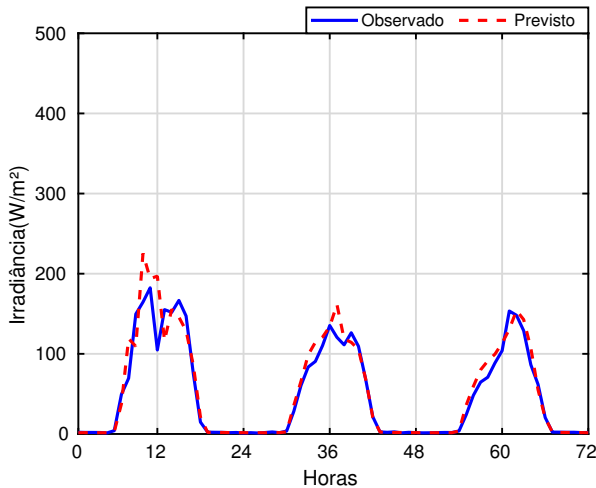


Fig. 7. Previsão do modelo PAR(p_m), horizonte de 3 dias e período de irradiação solar baixas.

do PAR(1)).

À título de comparação dos resultados obtidos neste trabalho com outros artigos, destaca-se a contribuição de [32] que usa modelos de *deep learning* para realizar previsões em séries temporais de irradiação. Os autores utilizam o RMSE como figura de desempenho para o modelo de predição, foram obtidos RMSE iguais a $50,89 \text{ W/m}^2$ para previsões que utilizam apenas informações dos dados de irradiação passados.

Assim, lançando mão do princípio da parcimônia, usar um modelo PAR de ordem 1 pode ser interessante, em especial se a ordem do modelo de previsão influenciar os demais processos. Portanto usar um modelo PAR de menor ordem, que oferece desempenho semelhante, é mais eficaz, pois diminui o custo computacional, podendo proporcionar uma redução

significativa no tempo de simulação.

IV. CONCLUSÃO

Este artigo propõe duas abordagens para modelagem da série temporal da irradiação solar de uma estação solarimétrica, localizada numa cidade no interior da Bahia, que se encontra na zona climática tropical. As abordagens consistiram na análise estatística exploratória da frequência horária, onde permitiu obter as características das séries, como: variação estocástica, assimetria, comportamento e distribuição das medições da irradiação solar anual. Em seguida, o desenvolvimento de um modelo periódico de séries temporais para previsão da irradiação solar que captasse a característica periódica e estocástica das medições.

Com base na análise estatística exploratória para todas as frequências horárias do dia durante um ano (Tabela I), observou-se que o horário das 11h apresentou os maiores valores de média e mediana. Dois testes paramétricos foram usados para verificar a hipótese de Normalidade dos dados anuais, o Anderson-Darling e o Shapiro-Wilk. O valor de p - valor $< 0,05$ indica a rejeição da hipótese H_0 , que pressupõe a normalidade. Tabela I mostrou que a rejeição da normalidade se deu em todos os horários, exceto no horário das 16h para ambos testes.

O *boxplot* permite um estudo gráfico da distribuição e simetria dos dados, bem como, da existência de pontos extremos (*outliers*). Foi realizado o *boxplot* para cada hora usando os dados anuais, assim como, em cada hora nos dados divididos por estações do ano. Pôde-se observar diversos *outliers* abaixo dos limites mínimos devido aos momentos de céu nublado. Nos *boxplots* por estação observa-se uma maior simetria no outono e inverno, e evidente assimetria nas estações primavera e verão, o que indica que a distribuição de frequência muda de acordo com a estações, evidenciando as características climáticas dos dados.

Observa-se pelos testes de normalidade, nos gráficos de *boxplot* e nos histogramas que a irradiação solar não segue a distribuição normal na maioria das medições de frequências horárias. A investigação do tipo de distribuição dos dados anuais e maior aprofundamento no estudo sazonal são temas para trabalhos futuros.

Com relação previsão de séries temporais de irradiação solar, qualidade de previsão foi avaliada pelo *Root Mean Square Error* (RMSE) e, de uma forma geral, os modelos PAR considerados apresentaram resultados bastante semelhantes, comprovando a ideia de comportamento periódico da série de irradiação solar. Como a diferença do RMSE entre o modelo periódico auto-regressivo de ordem 1 e ordem p_m não é significativa, usar um modelo de ordem 1 pode ser interessante neste caso. Menores erros de previsão foram verificados em períodos de irradiação solar baixa para os dois modelos. Trata-se de uma característica importante do modelo de previsão, já que em períodos com baixa irradiação solar a geração fotovoltaica está mais propensa a sofrer falhas e causar instabilidades, tornando ainda mais relevante se ter uma boa previsão em tais períodos.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por meio do programa PIBIC/UFBA - CNPq. Agradecemos também à Empresa de Energia Renovável que disponibilizou os dados de um sítio na Região Nordeste do Brasil.

REFERÊNCIAS

[1] A. G. Ferreira, *Meteorologia prática*. Oficina de Textos, 2006.
 [2] EPE, “Plano nacional de energia 2050,” *Rio De Janeiro*, 2015.
 [3] ANEEL, “Atlas de energia elétrica, do brasil,” *Brasília, Brasil. 3aEd*, 2008.
 [4] C. Tiba, N. Fraidenaich, F. Lyra, and A. Nogueira, “Atlas solarimétrico do brasil: banco de dados terrestres,” *Recife: Editora Universitária da UFPE*, p. 32, 2000.
 [5] C. G. Ribeiro, H. B. A. Neto, and T. S. Sene, “A oscilação do preço do petróleo: uma análise sobre o período entre 2010-2015,” *Estudos internacionais: revista de relações internacionais da PUC Minas*, vol. 6, no. 1, pp. 87–106, 2018.
 [6] C. A. Nobre, J. Reid, and A. P. S. Veiga, “Fundamentos científicos das mudanças climáticas,” *São José dos Campos, SP: Rede Clima/INPE*, 2012.
 [7] F. R. Martins, T. G. Soares, and F. J. L. Lima, “Generating solar irradiance data series with 1-minute time resolution based on hourly observational data,” *IEEE Latin America Transactions*, vol. 100, no. 1e, 2020.
 [8] ANEEL, “Atlas brasileiro de energia elétrica,” *Agência Nacional de Energia Elétrica, Brasília, Brasil. 3aEd*, 2008.
 [9] R. Bondarik, L. A. Pilatti, and D. J. Horst, “Uma visão geral sobre o potencial de geração de energias renováveis no brasil,” *Interciencia*, vol. 43, no. 10, pp. 680–688, 2018.
 [10] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, “Review of solar irradiance forecasting methods and a proposition for small-scale insular grids,” *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 65–76, 2013.
 [11] V. Kallio-Myers, A. Riihelä, P. Lahtinen, and A. Lindfors, “Global horizontal irradiance forecast for finland based on geostationary weather satellite data,” *Solar Energy*, vol. 198, pp. 68–80, 2020.
 [12] J. Mubiru and E. Banda, “Estimation of monthly average daily global solar irradiation using artificial neural networks,” *Solar energy*, vol. 82, no. 2, pp. 181–187, 2008.
 [13] J. Gómez, F. Carlesso, L. Vieira, and L. Da Silva, “Solar irradiance: basic concepts,” *Revista Brasileira de Ensino de Física*, vol. 40, no. 3, 2018.
 [14] F. Besharat, A. A. Dehghan, and A. R. Faghhi, “Empirical models for estimating global solar radiation: A review and case study,” *Renewable and Sustainable Energy Reviews*, vol. 21, pp. 798–821, 2013.
 [15] A. Angstrom, “Solar and terrestrial radiation. report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation,” *Quarterly Journal of the Royal Meteorological Society*, vol. 50, no. 210, pp. 121–126, 1924.
 [16] G. Mihalakakou, M. Santamouris, and D. Asimakopoulos, “The total solar radiation time series simulation in athens, using neural networks,” *Theoretical and Applied Climatology*, vol. 66, no. 3-4, pp. 185–197, 2000.
 [17] C. Voyant, J. G. De Gooijer, and G. Nottton, “Periodic autoregressive forecasting of global solar irradiation without knowledge-based model implementation,” *Solar Energy*, vol. 174, pp. 121–129, 2018.
 [18] L. F. N. Lourenço, M. B. de Camargo Salles, M. M. F. Gemignani, M. R. Gouvea, and N. Kagan, “Time series modelling for solar irradiance estimation in northeast brazil,” in *2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA)*. IEEE, 2017, pp. 401–405.
 [19] A. Mellit and A. M. Pavan, “A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected pv plant at trieste, italy,” *Solar Energy*, vol. 84, no. 5, pp. 807–821, 2010.
 [20] E. J. Balbinot, J. W. Scotton, S. M. Cerezer, and C. A. Martinazzo, “Modelos de séries temporais aplicados a previsão de radiação solar,” *Revista Perspectiva*, 2017.
 [21] J. Bosch, G. Lopez, and F. Batlles, “Daily solar irradiation estimation over a mountainous area using artificial neural networks,” *Renewable Energy*, vol. 33, no. 7, pp. 1622–1628, 2008.

[22] A. Mellit, M. Benganem, and S. A. Kalogirou, “An adaptive wavelet-network model for forecasting daily total solar-radiation,” *Applied Energy*, vol. 83, no. 7, pp. 705–722, 2006.
 [23] S. Pashiardis, S. A. Kalogirou, and A. Pelengaris, “Statistical analysis for the characterization of solar energy utilization and inter-comparison of solar radiation at two sites in cyprus,” *Applied energy*, vol. 190, pp. 1138–1158, 2017.
 [24] G. Novais, “Distribuição média dos climas zonais no globo: estudos preliminares de uma nova classificação climática (average distribution of zonal climates on the globe: preliminary studies of a new climatic classification),” *Revista Brasileira de Geografia Física*, vol. 10, no. 5, 2017.
 [25] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
 [26] R. Ballini, *Análise e previsões de vazões utilizando modelos de séries temporais, redes neurais e redes neurais nebulosas*. Tese de doutorado, Doutorado em Engenharia Elétrica, Unicamp, 2000.
 [27] C. W. Hansen, J. S. Stein, and A. Ellis, “Statistical criteria for characterizing irradiance time series,” *Sandia National Laboratories SAND2010-7314*, 2010.
 [28] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Prentice-Hall International, Inc, 2008.
 [29] L. Martinez *et al.*, “Políticas de controle malha fechada e malha aberta no planejamento da operação energética de sistemas hidrotérmicos,” 2001.
 [30] C. Chen, J. Twycross, and J. M. Garibaldi, “A new accuracy measure based on bounded relative error for time series forecasting,” *PLoS one*, vol. 12, no. 3, p. e0174202, 2017.
 [31] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” *International journal of forecasting*, vol. 22, no. 4, pp. 679–688, 2006.
 [32] B.-k. Jeon and E.-J. Kim, “Next-day prediction of hourly solar irradiance using local weather forecasts and lstm trained with non-local data,” *Energies*, vol. 13, no. 20, p. 5258, 2020.



Bruna Rodrigues é estudante de Engenharia Elétrica da Universidade Federal da Bahia. A partir de 2019, tornou-se membro do Grupo de Pesquisa em Sistemas Elétricos de Potência, que é um grupo filiado ao Capítulo Estudantil IEEE PES UFBA, e ingressou em um projeto de iniciação científica, como bolsista do PIBIC-CNPq, para estudar a modelagem de sistemas utilizando Séries Temporais e aplicá-las em sistemas fotovoltaicos.



Matheus Gomes é estudante de Engenharia Elétrica da Universidade Federal da Bahia. A partir de 2017, tornou-se membro do Grupo de Pesquisa em Sistemas Elétricos de Potência, que é um grupo filiado ao Capítulo Estudantil IEEE PES UFBA. Em 2018, ingressou em um projeto de iniciação científica, como bolsista do PIBIC-CNPq, para estudar a modelagem de sistemas utilizando Séries Temporais e aplicá-las em sistemas fotovoltaicos.



Ângelo Sant'Anna possui graduação em Estatística pela Universidade Federal da Bahia (2003), mestrado e doutorado em Engenharia de Produção pela Universidade Federal do Rio Grande do Sul (2006, 2009). Atualmente é professor do Departamento de Engenharia Mecânica da Universidade Federal da Bahia. Tem experiência na área de Estatística, Engenharia da Qualidade e Pesquisa Operacional, atuando principalmente nos seguintes temas: monitoramento de processos, otimização multiobjetivo.



Daniel Barbosa possui graduação em Engenharia Elétrica pela Universidade Federal da Bahia (2005), mestrado e doutorado em Engenharia Elétrica pela Universidade de São Paulo, Campus São Carlos (2007, 2010). Atualmente é professor da Universidade Federal da Bahia. Tem experiência na área de Engenharia Elétrica, com ênfase em Medição, Controle, Correção e Proteção de Sistemas Elétricos de Potência, atuando nos seguintes temas: EMTP-ATP, sistemas de elétricos e geração distribuída.



Luciana Martinez possui graduação em Bacharelado em Matemática pela Universidade de São Paulo (1994), mestrado em Ciências pela Universidade de São Paulo (1996) e doutorado em Engenharia Elétrica pela Universidade Estadual de Campinas (2001). Atualmente é professora da Universidade Federal da Bahia. Tem experiência em Otimização Linear e Não Linear, atuando principalmente nos seguintes temas: planejamento energético e sistemas de energia elétrica.