

Generating Solar Irradiance Data Series with 1-minute Time Resolution Based on Hourly Observational Data

Thaiane G. Soares, Francisco J. L. Lima, and Fernando R. Martins

Abstract—The large-scale integration of intermittent renewable energy sources into electricity distribution is one of the significant challenges in meeting energy demand in the near future. The solar energy resource has a fluctuating spatial and temporal character intrinsic to the radiative processes with the atmosphere and the Earth’s surface. Efficient use of solar energy requires technologies that consider such variability and its impacts on the electrical system. In this context, a reliable computational model to estimate the surface solar irradiance is beneficial for operational issues. This work presents a methodology based on Markov chains for the generation of a global irradiance time series with one-minute temporal resolution using observational database acquired by a basic weather station. The statistical model was developed using irradiance data acquired in São Martinho da Serra/RS and Petrolina/PE to verify their performance in two different climates. The transition probability matrices were determined for the following cloud conditions: cloudy, partly cloudy or clear sky using ten years of observational data. Next, Markov’s walking technique was used to generate a data series with one-minute temporal resolution taking into account the cloudiness classification. Model validation was performed using solar irradiance data not considered to calculate the transition probability matrix (MPT). The mean squared deviation of the relative distribution frequencies indicated a deviation of only 1.5%, and the Kolmogorov-Smirnov test indicated that the synthetic and observational series presented similar cumulative frequency distributions. Thus, the proposed methodology showed high reliability in reproducing the temporal variability due to the stochastic nature of the incoming solar energy in both sites.

Index Terms—Markov Processes, Numerical simulation, Remote sensing, Renewable energy sources, Solar energy.

I. INTRODUÇÃO

O sol é fonte de energia renovável e inesgotável; e o aproveitamento dessa energia é uma das alternativas energéticas no Brasil uma vez que a incidência de radiação solar no território brasileiro é elevada durante todo o ano [1] [2] e, o país possui grandes reservas de quartzo, matéria-prima utilizada na fabricação de painéis fotovoltaicos para conversão de energia solar em eletricidade [3].

Um dos maiores desafios para o atendimento da demanda crescente de energia será a integração em larga escala de fontes renováveis de energias intermitentes às redes de distribuição de eletricidade [4]. As fontes renováveis de energia, como a solar e a eólica, possuem um caráter flutuante espacial e temporal intrínsecos aos processos físicos da natureza e a sua utilização

de forma eficiente na geração de eletricidade deve adotar tecnologias operacionais que considerem essa variabilidade [5] de forma criteriosa tanto do ponto de vista econômico quanto do ponto de vista de segurança do sistema elétrico [6]. No Brasil, cabe ao Operador Nacional do Sistema Elétrico (ONS) e as companhias de distribuição de eletricidade lidar com dificuldades associadas às variações rápidas de inserção de energia intermitente e com sobrecargas no sistema de distribuição de eletricidade.

O aproveitamento do recurso de energia solar integrado ao sistema nacional de eletricidade apresenta desafios tecnológicos associados com as características intrínsecas de variabilidade espacial e temporal da fonte energética e que, em alguns momentos, pode proporcionar um desequilíbrio na relação geração/consumo de eletricidade, flutuações de tensão e outros aspectos técnicos com impacto direto na qualidade da energia distribuída aos consumidores. Com o crescimento da participação da energia solar na matriz elétrica brasileira, a confiabilidade do sistema elétrico torna-se dependente da capacidade do sistema para acomodar a variabilidade intrínseca do recurso solar, ao mesmo tempo em que mantém a qualidade e a continuidade do serviço aos consumidores [6] [7].

Neste contexto, um modelo numérico confiável de estimativa de irradiância solar incidente na superfície é bastante útil [6] [8]. A informação sobre a irradiância solar incidente na superfície pode contribuir no gerenciamento e na operação de uma planta solar, principalmente quando um sistema de armazenamento da energia excedente está disponível ou a tarifação da eletricidade é realizada com base em horários de ponta de consumo [9].

Um método numérico eficiente para estimativa da irradiância solar para horizontes de tempo curtos (dezenas de minutos) ou curtíssimos (alguns minutos) ajudará tanto o gerenciamento quanto o planejamento e a otimização dos regimes de operação das instalações de geração solar e do sistema elétrico considerando eventos repentinos (rampas de crescimento ou decrescimento) da incidência solar [6]. Os eventos repentinos estão associados principalmente com a passagem de nebulosidade e é um fenômeno presente em todo o território brasileiro, mesmo na região mais árida localizada no Nordeste do Brasil [10].

Os modelos físicos resolvem numericamente as equações de transferência radiativa que descrevem os processos de absorção e espalhamento de radiação solar na atmosfera e requerem a disponibilidade de dados observados em estações meteorológicas e dados observados por satélites [11]. Além

Thaiane G. Soares, Francisco J. L. Lima and Fernando R. Martins are with the Instituto do Mar, Universidade Federal de São Paulo (Unifesp), *campus* Baixada Santista, 11030-100 Brazil e-mail: (fernando.martins@unifesp.br).

Manuscript received August xx, 2019; revised August xx, 2019.

da demanda elevada de dados para geração de estimativas confiáveis, os modelos físicos, e demandam um tempo computacional elevado quando aplicados para áreas extensas. A combinação dessas características tornam os modelos físicos adequados para produzir estimativas para horizontes de tempo da ordem de dezenas de minutos a horas e resolução espacial de alguns quilômetros. No entanto, os modelos físicos apresentam desvantagens quando o foco reside na produção de estimativas para em locais específicos de interesse como uma planta de geração fotovoltaica ou um conjunto de sistemas fotovoltaicos operando em uma região, principalmente se há estações de coleta de dados em superfície nas proximidades da área de estudo (distâncias menores que 35km) [12].

Os modelos estatísticos são aplicados para estimativas de irradiância solar em superfície para horizontes de tempo da ordem de alguns minutos [13]. Em geral, os modelos estatísticos baseiam-se em dados históricos observados em estações automáticas de coleta de dados meteorológicos e estabelecem uma relação entre as variáveis de entrada (preditores) e as variáveis de saída – irradiância solar na superfície ou energia gerada [14]. Os modelos estatísticos são classificados em duas categorias: métodos estatísticos tradicionais e técnicas de aprendizado de máquina (ou inteligência artificial). Os modelos auto-regressivos (AR) de séries temporais, auto-regressivos de média móvel (ARMA), modelos de regressão múltipla são exemplos de métodos estatísticos tradicionais [15]. Os modelos de aprendizado de máquina empregam métodos de lógica difusa, redes neurais artificiais (RNAs) e algoritmos genéticos [15] [14].

Diversas bases de dados observados de irradiância solar incidente na superfície estão disponíveis para prover dados para aplicações no setor elétrico. Sejam comerciais ou públicas [16] [17] [18], estas bases de dados fornecem, em geral, irradiância solar incidente na superfície com uma resolução temporal horária (1 registro de irradiância solar a cada hora). Todavia, considerando os aspectos mencionados anteriormente, há necessidade de resolução temporal mais fina para realizações simulações que possam contribuir no gerenciamento e na operação de plantas de geração solar. Planejamento e gerenciamento de operação realizadas com estimativas horárias de irradiância solar negligenciam aumentos ou reduções de incidência de energia solar em intervalos curtos de tempo associadas com a nebulosidade e outras condições atmosféricas e, por consequência, podem resultar em procedimentos operacionais com incertezas elevadas.

O estudo apresentado neste artigo teve por objetivo o desenvolvimento de uma metodologia para prover uma série temporal de irradiância solar com resolução de um minuto tomando como ponto de partida uma base de dados horários de irradiância solar na superfície. Informações confiáveis da variabilidade da irradiância solar na superfície com resolução temporal de um minuto possibilitam uma simulação que represente as variações bruscas de incidência causadas pela passagem de pequenas nuvens e outros fenômenos meteorológicos típicos da região geográfica onde a planta de geração solar está operando. Este estudo teve como ponto de partida o trabalho desenvolvido por [19] baseado no uso de cadeias de Markov para a geração de uma série temporal de dados de irradiância

solar na superfície para intervalos de tempo de um minuto.

A metodologia adotada assume que a variabilidade de curtíssimo prazo da irradiância solar incidente na superfície é um processo estocástico no qual seu estado futuro depende apenas do seu estado atual. Assim, os estados passados (em tempos anteriores ao atual) não influenciam no estado futuro. Esta é a característica principal dos processos estocásticos de Markov. Se o espaço de estados possíveis é discreto (enumerável) como é o caso da base de dados de irradiância solar medida em intervalos de tempo definidos, então o modelo de Markov é denominado Cadeia de Markov [20].

II. MATERIAIS E MÉTODOS

A. Dados de Irradiância Solar Incidente em Superfície

Este trabalho foi desenvolvido com uso de dados coletados em duas estações radiométricas da Rede SONDA (<http://sonda.ccst.inpe.br>) operadas pelo Lab. de Modelagem e Estudos de Recursos Renováveis de Energia/Instituto Nacional de Pesquisas Espaciais (LABREN/INPE):

- estação em São Martinho da Serra/RS, na região Sul do Brasil – 29° 26' 34" S; 53° 49' 23" O; 489 m de altitude;
- estação Petrolina/PE, no Nordeste Brasileiro – 09° 04' 08" S; 40° 19' 11" O; 387 m de altitude.

As duas estações foram escolhidas para possibilitar avaliação em duas regiões com características climáticas distintas, e em razão da disponibilidade de uma série histórica de dados observados com 10 anos (2005 a 2015).

As estações registram valores médios de irradiância global e difusa no plano horizontal, e irradiância direta normal. As estações atendem os padrões de qualidade de operação estabelecidos para a "Baseline Surface Solar Radiation" (BSRN) [21]. Além da aquisição de dados radiométricos, as estações coletam dados de variáveis meteorológicas básicas: temperatura do ar, umidade relativa, pressão atmosférica, precipitação, direção e velocidade dos ventos [22].

A estação São Martinho da Serra situa-se na região central do estado do Rio Grande do Sul. A climatologia local sofre influência da Alta Subtropical do Atlântico e circulação associada. Outra forte influência é o sistema de baixa continental (Baixa do Chaco) mais intenso no verão e que condiciona a circulação em baixos níveis e os fluxos de umidade para a região Sul. A precipitação em São Martinho da Serra é bem distribuída ao longo do ano. O maior número de dias com chuva ocorre durante os meses de inverno, porém o volume total mensal da precipitação é semelhante em todos os meses do ano com valores médios entre 140 e 150 mm. A umidade relativa do ar mantém-se em média na faixa entre 65% e 75%. A temperatura máxima é, em média, 24°C, enquanto que a média da temperatura mínima é da ordem de 14°C [23] [24].

A estação de Petrolina-PE localiza-se no semiárido do nordeste brasileiro. Seu clima é tropical semiárido seco e quente. A precipitação no semiárido nordestino caracteriza-se pela escassez e irregularidade pluviométrica e está associada a sistemas sinóticos ou circulações atmosféricas regionais, como os Anticiclones Subtropicais do Atlântico Sul (ASAS) e Atlântico Norte (ASAN), e o cavado equatorial [25]. O total anual médio de precipitação é da ordem de 560 mm.

As chuvas ocorrem de janeiro a abril devido ao deslocamento da Zona de Convergência Intertropical (ZCIT) em direção ao Hemisfério Sul, que influencia na convergência de umidade e da convecção local [26] [27]. Também é importante mencionar a atuação das estruturas de brisa e a convecção local, que mesmo sendo mecanismos de menor escala, contribuem para os totais pluviométricos nas regiões litorâneas [24]. Nos meses mais úmidos, a umidade relativa do ar varia, em média, entre 66% e 72%. Menores valores acontecem nos meses mais quentes, de setembro a novembro, quando a umidade atinge valores abaixo de 55%. A temperatura do ar apresenta variações médias mensais entre 24 °C e 28 °C [23].

B. Modelagem Numérica

A metodologia adotada neste trabalho está baseada em [19]. Assim, o procedimento para a geração das séries sintéticas consiste em três etapas:

- processamento dos dados de irradiância global disponíveis para a localidade de interesse classificando-os em três diferentes condições de nebulosidade: céu claro (sem nuvens), céu parcialmente nublado e céu totalmente encoberto;
- determinação das matrizes de transição de probabilidade (MTP) para cada uma das três condições de tempo indicada na etapa anterior;
- aplicação do método da transformada inversa baseado em cadeias de Markov a tempo discreto para síntese de uma série temporal de irradiância global.

1) *Classificação das Condições de Nebulosidade*: A classificação dos dados para as diferentes condições de nebulosidade foi feita com base no índice de céu claro (k_t). O índice k_t é definido como a razão entre a irradiância global medida pela estação meteorológica e a irradiância global em condições de céu claro no mesmo local (eq. 1).

$$k_{t,i} = \frac{I_i}{I_{i,\text{clear}}} \quad (1)$$

onde $k_{t,i}$ é o índice de céu claro e I_i é a irradiância solar incidente na superfície num dado instante i do dia. A irradiância solar incidente na superfície para condições de céu sem nuvens (céu claro) no mesmo instante i está representado por $I_{i,\text{clear}}$.

A diferença entre a metodologia empregada neste estudo e a empregado por [19] reside na determinação da irradiação solar em condição de céu claro realizada nesta etapa. Os valores de $I_{i,\text{clear}}$ foram estimados com uso do modelo proposto em [28]. O modelo de Dumortier baseia-se no coeficiente de turbidez de Linke. Este é parâmetro conveniente para simular os processos radiativos associados à espessura óptica da atmosfera do vapor de água e de aerossóis atmosféricos [29] [30]. Quanto maior o valor do coeficiente de turbidez, maior a atenuação da radiação na atmosfera em condições de céu sem nebulosidade. Outros modelos de estimativa de I_{clear} foram avaliados incluindo o modelo de céu claro utilizado por Hofmann *et al.* em [19], mas apresentaram desempenho inferior ao modelo de Dumortier quando comparados com valores observados em superfície em condições de céu claro, principalmente no início e fim do dia.

A condição predominante de tempo para um dia específico resulta do padrão temporal característico dos valores de $k_{t,i}$. Este estudo adotou o valor médio de $\overline{k_{t,i}}$ ao longo do período de um dia ($\overline{k_{t,\text{dia}}}$) e seu desvio padrão ($\widetilde{k_{t,\text{dia}}}$) que apresentou resultados consistentes em estudos anteriores. [19]. Os parâmetros foram utilizados para classificar cada dia do período de estudo entre três classes de condição de nebulosidade – céu claro, parcialmente nublado, totalmente encoberto. A classificação foi feita de acordo com a Tabela I.

$$\overline{k_{t,\text{dia}}} = \frac{1}{n} \sum_{i=1}^{24} k_{t,i} \quad (2)$$

$$\widetilde{k_{t,\text{dia}}} = \frac{1}{n} \sum_{i=1}^{24} |k_{t,i} - k_{t,i-1}| \quad (3)$$

onde n é o número de horas em que a irradiância solar global foi positiva, i.e., quando o Sol está acima do horizonte.

TABELA I

LIMITES PARA A CLASSIFICAÇÃO DO DA CONDIÇÃO NEBULOSIDADE DIÁRIA CONSIDERANDO VALORES MÉDIOS ($\overline{k_{t,\text{DIA}}}$) DO ÍNDICE DE CÉU CLARO AO LONGO DO DIA ([19]).

Classificação de nebulosidade diária	Critério estabelecido
Encoberto	$0,6 - \overline{k_{t,\text{dia}}} > \widetilde{k_{t,\text{dia}}}$
Parcialmente Nublado	$0,72 - 0,8\overline{k_{t,\text{dia}}} \geq \widetilde{k_{t,\text{dia}}}$
Claro (sem nuvens)	Fora dos limites anteriores

2) *Matrizes de Transição*: Seja o processo estocástico $\{k_{t,i} = k_t(i), i = 0, 1, 2, \dots\}$ representando o conjunto de valores do índice de céu claro (k_t) correspondentes à instantes de tempo discretos (i). Consideraremos que $k_{t,i}$ assume valores possíveis num conjunto E , finito ou infinito enumerável, denominado *conjunto de estados* ou *espaço de estados*. Assim, a afirmação $k_{t,i} = k_t(i) = m$ indica que o processo estocástico está no estado m no instante de tempo i . Pode-se considerar que a transição de estados da variável estocástica $k_{t,i}$ representa uma cadeia de Markov se atende a eq. (4).

$$P[k_{t,i+1} = j | k_{t,i} = m, k_{t,i-1} = m_{i-1}, \dots, k_{t,0} = m_0] = P[k_{t,i+1} = j | k_{t,i} = m] = p_{jm} \quad (4)$$

onde p_{jm} representa a probabilidade de que a variável aleatória $k_{t,i}$ passe do estado m no instante i para o estado j no instante $i + 1$. Em resumo, a eq. (4) indica que a probabilidade de ocorrer o estado j no instante $i+1$ ($k_{t,i+1} = j$) depende somente de estado presente no instante i , e independe de toda a história anterior de valores assumidos (estados $k_{t,0}, k_{t,1}, \dots, k_{t,n-1}$) [20].

As probabilidades de alteração de estado da variável aleatória $k_{t,i}$ podem ser representadas através de uma matriz de transição de estados como representada na eq. (5).

$$\begin{pmatrix} p_{00} & \dots & p_{0j} \\ \vdots & \ddots & \vdots \\ p_{j,0} & \dots & p_{jj} \end{pmatrix} \quad (5)$$

Neste estudo, uma matriz de probabilidade de transição (MPT) foi criada para cada classe de condição de nebulosidade. A MPT dispõe da informação de quão provável será a mudança de valor (estado de nebulosidade) de k_t num instante i para outro valor (estado) no instante seguinte $i+1$. O cálculo das probabilidades de transição em cada classe de nebulosidade foi realizado a partir da contabilização das frequências de transição ocorrida na base de dados de irradiância solar incidente observados durante período de dois anos em cada uma das localidades do estudo. Em suma, a MPT contém as probabilidades de mudança de um valor específico de $k_{t,i}$ para um valor $k_{t,i+1}$ para a condição de nebulosidade específica a que corresponde a MPT em questão.

3) *Geração das Séries Temporais*: A geração de séries temporais de k_t com resolução temporal de um minuto foi realizada pelo Método da Transformada Inversa baseada em Cadeias de Markov a tempo discreto. A MPT utilizada para realização do processo é escolhida em função da classificação da condição de nebulosidade do dia em estudo. A base de dados horários (um valor de irradiância solar a cada 60 minutos) para a localidade de estudo é utilizada como parâmetro de entrada para geração da série sintética com 1-minuto de resolução.

Uma vez identificado a classificação de nebulosidade, o procedimento para geração da série tem início com o k_t horário correspondente ao nascer do Sol. Em seguida sorteia-se um número aleatório entre 0 e 1 para estabelecer o valor limite da probabilidade de transição de estados da variável estocástica k_t . O estado futuro j no instante $(i + 1)$ é determinado com base na soma das probabilidades de transição do estado presente m no instante i até que o valor limite seja alcançado ou superado [19].

A partir da série temporal de $k_{t,i}$, os valores de irradiância solar incidente na superfície foram calculados com a eq. (6).

$$I_{\text{sintético},i} = k_{t,i} \cdot I_{\text{clear},i} \quad (6)$$

onde $I_{\text{sintético},i}$ representa o valor de irradiância solar incidente na superfície no instante i da série temporal gerada sinteticamente. O procedimento é repetido até completar uma série de 60 valores que compõe o período de uma hora entre dois valores observados na localidade de estudo. O procedimento só seguirá para a determinação de estimativas para a hora seguinte se a diferença δ , obtida conforme descrito na eq. (7) seja inferior a uma precisão desejada. Neste estudo o valor da diferença $\delta = 0,05$ foi adotado como critério de parada do processo iterativo.

$$\delta = \left| \frac{\overline{I_{\text{hora}}}}{\sum_{i=1}^{60} I_{\text{sintético},i}} - 1 \right| < 0,05 \quad (7)$$

onde $\overline{I_{\text{hora}}}$ representa o valor médio entre as irradiâncias medidas no início e fim do intervalo de uma hora h observado na localidade em estudo.

C. Validação da Metodologia

A geração das matrizes MPT foi realizada utilizando os dados de irradiância global observados nas estações localizadas

em São Martinho da Serra/RS e Petrolina/PE ao longo dos anos de 2013 e 2014.

Os dados horários observados no ano de 2015 foram utilizados para gerar a série sintética de irradiância solar na superfície nas duas localidades. A série sintética gerada para 2015 foi comparada com a série temporal observada em superfície para a validação da metodologia. Os índices estatísticos utilizados na validação estão descritos a seguir.

1) *Média dos Gradientes de Irradiância Solar*: Com o intuito de verificar se a série sintética de irradiância solar apresenta uma variabilidade similar à série observada na localidade de estudo, comparou-se o valor médio dos gradientes da irradiância solar calculado conforme descrito na eq. (9). O gradiente (ou rampa) de irradiância solar é definido como a diferença absoluta entre o valor de irradiância global em minutos consecutivos (8).

$$\begin{aligned} g_{\text{sintético},i} &= |I_{\text{sintético},i} - I_{\text{sintético},i-1}| \\ g_{\text{obs},i} &= |I_{\text{obs},i} - I_{\text{obs},i-1}| \end{aligned} \quad (8)$$

onde $g_{\text{sintético},i}$ e $g_{\text{obs},i}$ representam os gradientes de irradiância solar para o instante i nas séries sintética e observada na localidade de estudo, respectivamente.

$$\begin{aligned} \text{avg}_{\text{sintético}} &= \frac{1}{n-1} \sum_{i=1}^n |I_{\text{sintético},i} - I_{\text{sintético},i-1}| \\ \text{avg}_{\text{obs}} &= \frac{1}{n-1} \sum_{i=1}^n |I_{\text{obs},i} - I_{\text{obs},i-1}| \end{aligned} \quad (9)$$

onde n representa o número de registros, $I_{\text{sintético},i}$ e $I_{\text{obs},i}$ representam a irradiância global nas séries sintética e observada, respectivamente, nas estações de superfície para os instantes i e $(i-1)$.

2) *Frequência de Gradientes*: Além de comparar o valor médio dos gradientes de irradiância solar, calculou-se o desvio entre as frequências dos gradientes de irradiância solar nas duas séries temporais ao longo do período de validação da metodologia. Os gradientes de irradiância solar foram classificados em classes de largura de 1 W/m^2 e a frequência em cada classe foi contabilizada. Para realizar a comparação entre as frequências dos gradientes foi utilizada a raiz do desvio quadrático médio ($RMSD_g$) entre as frequências dos gradientes de irradiância global em cada classe conforme descrito na eq. (10).

$$RMSD_g = \sqrt{\frac{1}{n_r} \sum_{r=1}^{n_r} |f_{\text{sintético},r} - f_{\text{obs},r}|^2} \quad (10)$$

onde n_r o número total de classes de gradientes, $f_{\text{sintético},r}$ e $f_{\text{obs},r}$ são as frequências dos gradientes de irradiância na classe r nas séries sintética e observada nas localidades de estudo, respectivamente. Valores de $RMSD_g$ próximos a zero indicam que a metodologia produziu séries sintéticas com frequência de gradientes (rampas) de irradiância solar similar à observada nas localidades de estudo.

3) *Comparação da Distribuição de Frequência da Irradiância Solar na Superfície*: O teste de Kolmogorov-Smirnov, ou teste KS pode ser utilizado para verificar a hipótese de que as séries sintética e observada de irradiância solar apresentam distribuições de probabilidade cumulativa similares. O teste KS compara a máxima diferença absoluta, D_n , entre as funções de distribuição de probabilidade cumulativa das duas séries temporais com um valor crítico, V_c para um dado nível de significância estatística (eq. 11) [31].

$$D_n = \max(|F_{\text{sintético}}(x) - F_{\text{obs}}(x)|)$$

$$V_c = \frac{1,63}{\sqrt{n}} \text{ para nível de significância de } 0,01 \quad (11)$$

onde $F_{\text{sintético}}(x)$ é a função de distribuição de probabilidade cumulativa na série temporal sintética e $F_{\text{obs}}(x)$ é a função de distribuição de probabilidade cumulativa assumida pela série observada na localidade de estudo. A hipótese de que as duas séries possuem a mesma distribuição de frequências é verdadeira quando D_n é menor que V_c .

III. RESULTADOS E DISCUSSÃO

A Tabela 2 apresenta um fragmento da MPT para a condição de céu totalmente encoberto por nuvens com os valores de probabilidade de mudança de estado do k_t obtida para a localidade de São Martinho da Serra/RS. A título de exemplo, considerando a Tabela II, a probabilidade do valor de k_t mudar de 0,02 para 0,03 no intervalo de 1 minuto é 9,26%; a probabilidade de manter o valor 0,02 é 83,53%; e a probabilidade de mudar de 0,02 para 0,01 é de 6,72%.

TABELA II

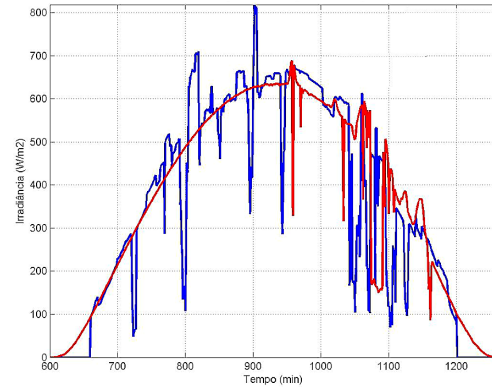
FRAGMENTO DA MPT PARA A CONDIÇÃO DE DIA COM CÉU ENCOBERTO EM SÃO MARTINHO DA SERRA/RS. A PROBABILIDADE ESTÁ APRESENTADA EM VALORES PORCENTUAIS "%".

$k_{t,i} k_{t,i+1}$	0	0,01	0,02	0,03	0,04	...
0	88,56	11,32	1,35	0	0	...
0,01	3,03	89,74	7,07	0,155	0	...
0,02	0	6,72	83,53	9,26	0,34	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

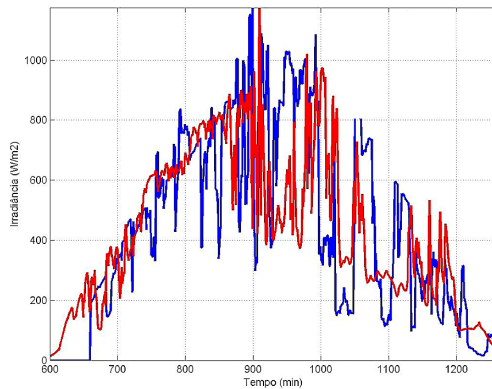
As MPTs determinadas para a geração das séries sintéticas de irradiância solar em condições de céu nublado ou parcialmente nublado são mais complexas que obtidas para às condições de céu claro, devido à grande variabilidade das condições atmosféricas nos dois primeiros casos que se refletem na variabilidade dos valores de k_t .

As Figs. 1 e 2 permitem a comparação das séries sintética e observada para as três condições de céu em São Martinho da Serra/RS e Petrolina/PE, respectivamente. Embora a evolução temporal das séries sintéticas não apresente simultaneidade com a evolução temporal das séries observadas, a variabilidade dos valores apresentada por elas é similar às encontradas nas séries observadas, qualquer que seja a condição de nebulosidade no dia em estudo. Como exemplo, o valor médio do gradiente da irradiância solar foi $avg_{\text{sintético}} = 60W/m^2$ na série sintética (Fig.1b), e $avg_{\text{obs}} = 54W/m^2$ na série

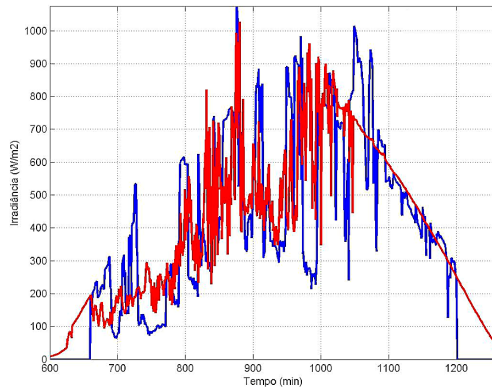
observada para o dia 03/09/2015 (parcialmente nublado em São Martinho da Serra).



(a)



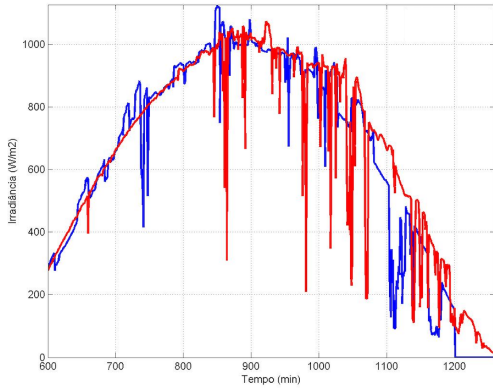
(b)



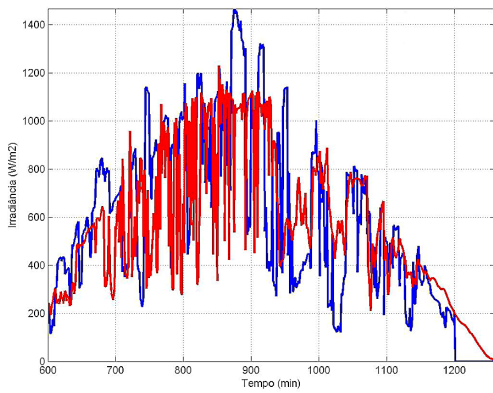
(c)

Fig. 1. Séries sintética (linha azul) e observada (linha vermelha) da irradiância global na superfície em três dias distintos em São Martinho da Serra/RS. (a) Condição de céu claro (19/05/2015). (b) Condição de céu parcialmente nublado (03/09/2015). (c) Condição de céu encoberto (25/03/2015).

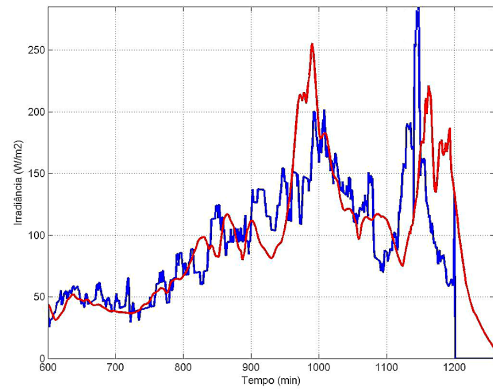
A frequência dos gradientes de irradiância solar no ano de 2015 estão apresentadas na Fig. 3. Observe que a série sintética consegue reproduzir a distribuição de frequência de rampas superiores a $10 W/m^2$ observada em ambas estações de superfície. No entanto, a metodologia adotada produziu séries



(a)



(b)

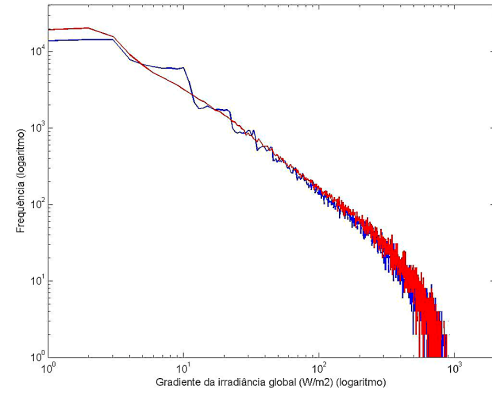


(c)

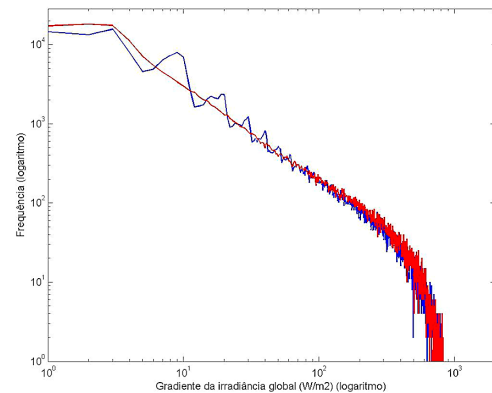
Fig. 2. Séries sintética (linha azul) e observada (linha vermelha) da irradiação global na superfície em três dias distintos em Petrolina/PE. (a) Condição de céu claro (01/01/2015). (b) Condição de céu parcialmente nublado (04/02/2015). (c) Condição de céu encoberto (25/03/2015).

sintéticas que subestimam, na ordem de 25%, a frequência de gradientes de irradiação inferiores a 5 W/m^2 .

O desvio entre as frequências de gradientes, $RMSD_g$, representa uma medida de dispersão entre valores da frequência de rampas de irradiação solar na superfície nas séries sintéticas e observadas na estação de coleta de dados. Baixos valores de $RMSD_g$ indicam que a série sintética reproduz



(a)



(b)

Fig. 3. Frequência de gradientes da irradiação global em (a) São Martinho da Serra/RS e em (b) Petrolina/PE. A linha azul corresponde a frequência de gradientes na série sintética gerada com a metodologia proposta neste trabalho. A linha vermelho representa a frequência de gradientes na série de dados observados nas estações.

a ocorrência de gradientes com a mesma frequência observada na localidade de estudo. Os resultados alcançados para $RMSD_g$ neste estudo foram inferiores aos encontrados por [19] para a cidade de Lindenberg na Alemanha utilizando metodologia similar à empregada neste estudo e outros modelos desenvolvidos por [32] [33], conforme listado na Tabela III.

TABELA III
DESVIO $RMSD_g$ PARA AS FREQUÊNCIAS DOS GRADIENTES DE IRRADIÂNCIA SOLAR NAS SÉRIES SINTÉTICA E OBSERVADA.

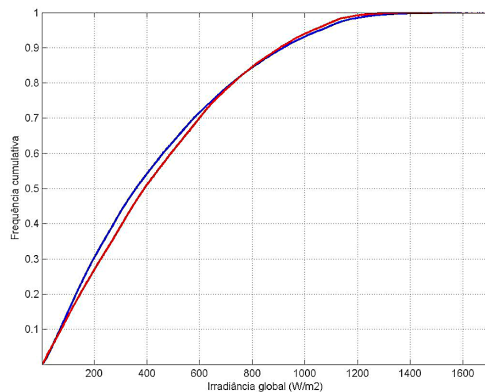
Metodologia adotada	$RMSD_g$	Localidade de Estudo
Aguiar&Collares-Pereira [19]	8131	Linderberg/Alemanha
Skarveit&Olseith [19]	4758	Linderberg/Alemanha
Hoffmann&Seckmeyer [19]	2787	Linderberg/Alemanha
Soares e Martins	470	S.Martinho Serra/RS
Soares e Martins	381	Petrolina/PE

A Fig. 4 apresenta as curvas de distribuição de probabilidade cumulativa dos valores de irradiação solar global incidente na superfície presentes nas séries sintética e observada nas duas localidades. O teste Kolmogorov-Smirnov foi utilizado para

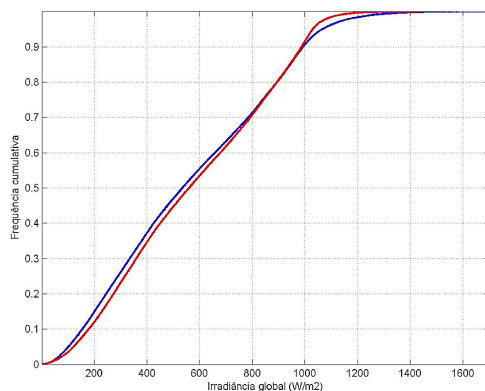
verificar a hipótese de que as séries sintéticas e observadas em cada localidade apresentam distribuição de probabilidade similares. O cálculo da máxima diferença (eq. 11) mostrou-se menor que o valor crítico ϵ , portanto, a hipótese não foi rejeitada com nível de confiança igual a 99%.

Pode-se observar na Fig. 4 que as curvas de distribuição de probabilidades das séries sintéticas e observadas são muito próximas nas duas localidades. Em São Martinho da Serra (Fig. 4a), as maiores discrepâncias, da ordem de 0,05, ocorrem entre 200 e 400 W/m^2 . A Fig. 4b mostra que as diferenças entre as duas distribuições de probabilidade cumulativa ocorrem para valores de irradiância solar superiores a 1000 W/m^2 .

Um fator que influenciou positivamente foi a seleção do modelo que melhor reproduziu a irradiação solar global no território brasileiro em condições de céu claro. Outro aspecto importante foi a determinação de MPTs específicas para cada localidade de estudo. Hoffmann *et. al* [19] construíram MPTs com base em observações realizadas em localidades distintas da Alemanha.



(a)



(b)

Fig. 4. Frequência de distribuição de probabilidade cumulativa das séries de dados de irradiância solar observada (linha vermelha) e sintética (linha azul) para o ano 2015 nas localidades: (a) São Martinho da Serra/RS e (b) Petrolina/PE.

IV. CONCLUSÃO

Este trabalho apresentou uma metodologia que emprega o método da Transformada Inversa baseada em cadeias de

Markov para produzir séries temporais sintéticas de irradiância global na superfície, com resolução de um minuto. A metodologia empregada está baseada em [19] em estudo desenvolvido com dados observados na França e Alemanha. Em razão das características típicas associadas à localização tropical e a climatologia de aerossóis no território brasileiro, foi necessário realizar adaptações para otimização dos resultados. A principal adequação foi identificação e implementação de um modelo numérico que reduziu incertezas das estimativas de irradiação solar global na superfície em condições de céu claro para o cálculo de $k_{t,i}$. A metodologia empregada mostrou-se eficiente com base em comparação estatística com dados observados em duas localidades que apresentam características climáticas distintas.

No entanto, é importante ressaltar que o objetivo da metodologia não é a previsão da irradiação solar, mas produzir séries sintéticas que apresentam distribuições de probabilidade similares às observadas em superfície. Os resultados alcançados foram superiores ao estudo desenvolvido por Hofmann *et al.* [19] uma vez que o procedimento de validação indicou menores valores de $RMSD_g$. O teste de Kolmogorov-Smirnov mostrou que as séries sintéticas e observadas apresentam distribuição de probabilidades similares tanto em São Martinho da Serra (RS) quanto em Petrolina (PE). A frequência de ocorrência de gradientes (ou rampas) de irradiância solar também mostrou similaridade nas séries sintéticas e observadas, mesmo que a ocorrência das rampas tenham ocorrido em momentos distintos do dia.

A metodologia empregada no estudo apresenta implementação simples e baixa demanda computacional não requerendo uso de computadores com elevada capacidade de processamento. O código computacional pode ser desenvolvido em qualquer linguagem de programação incluindo Python ou plataformas Matlab, Octave e similares.

Os resultados alcançados neste trabalho podem dar suporte às demandas de concessionárias, geradoras e distribuidoras de energia elétrica, e ao Operador Nacional do Sistema Elétrico ao reproduzir com confiança estatística os padrões de variabilidade da irradiância solar de minuto a minuto a partir de uma base de dados horários. A aplicação da metodologia possibilita o desenvolvimento de estudos voltados para o planejamento de operação de plantas fotovoltaicas e a contribuição destas para o sistema elétrico com um resolução temporal fina suficiente para acompanhar a influência das variações naturais da irradiância solar incidente na superfície.

AGRADECIMENTOS

Os autores agradecem ao Lab. de Modelagem e Estudos de Recursos Renováveis de Energia do Instituto Nacional de Pesquisas Espaciais (LABREN/INPE) pela cessão dos dados radiométricos. Agradecimentos são devidos ao CNPq pelo apoio financeiro dos autores. Os autores agradecem ao suporte institucional e financeiro do Instituto Nacional de Ciência e Tecnologia para Mudanças Climáticas (INCT-MC), através dos processos CNPq 573797/2008-0 e FAPESP 2008/57719-9.

REFERÊNCIAS

- [1] E. B. Pereira, F. R. Martins, A. R. Gonçalves, R. S. Costa, F. J. L. de Lima, R. Rütger, S. L. de Abreu, G. M. Tiepolo, S. V. Pereira, and J. G. de Souza, *Atlas Brasileiro de Energia Solar*, 2nd ed., INPE, Ed. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2017.
- [2] A. Bess Alcantara Amaral, A. L. Zambelli Loyola Gonzaga Mendonça, A. Araujo Martins Resende, and E. Eduardo Rego, "Solar energy and distributed generation: 2015, a year of inflection in Brazil?" *IEEE Latin America Transactions*, vol. 14, no. 8, pp. 3731–3737, Aug 2016.
- [3] EPE, "Análise da Inserção da Geração Solar na Matriz Elétrica Brasileira," Empresa Brasileira de Pesquisa Energética, Rio de Janeiro, Tech. Rep., 2012.
- [4] I. Souza Machado, B. Soares Moreira Cesar Borba, and R. Silva Maciel, "Modeling distributed pv market and its impacts on distribution system: A Brazilian case study," *IEEE Latin America Transactions*, vol. 14, no. 11, pp. 4520–4526, Nov 2016.
- [5] D. Heinemann, E. Lorenz, and M. Girodo, *Forecasting of solar radiation: Solar energy resource management for electricity generation from local level to global scale*. Hauppauge, NY, USA: Nova Science Publishers, 2006.
- [6] 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.
- [7] E. W. Luiz, F. R. Martins, R. S. Costa, and E. B. Pereira, "Comparison of methodologies for cloud cover estimation in Brazil - a case study," *Energy for Sustainable Development*, vol. 43, pp. 15 – 22, 2018.
- [8] E. W. Luiz, F. R. Martins, A. R. Gonçalves, and E. B. Pereira, "Analysis of intra-day solar irradiance variability in different Brazilian climate zones," *Solar Energy*, vol. 167, pp. 210 – 219, 2018.
- [9] C. Voyant, G. Nottton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Foulloy, "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105, pp. 569 – 582, 2017.
- [10] R. S. A. Palharini and D. A. Vila, "Climatological behavior of precipitating clouds in the northeast region of Brazil," *Advances in Meteorology*, vol. 2017, pp. 1–12, May 2017.
- [11] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, D. Renné, and T. E. Hoff, "Validation of short and medium term operational solar radiation forecasts in the us," *Solar Energy*, vol. 84, no. 12, pp. 2161 – 2172, 2010.
- [12] F. R. Martins and E. B. Pereira, "Estudo comparativo da confiabilidade de estimativas de irradiação solar para o sudeste brasileiro obtidas a partir de dados de satélite e por interpolação/extrapolação de dados de superfície," *Revista Brasileira de Geofísica*, 2011.
- [13] G. Reikard, "Predicting solar radiation at high resolutions: A comparison of time series forecasts," *Solar Energy*, vol. 83, no. 3, pp. 342 – 349, 2009.
- [14] F. J. L. Lima, F. R. Martins, E. B. Pereira, E. Lorenz, and D. Heinemann, "Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks," *Renewable Energy*, vol. 87, pp. 807–818, 2016.
- [15] P. F. Jiménez-Pérez and L. Mora-López, "Modeling and forecasting hourly global solar radiation using clustering and classification techniques," *Solar Energy*, vol. 135, pp. 682 – 691, 2016.
- [16] J. Remund and S. Müller, "Solar radiation and uncertainty information of meteoronorm 7," in *Proceedings of XXI International Photovoltaic Science and Engineering Conference*, Fukuoka, Japan, 2011.
- [17] T. Cebecauer and M. Suri, "Typical meteorological year data: Solargis approach," *Energy Procedia*, vol. 69, pp. 1958 – 1969, 2015, international Conference on Concentrating Solar Power and Chemical Energy Systems, SolarPACES 2014.
- [18] S. Wilcox and W. Marion, "Users manual for tmy3 data sets," National Renewable Energy Laboratory, Golden, USA, Tech. Rep. NREL/TP-581-43156, 2007.
- [19] C. C. R. M. Martin Hofmann, Stefan Riechelmann and G. Seckmeyer, "Improved synthesis of global irradiance with one-minute resolution for pv system simulations," *International Journal of Photoenergy*, vol. 24, pp. 1 – 10, 2014.
- [20] R. Howard, *Dynamic Probabilistic Systems*, 2nd ed. New York, NY, USA: Dover Publications, 2007, vol. 1.
- [21] A. Driemel and *et al.*, "Baseline surface radiation network (bsrn): structure and data description (1992–2017)," *Earth Syst. Sci. Data*, vol. 10, pp. 1491–1501, 2018.
- [22] F. R. Martins, R. Guarnieri, R. C. Chagas, N. S. L. M., P. E. B., A. E., and Thomaz C., "Projeto Sonda – Rede Nacional De Estações Para Coleta De Dados Meteorológicos Aplicados Ao Setor De Energia," in *Congresso Brasileiro de Energia Solar - CBENS I*, 2007.
- [23] INMET, *Normais Climatológicas do Brasil*, Instituto Nacional de Meteorologia, Brasília, Brasil, 2019. [Online]. Available: <http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas>
- [24] I. F. A. Cavalcanti, N. J. Ferreira, M. A. F. d. S. Dias, and M. G. A. Justi, *Tempo e Clima no Brasil*, 1st ed., Oficina de Textos, Ed., São Paulo, 2009.
- [25] M. T. Kayano and R. V. Andreoli, "Clima da região Nordeste do Brasil," in *Tempo e Clima no Brasil*, 1st ed., Off. Ed. São Paulo: Oficina de Texto, 2009, ch. 14, p. 464.
- [26] V. E. Kousky, "Frontal Influences on Northeast Brazil," *Monthly Weather Review*, 1979.
- [27] L. C. B. Molion and S. d. O. Bernardo, "Uma Revisão da Dinâmica das Chuvas no Nordeste Brasileiro," *Revista Brasileira de Meteorologia*, 2002.
- [28] D. Dumortier, "The satellight model of turbidity variations in europe," Report for the Sixth. Satellight Meeting, Tech. Rep., 1998.
- [29] L. Diabaté, J. Remund, and L. Wald, "Linke turbidity factors for several sites in africa," *Solar Energy*, vol. 75, no. 2, pp. 111 – 119, 2003.
- [30] F. Kasten, "The linke turbidity factor based on improved values of the integral rayleigh optical thickness," *Solar Energy*, vol. 56, no. 3, pp. 239 – 244, 1996.
- [31] M. P. Pes, E. B. Pereira, J. A. Marengo, F. R. Martins, D. Heinemann, and M. Schmidt, "Climate trends on the extreme winds in Brazil," *Renewable Energy*, vol. 109, pp. 110 – 120, 2017.
- [32] A. Skartveit and J. Olseth, "The probability density and autocorrelation of short-term global and beam irradiance," *Solar Energy*, vol. 49, no. 6, pp. 477 – 487, 1992.
- [33] R. Aguiar and M. Collares-Pereira, "Tag: A time-dependent, autoregressive, gaussian model for generating synthetic hourly radiation," *Solar Energy*, vol. 49, no. 3, pp. 167 – 174, 1992.



Fernando Ramos Martins Docente do Instituto do Mar, Universidade Federal de São Paulo, campus Baixada Santista. Bacharel em Física (IF/USP), Mestre em Tecnologia Nuclear (IPEN/CNEN) e Doutor em Geofísica Espacial (INPE). A área de atuação principal está focada no uso da modelagem numérica para avaliação e previsão de recursos renováveis de energia, principalmente, solar e eólica. Também contribui com análise de dados e séries temporais com base em metodologias de aprendizado de máquina e estatística multivariada.



Francisco José Lopes de Lima é Doutor em Meteorologia pelo Instituto Nacional de Pesquisas Espaciais (INPE). Realizou pós-doutorado na Universidade Federal de São Paulo e na Universidade de Oldenburgo (Alemanha) em modelagem numérica e previsão de recursos energéticos renováveis. Atualmente ocupa uma posição de pós-doutor com atividades em modelagem computacional, física, meteorologia dinâmica.



Thaiane Gamarra Soares Bacharel em Ciência e Tecnologia do Mar e Bacharel em Engenharia de Petróleo pela Universidade Federal de São Paulo, campus Baixada Santista. Bolsista TT FAPESP em Projeto PIPE desenvolvido por Virtux S.A. ao longo de 2019. Membro da equipe de pesquisadores do Laboratório Interdisciplinar de Computação Aplicada (LaICA/Unifesp).