Simulação de geração multivariada de renováveis via cadeias de Markov

Gustavo de Andrade Melo
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900
gustavo.melo.rio@gmail.com

Paula Medina Maçaira
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900
paulamacaira@puc-rio.br

Fernando Luiz Cyrino Oliveira
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900
cyrino@puc-rio.br

Alessandro Soares da Silva Junior
PSR – Energy Consulting and Analytics
Praia de Botafogo, 370 / 1º andar, Botafogo - Rio de Janeiro, RJ – Brasil - 22250-040
alessandro@psr-inc.com

RESUMO

O aumento da participação de fontes de energias renováveis variáveis (ERVs) na matriz elétrica brasileira traz diversos desafios para o planejamento e operação do Sistema Elétrico Brasileiro (SEB), devido à estocasticidade presente na ERVs. Tais desafios envolvem a modelagem e simulação dos processos intermitentes de geração e, neste contexto, este trabalho tem como objetivo simular cenários de geração multivariada de uma usina híbrida do Nordeste brasileiro para o horizonte de um ano, considerando as fontes eólica e solar. Foram utilizados dados técnicos e de localização da usina, e a metodologia de simulação proposta baseia-se na discretização das séries temporais históricas para geração de cenários a partir de Cadeias de Markov. Os resultados obtidos são promissores: os cenários simulados reproduziram satisfatoriamente as características dos dados históricos de geração eólica e solar.[1]

PALAVRAS CHAVE. Fontes Renováveis; Modelagem e Simulação; Cadeias de Markov.

Tópicos. SIM – Simulação; EST&MP – Estatística e Modelos Probabilísticos; EN&PG – PO na Área de Energia, Petróleo e Gás.

ABSTRACT

The increased participation of variable renewable energy sources (RES) in the Brazilian electricity matrix brings several challenges to the planning and operation of the Brazilian Electric System (BES), due to the stochasticity present in RES. Such challenges involve the modeling and simulation of intermittent generation processes and, in this context, this work aims to simulate multivariate generation scenarios of a hybrid plant in Northeastern Brazil for a one-year horizon, considering wind and PV sources. Plant technical and location data were used, and the simulation methodology is based on the discretization of historical time series to generate scenarios from

[1] Nota de agradecimento: os autores agradecem à equipe da PSR pelo imprescindível suporte prestado no desenvolvimento deste trabalho.

Markov Chains. The results obtained are promising: the simulated scenarios satisfactorily reproduced the characteristics of historical data on wind and PV generation.

KEYWORDS. Renewable Sources; Modeling and Simulation; Markov Chain.

Paper topics. Simulation; Statistics and Probabilistic Models; Operational Research applied to the Energy, Oil and Gas Industry.

1. Introdução

As crescentes preocupações relacionadas às mudanças climáticas e à independência energética têm apoiado os desenvolvimentos tecnológicos e regulatórios necessários para expandir o mix de energia em todo o mundo, com particular ênfase nas fontes de energia renováveis. Segundo a *International Renewable Energy Agency* [IRENA, 2021], a despeito da pandemia de COVID-19, mais de 260 GW de capacidade de energia renovável foi adicionada à matriz elétrica global em 2020, superando a expansão em 2019 em cerca de 50%. Destacam-se as fontes eólica e fotovoltaica (FV), que juntas respondem por 91% do crescimento. Em contrapartida, a adição de combustíveis fósseis na matriz global caiu de 64 GW em 2019 para 60 GW em 2020 [IRENA, 2021], evidenciando uma tendência contínua de redução da expansão dos combustíveis fósseis.

Seguindo a tendência mundial, o Brasil vem investindo substancialmente, nas últimas duas décadas, no aumento da participação de fontes de energia renováveis variáveis (ERVs) em sua matriz elétrica. Embora as hidrelétricas ainda correspondam por 60% da capacidade instalada no país, projeções do Plano Decenal de Expansão de Energia do Ministério de Minas e Energia [MME, 2020] apontam que, em 2029, a fonte hidráulica reduzirá sua participação para 42% da matriz elétrica. Já a FV saltará dos atuais 2% para 9%, enquanto a eólica, de 9% para 16%.

No entanto, devido à variabilidade e às incertezas inerentes às fontes ERVs, sua integração em larga escala impacta fortemente o planejamento e operação do Sistema Elétrico Brasileiro (SEB), representando um desafio para o despacho da geração de energia e exigindo grandes mudanças na sua gestão e otimização [Esteves et al, 2019]. Na prática, segundo Pinson [2013], tais desafios requerem a modelagem e simulação dos processos estocásticos de geração de energia renovável em escalas temporais e espaciais, para serem posteriormente usadas como base para a tomada de decisões nos setores público e privado.

Neste contexto, as técnicas probabilísticas (e.g., analíticas e de simulação) são os meios mais comumente utilizados para modelar a geração de fontes renováveis. Segundo Almutairi et al. [2016], extensas medições históricas de variáveis ligadas a fontes intermitentes, como velocidade do vento e radiação solar, são necessárias para modelar o comportamento aleatório de tais variáveis em um determinado local. No entanto, não são raras as vezes em que os dados disponíveis são insuficientes, destacando a necessidade do desenvolvimento de técnicas confiáveis de simulação.

Dessa forma, nos últimos anos, um volume considerável de pesquisas utilizando dados de diversos países tem sido direcionado à modelagem e simulação de geração de fontes renováveis intermitentes. Muitos desses trabalhos utilizam modelos estocásticos baseados em cadeias de Markov. Luo et al. [2014] aplicaram cadeias de Markov Monte Carlo (MCMC), usando amostragem de Gibbs, para simular séries temporais de energia eólica e FV na Alemanha. Almutairi et al. [2016] simularam cenários de geração eólica do Canadá, a partir de uma metodologia desenvolvida com base em cadeias de Markov. Os índices de confiabilidade obtidos pelo método proposto foram comparados com índices produzidos por modelos ARMA de Box e Jenkins [1970], frequentemente utilizados em estudos de confiabilidade, apontando a eficácia da técnica proposta. Miao et al. [2017], usando dados norteamericanos, propuseram um modelo Markoviano de geração de parque eólico, considerando o parque como uma unidade de múltiplos estados de geração. Baldioti e Souza [2018], a partir de uma metodologia combinando modelos autorregressivos periódicos (PAR(p)), Estimativa de Densidade por Kernel (KDE) e MCMC, geraram séries sintéticas de Energia Natural Afluente (ENA) para o Sudeste do Brasil. Maçaira et al. [2019] basearam-se no método desenvolvido por Almutairi et al. [2016] para simular cenários de energia eólica no Nordeste do Brasil. Os autores integraram os cenários simulados ao modelo de despacho hidrotérmico brasileiro, sem a necessidade de qualquer alteração estrutural no modelo de otimização, e obtiveram resultados satisfatórios ao reduzirem a chance de esgotamento dos reservatórios de água e de um despacho ineficaz. Ma et al. [2020] simularam a geração de energia FV de uma província chinesa combinando clusterização via K-*means* e cadeias de Markov.

Deve-se destacar que há outras linhas de pesquisa relacionadas à modelagem e simulação de geração de fontes renováveis. Dias e Borges [2014], a partir de séries hidrológicas do Brasil, aplicaram o método não paramétrico KDE para a construção de um modelo autorregressivo de vazões afluentes que não assume nenhuma distribuição de probabilidade específica. As séries sintéticas geradas pelo modelo proposto reproduziram melhor a densidade de probabilidade das séries históricas em comparação a modelos PAR(p). Schmidt et al. [2016] usaram dados brasileiros para estimar, a partir de um modelo de otimização, uma combinação ideal de expansão de capacidade de geração eólica, FV e hidrelétrica para o país, e simularam cenários de despachos das usinas por *bootstrap* nas séries de produção. Alipour et al. [2020], com base nas incertezas decorrentes da geração eólica e FV e do consumo de carga elétrica, propuseram um modelo de Rede Neural Profunda para simulação da operação elétrica com dados de 37 países europeus. Duca et al. [2021] modelaram a velocidade do vento ao longo do tempo através dos modelos dinâmicos de espaço de estado Weibull e Gamma. Os resultados, obtidos a partir de dados de diferentes regiões do Brasil, apontaram que o modelo dinâmico Gamma é competitivo com o Weibull, mais comumente usado para modelagem e previsão de velocidade de vento. Já no *software Time Series Lab* (TSL) [PSR, 2021], foi implementada uma metodologia para gerar cenários futuros de geração de renováveis como base em KDE e em redes bayesianas, as quais visam capturar os efeitos de complementaridade e correlação entre as ERVs.

É importante sublinhar a relevância de técnicas de simulação que considerem o caráter complementar entre diferentes fontes de energia intermitentes, como eólica e solar. Segundo Jurasz et al. [2021], a complementaridade temporal entre recursos renováveis resulta em uma potência combinada com menor variabilidade e intermitência, diminuindo a demanda por armazenamento e suavizando a operação do sistema elétrico. Buscando abordar esse caráter complementar, o objetivo deste trabalho é simular cenários de geração multivariada de renováveis para o horizonte de um ano, considerando as fontes eólica e solar. Os dados para a simulação foram obtidos a partir da combinação de parâmetros técnicos do parque híbrido Caetité, localizado no sertão da Bahia, e de dados de reanálise correspondentes à região da usina. A metodologia utilizada para gerar os cenários híbridos é baseada em Almutairi et al. [2016], que representam o processo de geração eólica por cadeias de Markov. É relevante destacar a inovação que este artigo traz, visto que, até onde se tenha conhecimento, considerando a data de conclusão desta pesquisa, nenhum trabalho havia realizado simulação de cenários de geração multivariada de renováveis com a metodologia proposta.

Com o objetivo de verificar se os cenários sintéticos reproduzem satisfatoriamente o comportamento dos dados iniciais, são avaliados fatores estatísticos, como a função de distribuição de probabilidade (FDP) e características mensais, isto é, média e desvio padrão das ERVs. Além disso, os cenários simulados foram comparados com as gerações eólica e solar do último ano da série histórica, cujos dados não foram considerados no processo de simulação.

O restante deste artigo está subdividido da seguinte forma. O capítulo 2 detalha a metodologia de simulação utilizada. A seção 3 apresenta os resultados e as discussões. Por fim, o capítulo 4 resume as conclusões e apresenta sugestões para trabalhos futuros.

2. Metodologia

Esta seção é subdividida em duas partes. Na subseção 2.1, é apresentado o processo de obtenção dos dados de geração usados na simulação. A subseção 2.2 detalha o método de simulação utilizado, baseado em cadeias de Markov.

2.1 Medições históricas

Para simular cenários futuros de energia eólica e solar, são necessários dados históricos de geração das ERVs. No entanto, muitas vezes os dados disponíveis são inexistentes ou

insuficientes, como no caso do parque híbrido considerado neste estudo. Dessa forma, para obtenção dos dados de geração foi utilizado o *software Time Series Lab* (TSL) [PSR, 2021], no qual é implementada uma metodologia com o objetivo de criar registros históricos sintéticos de geração horária de até 40 anos usando características técnicas das usinas e informações disponíveis no banco de dados de reanálise global Merra-2 [NASA, 2021]. Um resumo da metodologia do TSL [PSR, 2021] é dado a seguir. Com o *software*, obtivemos um registro histórico de 1980 a 2018 de geração correspondente ao parque híbrido Caetité (Bahia), cuja capacidade instalada é de 26,4 MW, dos quais 21,6 MW provêm da fonte eólica e 4,8 MW da FV. A base de dados obtida é composta de duas séries horárias de energia para o período mencionado, uma para fonte eólica e outra para FV.

É importante destacar, primeiramente, as vantagens em simular séries sintéticas a partir de dados medidos de energia, em comparação a dados de recursos primários, como velocidade do vento e irradiância solar. Segundo Papaefthymiou e Klöckl [2008], são duas as principais razões para simular diretamente as séries de geração: evita erros de cálculo de conversão dos recursos primários em energia; e ao discretizar as séries para realizar as simulações via cadeias de Markov, o número de estados é menor, especificamente para a série eólica, visto que a produção eólica só é observada para uma determinada faixa de valores de velocidade do vento.

No TSL [PSR, 2021], a geração eólica é calculada utilizando um modelo baseado no *Virtual Wind Farm* (VWF), desenvolvido por Staffell e Pfenninger [2016]. O primeiro passo é informar as coordenadas geográficas do local para que os dados de velocidade do vento sejam obtidos do Merra-2 [NASA, 2021]. No segundo passo, devem ser definidas as características técnicas referentes à altura e ao modelo das turbinas eólicas da usina. Então, uma função extrapola as velocidades iniciais para a altura das turbinas e, por fim, a velocidade do vento extrapolada é convertida em energia através da curva de potência correspondente ao modelo de turbina informado.

Já a geração solar é obtida a partir do método *Global Solar Energy Estimator* (GSEE), com base em Pfenninger e Staffell [2016]. O primeiro passo, de forma análoga ao processo de obtenção de geração eólica, é informar as coordenadas da usina para extrair, do Merra-2 [NASA, 2021], os dados de irradiação e temperatura necessários para o cálculo da Irradiação Horizontal Global (IHG). Então, devem ser definidas as características técnicas do parque solar, isto é, o ângulo relacionado à inclinação dos painéis solares, bem como o sistema de rastreamento instalado, para que, finalmente, seja calculada a energia FV a partir da IHG. É importante mencionar que o foco deste trabalho não está na conversão dos recursos primários em geração de energia, mas na simulação de cenários de ERVs. Para maiores detalhes do método utilizado para geração dos dados históricos, consulte TSL [PSR, 2021], Staffell e Pfenninger [2016] e Pfenninger e Staffell [2016].

Dessa forma, obtivemos, para cada fonte renovável do parque híbrido, uma série temporal com granularidade horária de geração por unidade (p.u.) de capacidade instalada, em valores no intervalo [0,1], para o período de 1980 a 2018. Para chegarmos aos valores de geração eólica e solar em MW, multiplicamos os registros de cada uma das séries históricas pela capacidade instalada referente à fonte renovável da série, isto é, 21,6 MW no caso da eólica e 4,8 MW no caso da FV. Os valores de capacidade instalada, bem como os demais parâmetros técnicos do parque foram obtidos em Renova [2016].

O gráfico da figura 1 apresenta a geração p.u. média por hora das fontes eólica e solar, a partir da base histórica obtida. Note o caráter complementar temporal entre as duas fontes, visto que a média máxima da geração eólica ocorre no período noturno, no qual a produção solar é nula. Em contrapartida, os maiores níveis médios de geração FV ocorrem em períodos do dia nos quais, geralmente, a produção eólica atinge seus menores patamares.

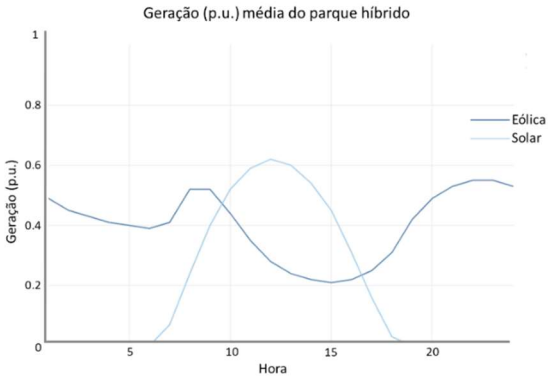


Figura 1. Geração (p.u.) média do parque híbrido. Fonte: TSL [PSR, 2021].

2.2 Simulação via cadeias de Markov

O procedimento aplicado neste estudo para simular estocasticamente os dados de energia eólica e solar é baseado em Almutairi et al. [2016]. O processo pode ser resumido de acordo com o fluxograma apresentado na figura 2 e os detalhes de cada etapa são dados em seguida.

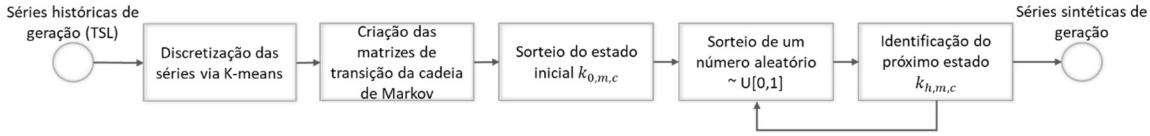


Figura 2. Processo de simulação das séries sintéticas de geração. Fonte: Almutairi et al. [2016]. Elaboração dos autores.

Após a obtenção das séries históricas de geração a partir do TSL [PSR, 2021], é necessário agrupar os dados de energia medidos em um número finito de estados. Para realizar a discretização, utilizamos a técnica de agrupamento K-*means* [MacQueen, 1967]. Na técnica, dado um número k de clusters representados por centroides iniciais, a distância de cada valor da série histórica para cada centroide é calculada e todos os valores de energia são atribuídos ao centroide mais próximo. Novos centroides de cada cluster são calculados usando a média dos dados de energia no cluster k e as distâncias são recalculadas. Esse processo é repetido até que os centroides permaneçam fixos após várias iterações. No final, os valores de energia do histórico são substituídos pelos centroides do cluster a que pertencem. No caso deste trabalho, dado que o objetivo é simular geração multivariada, a partir de dois tipos de ERVs – eólica e solar, cada cluster possui um centroide formado por dois atributos, equivalentes aos dois tipos de geração considerados.

É importante destacar que o agrupamento foi realizada nos dados mensais, isto é, com os dados de geração pertencentes ao mesmo mês considerando todo o histórico, a fim de preservar, nas séries simuladas, as características sazonais anuais de geração de ERVs. Além disso, cada conjunto de dados mensais foi agrupado de acordo com o período do dia, noturno ou diurno, com o objetivo de manter as características de produção diárias nas séries sintéticas. O período noturno compreende os horários em que a geração solar é nula; o diurno, portanto, abrange os horários nos quais a geração solar é diferente de zero.

A segunda etapa corresponde à criação das matrizes de transição da cadeia de Markov, kxk, (P_{ind}), por período, noturno e diurno, para cada mês, onde k é o número de estados de cada período calculados na etapa anterior. A probabilidade de transição $(p_{a,b})$ do estado a para o estado b, para todos os índices $1 \leq (a, b) \leq k$, pode ser calculada pela Equação 1 a seguir.

$$p_{a,b} = \frac{n_{a,b}}{\sum_k n_{a,b}} \quad (1)$$

onde $n_{a,b}$ é o número de transições de a para b. Após obter as probabilidades de transição para cada estado, é possível construir a matriz de transição para cada período mensal, conforme a Equação 2:

$$P_{ind} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,k} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k,1} & p_{k,2} & \cdots & p_{k,k} \end{bmatrix} \quad (2)$$

A partir da matriz de transição individual P_{ind}, a matriz de transição de probabilidade acumulada P_{acum} pode ser criada, de modo que sua última coluna seja igual a um para cada linha da matriz do período mensal, conforme a Equação 3:

$$P_{acum} = \begin{bmatrix} p_{1,1} & p_{1,1} + p_{1,2} & \cdots & p_{1,1} + p_{1,2} + \cdots + p_{1,k} \\ p_{2,1} & p_{2,1} + p_{2,2} & \cdots & p_{2,1} + p_{2,2} + \cdots + p_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k,1} & p_{k,1} + p_{k,2} & \cdots & p_{k,1} + p_{k,2} + \cdots + p_{k,k} \end{bmatrix} \quad (3)$$

A terceira etapa consiste em selecionar, aleatoriamente, um estado inicial $k_{0,m,c}$ para o período mensal desejado, sendo m o índice do mês e c, o índice referente ao cenário simulado. O quarto passo é gerar, via simulação de Monte Carlo, um número aleatório de uma distribuição uniforme $U \sim [0,1]$. A partir deste número aleatório e da matriz de transição acumulada da cadeia de Markov, é possível obter, na quinta etapa, o próximo estado simulado de geração multivariada. Para isso, compara-se o valor do número aleatório com os elementos da n-ésima linha da transição de probabilidade acumulada. Se o número gerado aleatoriamente for maior que a probabilidade acumulada do estado anterior e menor ou igual à probabilidade acumulada do estado subsequente, este é escolhido para representar o próximo estado simulado. Este procedimento é repetido para simular séries sintéticas horárias de geração eólica e solar para T anos e está representado na Equação 4:

$$GES_{h,p,d,m,y,c} = c_k \quad se \quad c_{k-1} < p_{unif[0,1]} \leq c_k \quad (4)$$

onde $GES_{h,p,d,m,y,s}$ é a variável que representa a geração multivariada simulada para hora h, período p, dia d, mês m, ano y ($\sum y = T$) e cenário c; c_k é o valor do estado k, constituído de duas dimensões, uma para geração eólica e outra para solar; e $p_{unif[0,1]}$ é o número aleatório obtido a partir da distribuição uniforme.

3. Resultados e discussões

Na primeira etapa da metodologia de simulação, obteve-se um número finito de estados para cada conjunto de dados mensais por período, noturno ou diurno, a partir da técnica K-*means*. É importante mencionar que o último ano do histórico, 2018, foi excluído do processo de agrupamento, a fim de avaliar os cenários simulados, comparando-os com os dados reais deste ano. Para cada série, foram analisados diversos agrupamentos, variando o número de clusters entre 2 e 25. Assim, foi escolhido, para cada série, o agrupamento que representava ao menos 98% da variabilidade dos dados. Para o período noturno dos meses, o número de clusters selecionado variou de 14 a 17. Já para o período diurno, a quantidade de clusters selecionada foi de 23 para todos os meses.

A próxima etapa consiste em gerar a matriz de transição da cadeia de Markov, construída pela frequência com a qual um estado transita para outro ou para si mesmo. A partir da matriz de transição individual P_{ind}, cria-se a matriz de transição de probabilidade acumulada P_{acum}. Nas

figuras 3 e 4 são apresentadas as matrizes de transição de probabilidade acumulada do mês de janeiro para os períodos noturno e diurno, respectivamente. Como foram obtidas 24 matrizes no total, considerando todos os meses, nos limitaremos a apresentar os exemplos apenas do mês de janeiro. Note que o número de estados do período noturno escolhido no método K-*means* foi igual a 15, e para o período diurno, igual a 23. Os centroides de cada estado para cada período estão descritos nas tabelas 1 e 2, onde GE representa o centroide da geração eólica e GS, da geração solar, ambos em MW. Além disso, no período noturno não há geração solar. Portanto, o centroide de cada estado deste período é representado apenas pela geração eólica.

Estado a	Estado b														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0,29	0,29	0,34	0,34	0,52	0,86	0,86	0,87	0,87	0,87	0,87	0,88	1	1	1
2	0	0,34	0,34	0,37	0,37	0,38	0,40	0,44	0,47	0,62	0,64	0,64	0,64	0,76	1
3	0,06	0,06	0,69	0,70	1	1	1	1	1	1	1	1	1	1	1
4	0	0	0	0,79	0,79	0,79	0,94	0,94	0,98	1	1	1	1	1	1
5	0,36	0,36	0,52	0,52	0,88	0,98	0,98	0,98	0,98	0,98	0,98	0,99	1	1	1
6	0,15	0,15	0,18	0,18	0,24	0,55	0,55	0,56	0,56	0,56	0,57	0,68	0,99	1	1
7	0	0,02	0,02	0,20	0,20	0,20	0,59	0,59	0,87	0,99	0,99	0,99	0,99	0,99	1
8	0	0,04	0,04	0,06	0,07	0,10	0,11	0,40	0,41	0,42	0,58	0,67	0,72	0,94	1
9	0	0,11	0,11	0,17	0,17	0,17	0,32	0,33	0,68	0,95	0,95	0,96	0,96	0,96	1
10	0	0,27	0,27	0,31	0,31	0,31	0,35	0,35	0,50	0,84	0,84	0,84	0,84	0,88	1
11	0,02	0,03	0,03	0,05	0,05	0,10	0,11	0,41	0,43	0,44	0,74	0,87	0,94	0,99	1
12	0,05	0,06	0,06	0,06	0,08	0,16	0,16	0,22	0,23	0,23	0,55	0,85	0,99	0,99	1
13	0,09	0,09	0,09	0,09	0,13	0,27	0,27	0,28	0,28	0,28	0,38	0,71	0,99	1	1
14	0	0,04	0,04	0,08	0,09	0,10	0,12	0,31	0,33	0,34	0,45	0,51	0,54	0,82	1
15	0	0,14	0,14	0,19	0,19	0,19	0,21	0,32	0,34	0,38	0,44	0,46	0,47	0,70	1

Figura 3. Matriz de transição de probabilidade acumulada – Janeiro/Noturno. Elaboração dos autores.

Estado a	Estado b																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	0,27	0,27	0,27	0,62	0,70	0,70	0,70	0,70	0,70	0,70	0,70	0,90	0,90	0,90	1	1	1	1	1	1	1	1	1
2	0	0,45	0,45	0,45	0,45	0,72	0,73	0,73	0,73	0,87	0,87	0,87	0,87	0,87	0,87	0,92	0,92	0,92	0,96	0,96	1	1	1
3	0	0	0,47	0,47	0,47	0,47	0,47	0,47	0,52	0,52	0,64	0,65	0,65	1	1	1	1	1	1	1	1	1	1
4	0,21	0,21	0,21	0,45	0,80	0,80	0,80	0,80	0,80	0,80	0,80	0,80	0,80	0,80	0,82	0,82	0,83	0,83	1	1	1	1	1
5	0,02	0,02	0,02	0,22	0,42	0,42	0,42	0,42	0,42	0,57	0,57	0,57	0,57	0,57	0,59	0,59	0,61	0,61	0,99	0,99	1	1	1
6	0	0,06	0,06	0,06	0,06	0,23	0,24	0,54	0,54	0,57	0,57	0,57	0,57	0,57	0,57	0,66	0,66	0,67	0,67	0,67	1	1	1
7	0	0	0	0	0	0,04	0,23	0,24	0,24	0,24	0,24	0,24	0,24	0,24	0,24	0,32	0,32	0,43	0,43	0,43	0,43	0,43	1
8	0	0	0	0	0	0,07	0,09	0,15	0,15	0,15	0,16	0,16	0,16	0,16	0,16	0,32	0,34	0,71	0,71	0,71	0,73	0,94	1
9	0	0	0,34	0,34	0,34	0,34	0,34	0,34	0,49	0,49	0,67	0,67	0,67	0,74	0,74	0,74	0,74	0,87	0,87	0,94	0,94	0,99	1
10	0	0,55	0,55	0,55	0,64	0,66	0,66	0,66	0,66	0,81	0,81	0,81	0,81	0,81	0,81	0,81	0,84	0,84	0,99	0,99	1	1	1
11	0,09	0,09	0,11	0,11	0,11	0,11	0,11	0,11	0,13	0,13	0,36	0,37	0,37	0,62	0,84	0,84	0,87	0,87	0,87	0,87	0,87	1	1
12	0,33	0,33	0,33	0,33	0,33	0,33	0,33	0,33	0,33	0,33	0,33	0,99	0,99	1	1	1	1	1	1	1	1	1	1
13	0	0	0,02	0,02	0,02	0,02	0,02	0,02	0,03	0,03	0,03	0,03	0,47	0,47	0,47	0,47	0,47	0,48	0,48	0,97	0,97	0,97	1
14	0,09	0,09	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,23	0,45	0,45	0,98	1	1	1	1	1	1	1	1	1
15	0,19	0,19	0,19	0,26	0,31	0,31	0,31	0,32	0,32	0,32	0,35	0,35	0,35	0,36	0,56	0,56	0,94	0,94	0,96	0,96	0,97	1	1
16	0	0	0	0	0	0,08	0,51	0,81	0,81	0,81	0,81	0,81	0,81	0,81	0,81	0,96	0,96	0,98	0,98	0,98	0,99	0,99	1
17	0	0	0	0,06	0,14	0,22	0,22	0,27	0,27	0,30	0,31	0,31	0,31	0,31	0,45	0,46	0,64	0,64	0,67	0,67	0,97	1	1
18	0	0	0	0	0	0	0,07	0,14	0,63	0,63	0,71	0,71	0,73	0,73	0,73	0,78	0,78	0,91	0,91	0,91	0,91	0,98	1
19	0	0,19	0,19	0,23	0,37	0,37	0,37	0,37	0,37	0,72	0,72	0,72	0,72	0,72	0,73	0,73	0,74	0,74	0,92	0,92	1	1	1
20	0	0	0,44	0,44	0,44	0,44	0,44	0,44	0,55	0,55	0,55	0,55	0,56	0,57	0,57	0,57	0,57	0,57	0,57	0,99	0,99	0,99	1
21	0	0,07	0,07	0,07	0,12	0,36	0,37	0,39	0,39	0,40	0,40	0,40	0,40	0,40	0,42	0,51	0,75	0,76	0,79	0,79	0,88	1	1
22	0	0	0	0,01	0,01	0,01	0,01	0,19	0,23	0,23	0,53	0,53	0,53	0,53	0,70	0,70	0,78	0,80	0,80	0,80	0,82	1	1
23	0	0	0	0	0	0	0,05	0,06	0,24	0,24	0,24	0,24	0,81	0,81	0,81	0,81	0,81	0,84	0,84	0,85	0,85	0,85	1

Figura 4. Matriz de transição de probabilidade acumulada – Janeiro/Diurno. Elaboração dos autores.

	Estado														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
GE (MW)	15,03	3,35	19,55	0,07	17,11	13,07	0,79	6,82	1,56	2,41	8,15	9,71	11,29	5,61	4,43

Tabela 1. Valores (MW) dos centroides por estado – Janeiro/Noturno. Elaboração dos autores.

	Estado																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
GE (MW)	1,02	0,96	8,07	0,68	0,83	5,21	14,99	8,51	10,96	0,63	5,97	0,89	17,60	4,22	3,57	9,93	3,98	11,96	0,77	12,73	3,82	7,34	17,45
GS (MW)	3,12	0,23	3,56	2,59	2,01	0,10	0,69	0,78	2,70	0,85	2,89	3,61	3,01	3,58	2,43	0,11	1,57	1,77	1,43	3,49	0,71	1,91	1,90

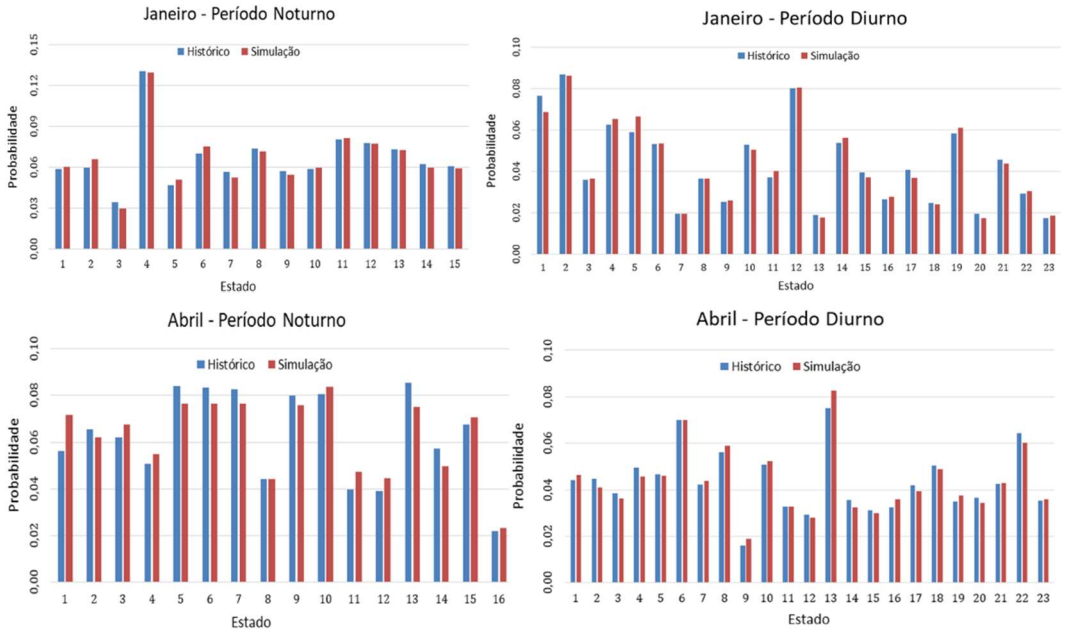
Tabela 2. Valores (MW) dos centroides por estado – Janeiro/Diruno. Elaboração dos autores.

Para simular os valores de geração multivariada por hora, é necessário, primeiro, selecionar aleatoriamente o estado inicial. Por exemplo, considerando a simulação do período diurno, seleciona-se o estado 9 (10,96 MW de geração eólica e 2,70 MW de FV) como estado inicial e, então, um valor da distribuição uniforme [0, 1] é sorteado. Suponha que seja 0,30. Como 0,30 é maior que 0 (estado 2) e menor que 0,34 (estado 3), o segundo estado sorteado é o 3 (8,07 MW de geração eólica e 3,56 MW de FV). Repetimos este procedimento até completar o horizonte de um ano de simulação. Além disso, com base no estudo de Maçaira et al. [2019] e por motivos

de convergência, todo o processo de simulação deve ser repetido 200 vezes para cada período mensal, gerando assim o mesmo número de cenários possíveis.

Para verificar se os cenários sintéticos reproduzem satisfatoriamente o comportamento dos dados iniciais, foram realizados os seguintes procedimentos: (i) avaliação das funções de distribuição de probabilidade (FDPs); (ii) avaliação das características mensais, como média e desvio padrão; (iii) e comparação dos cenários simulados com as gerações históricas do ano de 2018, cujos dados não foram utilizados no processo de simulação. A FDP mostra os valores do comportamento de distribuição de um determinado conjunto de dados. Neste trabalho é utilizado para comparar a proporção dos valores gerados através das simulações e das séries históricas.

Na figura 5, são apresentadas as FDPs das séries históricas e simuladas dos períodos noturno e diurno de janeiro, abril, julho e dezembro, representando, assim, as quatro estações do ano. Além disso, na figura 5, os valores de geração por estado podem ser observados logo abaixo das FDPs mensais. Em relação ao mês de janeiro, para maiores detalhes dos valores de geração eólica e solar por estado, consulte as tabelas 1 e 2. Observe, na figura 5, que o número de valores pertencentes a um estado na série simulada é aproximadamente igual ao da série histórica para ambos os períodos, apontando que o método de geração de cenários sintéticos utilizado reproduz satisfatoriamente o comportamento dos dados reais do parque híbrido.

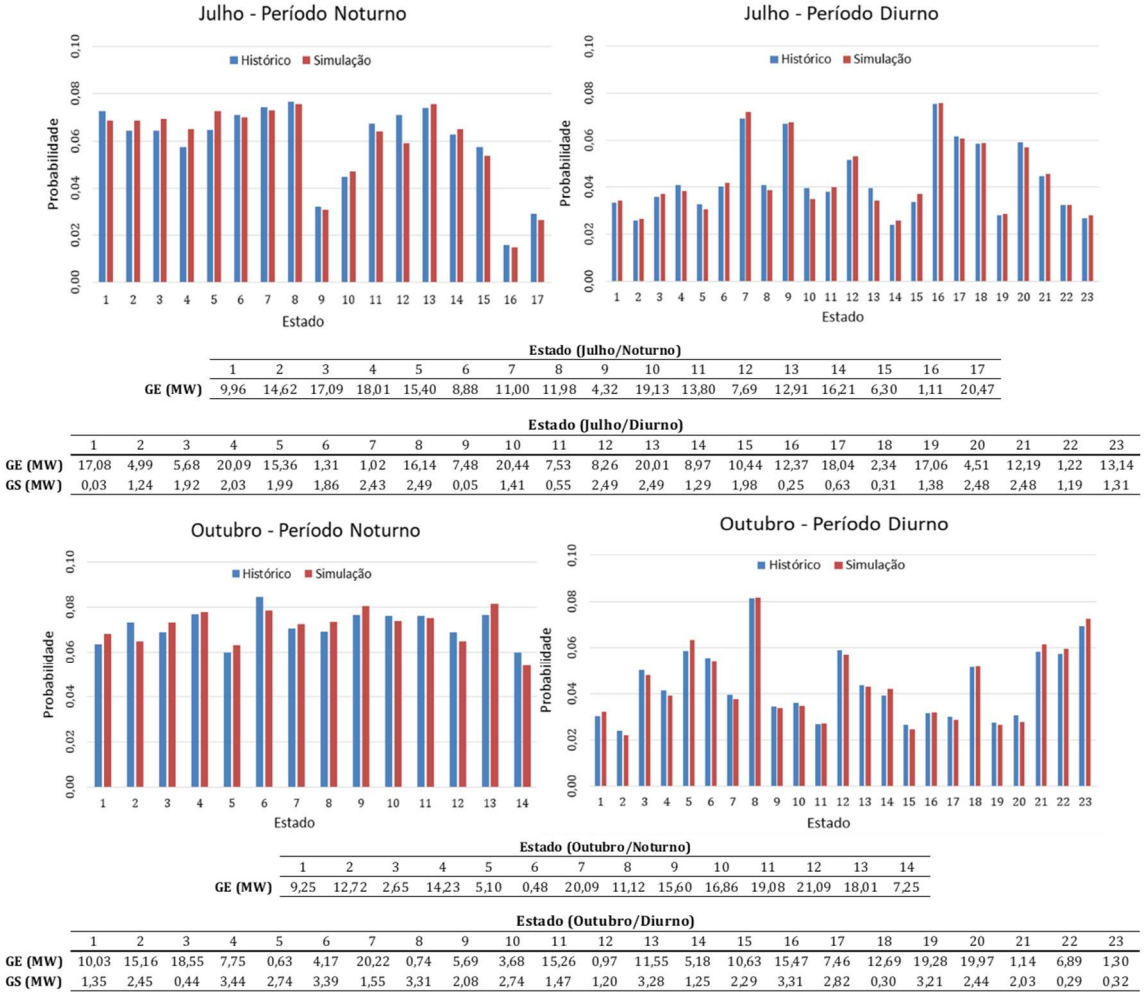


Estado (Abril/Noturno)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
GE (MW)	0,17	14,74	4,80	3,69	11,81	9,24	8,07	17,78	13,22	6,96	2,58	1,37	10,50	16,23	5,88	19,54

Estado (Abril/Diurno)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
GE (MW)	4,44	0,98	1,22	4,25	0,85	1,03	13,38	1,73	15,99	8,53	11,31	18,76	4,58	8,89	18,60	17,50	6,21	1,40	13,11	6,99	7,63	0,89	11,59
GS (MW)	2,26	0,13	1,17	0,15	2,10	2,95	2,76	1,62	0,17	2,89	0,15	1,87	2,89	2,26	2,67	1,01	0,83	0,68	1,82	1,59	0,15	2,56	0,95



Estado (Julho/Noturno)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
GE (MW)	9,96	14,62	17,09	18,01	15,40	8,88	11,00	11,98	4,32	19,13	13,80	7,69	12,91	16,21	6,30	1,11	20,47

Estado (Julho/Diurno)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
GE (MW)	17,08	4,99	5,68	20,09	15,36	1,31	1,02	16,14	7,48	20,44	7,53	8,26	20,01	8,97	10,44	12,37	18,04	2,34	17,06	4,51	12,19	1,22	13,14
GS (MW)	0,03	1,24	1,92	2,03	1,99	1,86	2,43	2,49	0,05	1,41	0,55	2,49	2,49	1,29	1,98	0,25	0,63	0,31	1,38	2,48	2,48	1,19	1,31

Estado (Outubro/Noturno)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
GE (MW)	9,25	12,72	2,65	14,23	5,10	0,48	20,09	11,12	15,60	16,86	19,08	21,09	18,01	7,25

Estado (Outubro/Diurno)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
GE (MW)	10,03	15,16	18,55	7,75	0,63	4,17	20,22	0,74	5,69	3,68	15,26	0,97	11,55	5,18	10,63	15,47	7,46	12,69	19,28	19,97	1,14	6,89	1,30
GS (MW)	1,35	2,45	0,44	3,44	2,74	3,39	1,55	3,31	2,08	2,74	1,47	1,20	3,28	1,25	2,29	3,31	2,82	0,30	3,21	2,44	2,03	0,29	0,32

Figura 5. Comparação das FDPs dos períodos noturno e diurno de janeiro, abril, julho e outubro. Elaboração dos autores.

Como as séries temporais de energia eólica e solar se comportam de forma diferente ao longo do ano, é importante verificar se esse comportamento é replicado nas séries simuladas. Essa análise pode ser feita através da comparação mensal das médias e desvios padrão entre as séries reais e sintéticas, e está apresentada nas tabelas 3 e 4, que se referem, respectivamente, aos períodos noturno e diurno dos doze meses do ano. Note que os valores de ambas as séries simuladas e observadas nas duas tabelas são muito próximos, com erros percentuais variando entre 0,02% e 8,03% para a média e entre 0,002% e 7,57% para o desvio padrão. Isso indica que o método de geração de cenários replica a sazonalidade anual das gerações eólica e solar.

Noturno				
Mês	Média GE (MW)		Desvio-padrão GE (MW)	
	Histórico	Simulação	Histórico	Simulação
Jan	7,1075	7,1523	5,6066	5,5683
Fev	8,3321	8,1116	5,8744	5,8543
Mar	7,6001	7,5609	5,4825	5,4350
Abr	9,0943	8,7037	5,0307	5,1631
Mai	9,9412	9,8648	4,8201	4,8087
Jun	11,3306	11,5642	4,4380	4,4003
Jul	12,5797	12,7964	4,3025	4,2438
Ago	14,2650	14,1529	4,4674	4,4673
Set	14,7856	14,9157	5,1972	5,1607
Out	12,5116	12,5376	6,4796	6,4250
Nov	7,9619	8,2766	6,3168	6,3840
Dez	6,5485	7,0742	5,6507	5,7365

Tabela 3. Resultados mensais – Período noturno. Elaboração dos autores.

Mês	Diurno							
	Média GE (MW)		Desvio padrão GE (MW)		Média GS (MW)		Desvio padrão GS (MW)	
	Histórico	Simulação	Histórico	Simulação	Histórico	Simulação	Histórico	Simulação
Jan	4,4403	4,4298	4,6868	4,5388	1,9546	1,9541	1,2287	1,2105
Fev	5,5041	5,5964	5,2456	5,1569	1,9893	1,9626	1,2041	1,1880
Mar	4,9910	5,0823	5,1038	5,0203	1,8727	1,8870	1,1012	1,0774
Abr	6,5183	6,6223	5,7412	5,6179	1,6868	1,7028	1,0250	1,0109
Mai	7,3487	7,1303	6,2155	5,9711	1,4618	1,4542	0,9292	0,9180
Jun	9,0801	8,8672	6,3171	6,1796	1,3621	1,3711	0,8833	0,8715
Jul	10,2616	10,0681	6,5590	6,3537	1,4461	1,4369	0,9222	0,9093
Ago	10,5803	10,3820	6,7794	6,6892	1,7020	1,7154	1,0164	1,0013
Set	9,7462	9,4904	7,0191	6,8807	1,9673	1,9910	1,0876	1,0712
Out	7,8613	7,6955	6,7366	6,5772	1,9942	1,9801	1,1607	1,1439
Nov	5,4256	5,0238	5,7174	5,2843	1,8132	1,7829	1,1873	1,1652
Dez	4,3471	4,3460	4,8244	4,7067	1,8780	1,9168	1,2128	1,2014

Tabela 4. Resultados mensais – Período diurno. Elaboração dos autores.

O último procedimento a ser testado é a comparação dos cenários simulados com as gerações reais do ano de 2018. Primeiro, foi calculada a geração média mensal por período - noturno e diurno, de cada um dos 200 cenários simulados. Com isso, foi possível construir os gráficos das figuras 6 e 7, que mostram, mês a mês, a geração horária média sintética e a real do ano de 2018. Os gráficos também mostram as médias simuladas máxima e mínima. O período diurno é apresentado em dois gráficos na figura 7, já que há dois tipos de ERVs nesse período, eólica e FV. Pode-se perceber que a média real se encontra entre a mínima e a máxima simuladas em 94,4% dos casos, com exceção de setembro e novembro da geração eólica do período diurno. Além disso, é possível notar que o movimento sazonal da geração média real é acompanhado pela média sintética de maneira adequada nos três gráficos. Portanto, podemos afirmar que as séries simuladas reproduzem satisfatoriamente o comportamento de uma série real fora da amostra da base histórica usada no processo de simulação.

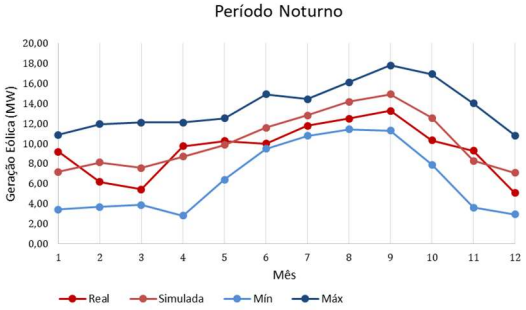


Figura 6. Médias de geração de 2018 e dos cenários simulados - Noturno. Elaboração dos autores.

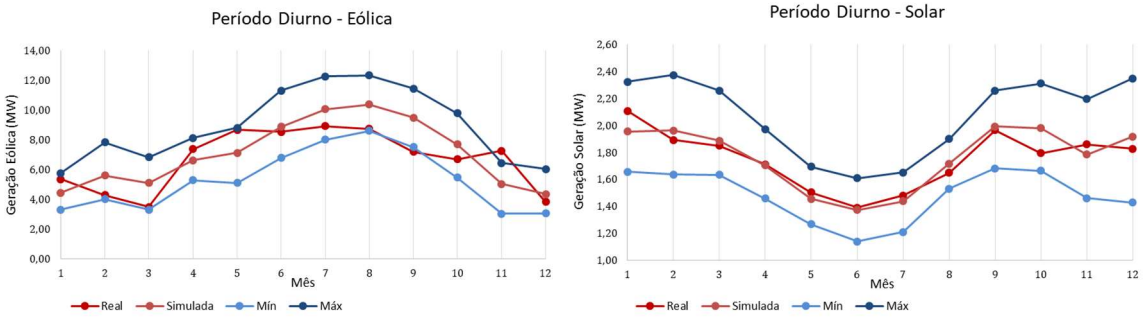


Figura 7. Médias de geração de 2018 e dos cenários simulados - Diurno. Elaboração dos autores.

4. Conclusões

O aumento da participação de ERVs na matriz elétrica brasileira traz diversos desafios para o planejamento e operação do SEB, devido à estocasticidade presente na geração de fontes alternativas. Tais desafios envolvem a modelagem e simulação dos processos intermitentes de geração de renováveis e, neste contexto, este trabalho teve como objetivo simular cenários de

geração multivariada de uma usina híbrida para o horizonte de um ano. Foram utilizados dados técnicos e de localização da usina de Caetité, e a metodologia de simulação proposta baseia-se na discretização das séries temporais históricas para geração de cenários a partir de cadeias de Markov.

Pode-se afirmar que os cenários obtidos a partir da metodologia de simulação aplicada reproduzem adequadamente as características da base de dados histórica de 39 anos utilizada, a partir da avaliação das funções de distribuição de probabilidade; das características mensais, como média e desvio padrão; e, por fim, da comparação entre os cenários simulados e as gerações históricas do ano de 2018, cujos dados não foram utilizados no processo de simulação. Portanto, o objetivo deste artigo de gerar cenários multivariados de energia renovável compatíveis com a realidade foi atingido satisfatoriamente.

Como proposta para estudos futuros, sugerem-se: (i) aplicar a metodologia de simulação discretizando as séries históricas de forma alternativa aos agrupamentos noturno e diurno, considerando, por exemplo, grupos de granularidade temporal menor, como o horário; (ii) avaliar a metodologia de geração multivariada considerando outras fontes de energia, como a hidráulica; (iii) aplicar a metodologia de simulação a usinas de diferentes locais, permitindo, assim, a avaliação do método considerando os efeitos de complementaridade não apenas temporal, mas também espacial entre as ERVs.

Referências

Alipour, M., Aghaei, J., Norouzi, M., Niknam, T., Hashemi, S. e Lehtonen M. (2020). A novel electrical net-load forecasting model based on deep neural networks and wavelet transform integration. *Energy*, Vol. 205, N. 118106.

Almutairi, A., Hassan Ahmed, M. e Salama, M. M. A. (2016). Use of MCMC to incorporate a wind power model for the evaluation of generating capacity adequacy. *Electric Power Systems Research*, Vol. 133, 63–70.

Baldioti, H. R. e Souza, R. C. (2018). Using a Markov chain Monte Carlo technique to simulate synthetic natural inflow energy scenarios. *2018 International Conference on Probabilistic Methods Applied to Power Systems*, N. 8440277.

Box, G. E. E. e Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*, HoldenDay, San Franscisco, Cal.

Dias, J. A. S. e Borges, C. L. T. (2014). A Non Parametric Stochastic Model for River Inflows based on Kernel Density Estimation. *2014 International Conference on Probabilistic Methods Applied to Power Systems*, N. 6960626.

Duca, V. E. L. A., Fonseca, T. C. O. e Cyrino Oliveira, F. L. (2021). A generalized dynamical model for wind speed forecasting. *Renewable and Sustainable Energy Reviews*, Vol. 136, N. 110421.

Esteves, G. R. T., Maçaira, P. M., Oliveira. F. L. C., Amador, G. e Souza, R. C. (2019). Improvements in the current Brazil's energy dispatch optimization: Load forecast and wind power. *ICORES 2019 - Proceedings of the 8th International Conference on Operations Research and Enterprise Systems*, 398-405.

IRENA. (2021). Web page. https://www.irena.org/newsroom/pressreleases/2021/Apr/World-Adds-Record-New-Renewable-Energy-Capacity-in-2020. Acessado: 04/05/2021.

Jurasz, J., Mikulik, J., Dabek, P. B., Guezgouz, M. e Kazmierczak, B. (2021). Complementarity and 'Resource Droughts' of Solar and Wind Energy in Poland: An ERA5-Based Analysis. *Energies*, Vol. 14, N. 1118.

Luo, G., Shi, D., Chen, J. e Wu, X. (2014). A Markov chain Monte Carlo method for simulation of wind and solar power time series. *Dianwang Jishu/Power System Technology*, Vol. 38, 321-327.

Ma, M., Ye, L., Li, J., Li, P., Song, R. e Zhuang, H. (2020). Photovoltaic Time Series Aggregation Method Based on K-means and MCMC Algorithm. *Asia-Pacific Power and Energy Engineering Conference*, Vol. 2020-September, N. 9220338.

Maçaira, P. M., Cyrillo, Y. M., Oliveira, F. L. C. e Souza, R. C. (2019). Including wind power generation in Brazil's long-term optimization model for energy planning. *Energies*, Vol. 12, N. 826.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Oakland, CA, USA, p. 14.

Miao, S., Xie, K., Yang, H., Tai H.-M. e Hu, B. (2017). A Markovian wind farm generation model and its application to adequacy assessment. *Renewable Energy*, Vol. 113, 1447-1461.

MME. (2020). Web page. https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/plano-decenal-de-expansao-de-energia-2029. Acessado: 04/05/2021.

NASA. (2021). Web page. https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/. Acessado: 12/05/2021.

Papaefthymiou, G. e Klöckl, B. (2008). MCMC for wind power simulation. *IEEE Transactions on Energy Conversion*, Vol. 23, 234–240.

Pfenninger, S. e Staffell, I. (2016). Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy*, Vol. 114, 1251-1265.

Pinson, P. (2013). Wind energy: Forecasting challenges for its operational management. *Statistical Science*, Vol. 28, 564–585.

PSR. (2021). Web page. https://www.psr-inc.com/softwares/?current=p13880. Acessado em 12/05/2021.

Renova Energia. (2016). Web page. http://www.renovaenergia.com.br/. Acessado: 10/05/2021.

Schmidt, J., Cancella, R. e Pereira, A. O. Jr. (2016). An optimal mix of solar PV, wind and hydro power for a low-carbon electricity supply in Brazil. *Renewable Energy*, Vol. 85, 137-147.

Staffell, I. e Pfenninger, S. (2016). Using Bias-Corrected Reanalysis to Simulate Current and Future Wind Power Output. *Energy*, Vol. 114, 1224-1239.