Simulação de Cenários de Geração Renovável Explorando as Características Complementares entre Diferentes Recursos

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

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

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

Erick Meira
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900
research@erickmeira.com

RESUMO

O aumento da participação de fontes de Energias Renováveis Variáveis (ERVs) nas matrizes elétricas dos países traz diversos desafios ao planejamento e à operação de sistemas de potência, devido, sobretudo, à intermitência das ERVs. Tais desafios envolvem capturar a estocasticidade das produções intermitentes, visando simular possíveis cenários de geração por fontes renováveis. Nesse contexto, a exploração dos efeitos complementares entre diferentes recursos renováveis é um tópico de crescente importância na literatura, dada a possibilidade de se obter uma potência combinada com menor variabilidade e intermitência. Nesse sentido, o presente trabalho propõe um modelo de simulação estocástica de geração de renováveis que explora diferentes efeitos complementares entre distintas ERVs no contexto do Brasil. A modelagem considera o cálculo de estados de geração de energia para se representar simultaneamente a geração de múltiplas fontes renováveis, bem como a elaboração de matrizes de probabilidades de transição com base nos dados históricos, tornando possível a obtenção de diferentes cenários de geração por meio de técnicas de simulação. As séries sintéticas obtidas são subsequentemente avaliadas a partir de técnicas estatísticas que sugerem uma boa aderência dos cenários gerados à grande parte das características reais das séries históricas.

PALAVRAS CHAVE. Energias Renováveis. Simulação Estocástica. Complementaridade.

Tópicos. EM&GP – PO na Área de Energia, Petróleo e Gás. EST&MP – Estatística e Modelos Probabilísticos.

ABSTRACT

The increased participation of variable renewable energy sources (VREs) in electrical matrices worldwide brings several challenges to the planning and operation of power systems, specially due to the intermittency of VREs. Such challenges involve capturing the stochasticity of intermittent productions, aiming to simulate possible scenarios of the stochastic processes of renewables, and

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

constituting a powerful tool to support decision-making in the public and private sectors. In this context, a topic of increasing importance in the literature is exploring the complementary effects presented between different resources since such properties result in a combined potency with less variability and intermittency. In this sense, this work proposes a stochastic simulation model of renewable generation that explores different complementary effects between different VREs in the context of Brazil. The approach considers the calculation of energy generation states to simultaneously represent the generation of multiple renewable sources, as well as the use of probability transition matrices based on historical data, making it possible to obtain different generation scenarios through simulation techniques. The synthetic series are subsequently evaluated using statistical techniques that indicate good adherence of the generated scenarios to most stylized facts of the historical series.

KEYWORDS. Renewable Energies. Stochastic Simulation. Complementarity.

Paper topics. OR in the Energy, Oil, and Gas Area. Statistics and Probabilistic Models.

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

1. Introdução

O planeta passa por profundas mudanças quanto à forma de se explorar as diferentes fontes energéticas, em particular aquelas necessárias à produção de eletricidade, recurso essencial para a sociedade moderna e fator chave para o desenvolvimento socioeconômico [GESEL 2017]. Neste sentido, nos últimos anos, as chamadas Energias Renováveis Variáveis (ERVs) vêm se destacando como alternativas econômicas e ecologicamente corretas para o abastecimento de energia tanto em redes elétricas isoladas como em redes integradas em sistemas de potência. Segundo a *International Renewable Energy Agency* [IRENA 2023], em 2020, cerca de US$ 348 bilhões haviam sido investidos em energias renováveis no planeta, enquanto, em 2022, os investimentos subiram para US$ 499 bilhões, representando um crescimento de 43%. Apesar dos avanços expressivos, a agência reforça a importância de se investir ainda mais no setor para que a tolerância máxima de aumento de 1,5°C, estipulada no âmbito do Acordo de Paris, seja respeitada. Nesse contexto, estima-se ser necessário um investimento global médio de cerca de US$ 4,4 trilhões por ano até 2030 para que se tenha sucesso [IRENA 2023].

No contexto brasileiro, os investimentos em usinas de fontes renováveis variáveis vêm aumentando expressivamente ao longo das duas últimas décadas, acompanhando a tendência mundial, com destaque para as fontes eólica e solar. O Plano Decenal de Expansão de Energia (PDE 2029) elaborado pela Empresa de Pesquisa Energética [EPE 2020] estima que as gerações a partir das fontes supracitadas terão crescimento médio anual aproximado de 6,9% até 2029, superando as projeções de crescimento médio para outras fontes, como hidráulica (2% a.a.) e carvão vegetal (1% a.a.). Assim, embora as hidrelétricas respondessem por cerca de 65% da capacidade instalada no país em 2021, a EPE planeja que, até 2030, as fontes hídricas terão sua participação reduzida para 42%, enquanto a solar fotovoltaica (FV) saltará dos atuais 2% para 9% e a eólica, de 10% para 16% [EPE 2020].

A Figura 1 apresenta a evolução das curvas de gerações eólica e solar anuais no Brasil entre 2010 e 2020, evidenciando a tendência de crescimento para ambas as fontes ao longo da última década. Observe o crescimento exponencial da curva de geração eólica a partir de 2014 e da curva solar a partir de 2017, segundo dados do Operador Nacional do Sistema Elétrico [ONS 2022].

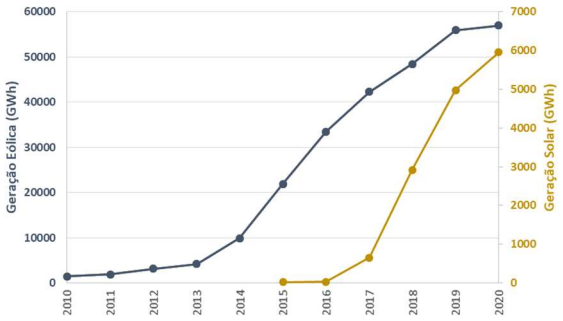


Figura 1 - Séries históricas de geração eólica e solar no Brasil entre 2010 e 2020.
Fonte: Elaborado a partir de dados do ONS [2022].

Apesar dos números crescentes e promissores, destaca-se que a inserção em larga escala de ERVs em sistemas de potência apresenta desafios únicos para planejadores e operadores de sistemas, que devem tomar ações preventivas e corretivas a fim de se manter a segurança e a confiabilidade das redes [Morales et al. 2010]. Assim, à medida que a parcela de recursos de energia renovável aumenta rapidamente nas matrizes elétricas, o reconhecimento, a representação, a quantificação e a interpretação de suas incertezas se tornam cada vez mais importantes [Sakki et al. 2022]. Na prática, segundo Pinson [2013], um dos principais desafios envolve a modelagem e simulação dos processos estocásticos de geração de energia renovável em escalas temporais e espaciais com o objetivo de embasar o processo de tomada de decisões, tanto no setor público quanto no privado.

Nos últimos anos, um volume considerável de pesquisas tem sido direcionado à modelagem e previsão ou simulação de ERVs. Nesse contexto, uma temática crescente na literatura

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

relaciona-se ao desenvolvimento de metodologias para simular de maneira conjunta diferentes recursos renováveis com características complementares, como, por exemplo, as fontes eólica e solar, trazendo oportunidades de contribuição científica. Segundo Jurasz et al. [2021], a complementaridade temporal entre ERVs resulta em uma potência combinada com menor variabilidade e intermitência, diminuindo a demanda por armazenamento e suavizando a operação do sistema elétrico.

Na Figura 2, para exemplificar o caráter complementar temporal e espacial entre as fontes eólica e solar, é apresentada, para duas usinas de diferentes locais do Rio Grande do Norte, a geração horária por unidade (p.u.). A fonte solar segue o típico formato de sino, com o pico de geração nos horários próximos ao meio-dia e valores nulos entre 18h e 04h da manhã. Já a geração eólica apresenta comportamento mais estável ao longo do dia, com os maiores níveis de produção no meio da tarde, além de manter os níveis de produção durante o período da noite.

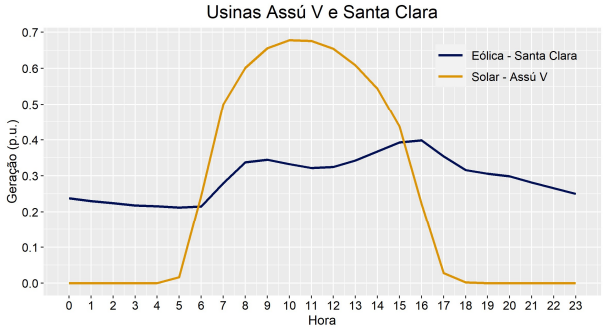


Figura 2 - Exemplo de complementaridade temporal entre as fontes eólica e solar.
Fonte: Elaborado a partir de dados do ONS [2022].

Ante o exposto, buscando-se contribuir com o tema em questão - simulação de energias renováveis - o presente trabalho propõe uma metodologia para simular cenários conjuntos de geração de ERVs, mais especificamente as fontes eólica e solar fotovoltaica, buscando, assim, capturar o caráter complementar entre esses recursos em diferentes localidades no Brasil.

No que tange ao referencial teórico da pesquisa, Li et al. [2022] subdividem os diferentes métodos existentes para simulação ou previsão de recursos renováveis em paramétricos e não paramétricos. Os métodos paramétricos podem ser divididos em probabilísticos e modelos de séries temporais. Os autores destacam duas principais desvantagens dos modelos probabilísticos: as funções de distribuição predeterminadas tendem a causar erros significativos e as informações históricas não podem ser capturadas inteiramente. Quanto aos modelos de séries temporais, muitos costumam assumir pressupostos estatísticos, como a normalidade dos resíduos, possibilitando o ajuste dos dados em uma distribuição paramétrica [Baldioti e Souza 2018]. No entanto, uma das desvantagens é que, na maioria dos casos reais referentes aos recursos renováveis, o ruído não pode ser tratado dessa forma, pois apresenta comportamentos de cauda intrinsecamente assimétricos [Baldioti e Souza 2018].

Dadas essas limitações, é possível perceber um volume elevado de publicações recentes nesta área de pesquisa envolvendo métodos não paramétricos, além de métodos híbridos ou combinados. Segundo Tawn e Browell [2022], cerca de 60% dos trabalhos encontrados na literatura, com foco nas fontes eólica e solar, incluem métodos não paramétricos, como, por exemplo, modelos de aprendizado de máquina, com destaque para algoritmos de redes neurais e suas variações. É importante sublinhar que, embora apresentem vantagens na representação da aleatoriedade das ERVs em relação a outros métodos, os modelos de *machine learning* costumam exigir grande esforço computacional, além de serem usualmente difíceis de interpretar e controlar.

Nesse meio, verifica-se um vasto número de trabalhos que modelam os processos estocásticos de recursos renováveis a partir de cadeias de Markov [Baldioti e Souza 2018; Luo et al. 2014; Almutairi et al. 2016; Miao et al. 2017; Maçaira et al. 2019; Ma et al. 2020; Melo et al. 2021], mais fáceis de interpretar e controlar comparativamente a modelos de *machine learning*. Em maiores detalhes, o presente estudo se baseia na modelagem dos processos estocásticos de

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

geração de renováveis via cadeias de Markov. No entanto, a metodologia proposta inova ao introduzir matrizes de transição de estados conjuntos, isto é, estados que representam a geração de mais de uma fonte renovável simultaneamente, visando capturar as propriedades complementares entre os recursos, dentre outras inovações propostas, detalhadas mais à frente. Os cenários são avaliados com base em diversos critérios estatísticos, como comparação da distribuição dos dados históricos e simulados, funções de autocorrelação e estatísticas descritivas. As simulações são realizadas a partir de dados de duas usinas de locais distintos do estado do Rio Grande do Norte, sendo uma eólica e outra solar FV, buscando abordar tanto a complementaridade temporal entre os recursos, quanto a espacial, já que as plantas estão situadas em diferentes localizações geográficas.

O restante deste estudo está organizado da seguinte forma: a segunda seção apresenta a metodologia de simulação e seu embasamento teórico; a terceiro seção traz uma análise descritiva dos dados utilizados no estudo; os resultados obtidos nos estudos de caso, bem como as análises dos cenários gerados, são apresentados na seção 4; por fim, a quinta seção reúne as principais conclusões da pesquisa, bem como suas implicações e sugestões para trabalhos futuros.

2. Metodologia

A Figura 3 resume os principais passos da metodologia proposta, que pode ser dividida em dois estágios principais: modelagem e simulação. A seguir, são apresentados os detalhes de cada passo, bem como são ressaltadas as inovações frente a trabalhos que modelam os recursos renováveis a partir do conceito de cadeias de Markov, como Maçaira et al. [2019] e Almutairi et al. [2016], que serviram de base para a proposta.

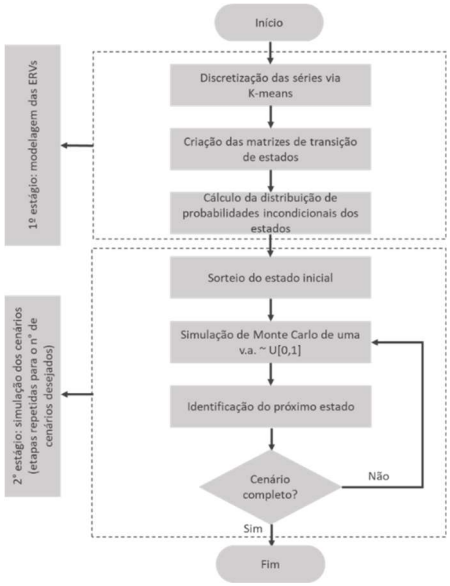


Figura 3 – Metodologia de modelagem e simulação proposta.

O primeiro passo da modelagem consiste em discretizar as séries temporais de geração, isto é, em agrupar os dados de energia eólica e solar em um número finito de estados. Para esse propósito, pode-se utilizar a técnica de agrupamento *k-means* [Macqueen 1967], devido à simplicidade e à precisão fornecidas [Almutairi et al. 2016]. O objetivo do algoritmo é classificar objetos em vários grupos (*clusters*), de modo que os objetos dentro do mesmo *cluster* sejam tão semelhantes quanto possível (alta similaridade intraclasse), enquanto os objetos de diferentes *clusters* sejam tão diferentes quanto possível (baixa similaridade interclasses) [Macqueen 1967]. No agrupamento *k-means*, cada *cluster* é representado por seu centro (centroide), que corresponde à média dos pontos atribuídos ao *cluster*.

Como primeira contribuição deste trabalho, dado que o objetivo é desenvolver uma metodologia de simulação conjunta das gerações eólica e solar, tem-se que cada dado de energia no tempo t, bem como cada centroide dos *clusters*, é constituído por dois atributos, correspondentes aos dois tipos de geração considerados. No final, o valor de cada tipo de energia

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

observado é substituído pelo atributo correspondente do centroide do *cluster* a que pertence, finalizando o processo de discretização das séries temporais. Ressalta-se que a discretização considera uma periodicidade horária média ao longo de um mês (24 discretizações por mês), evitando, assim, que os cenários apresentem perfis horários noturnos com características diurnas de produção de energia e vice-versa. Assim, os estados buscam representar tanto a sazonalidade mensal, quanto sazonalidades horárias de produção das fontes renováveis.

O segundo passo da modelagem corresponde à criação das matrizes de transição de estados, P_{ind}. Como as discretizações no passo anterior são realizadas hora a hora, as matrizes também são criadas considerando tal periodicidade. A probabilidade de transição estacionária $p_{a,b}$ do estado a para o estado b, para todos os índices $1 \leq (a, b) \leq k$, pode ser calculada pela Equação (1):

$$p_{a,b} = \frac{n_{a,b}}{\sum_{i=1}^{k} n_{a,i}} \tag{1}$$

onde $n_{a,b}$ é o número de transições de a para b no histórico de dados, e o denominador é o total de transições de a para todos os demais estados possíveis. Após obter as probabilidades de transição para cada estado, é possível construir a matriz de transição entre as horas de cada mês, conforme a Equação (2):

$$P_{ind} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,z} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,z} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k,1} & p_{k,2} & \cdots & p_{k,z} \end{bmatrix} \tag{2}$$

A partir da matriz de transição individual P_{ind}, cria-se a matriz de transição de probabilidade acumulada P_{acum}, de modo que sua última coluna seja igual a um para cada linha da matriz, conforme (3):

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

Para simular cenários concatenados cronologicamente no nível horário, observou-se ser necessário elaborar matrizes cujas linhas sejam compostas pelos estados da hora h e cujas colunas, pelos estados da hora $h + 1$. Dessa maneira, na proposta, são elaboradas 24 matrizes por mês, representando as transições entre as horas de cada mês. Sublinha-se que as matrizes não são necessariamente quadradas e, portanto, de Markov, visto que os estados da hora h não são obrigatoriamente idênticos, em número e valor, aos estados da hora $h + 1$. Essa é outra inovação da metodologia proposta, cujo objetivo é manter, nos cenários, as características sazonais de nível horário dos dados históricos.

Além das matrizes de transição entre as horas do mesmo mês, denominadas matrizes intramensais, propõe-se também a criação de matrizes de transição entre os meses, ou intermensais, com o objetivo de tornar todo o processo de simulação concatenado cronologicamente. Assim, devem ser criadas 11 matrizes de transição intermensais, nas quais as linhas são representadas pelos estados da última hora ($h = 23$) do mês m, e as colunas, pelos estados da primeira hora ($h = 0$) do mês $m + 1$.

Como último passo da modelagem, realiza-se o cálculo da distribuição de probabilidade incondicional dos estados, para futuro sorteio do estado inicial dos cenários. Dessa forma, cada hora de cada mês pode ter sua distribuição incondicional calculada, que pode ser assim chamada porque é independente do estado da hora anterior.

Para iniciar o processo de simulação, gera-se, via simulação de Monte Carlo, um número aleatório de uma distribuição uniforme $U \sim [0,1]$. Em seguida, compara-se o valor do número aleatório com os elementos do vetor de distribuição de probabilidade inicial acumulada da hora h a partir da qual se deseja iniciar a simulação. Se o número gerado aleatoriamente for maior que a

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

probabilidade acumulada de um estado qualquer s e menor ou igual à probabilidade acumulada do estado $s+1$ subsequente a s, o estado $s+1$ é selecionado para ser o primeiro do cenário sintético.

No segundo passo, gera-se, novamente, um número aleatório de uma distribuição uniforme $U \sim [0,1]$. A partir deste número aleatório e da matriz de transição acumulada para a hora h, cujos estados compõem as linhas da matriz, e da hora $h+1$, cujos estados compõem as colunas, obtém-se o próximo estado simulado de geração conjunta. Para isso, compara-se o valor do número aleatório com os elementos da n-ésima linha da matriz de transição de probabilidade acumulada, correspondente ao estado simulado para a hora h. 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 a geração conjunta sintética da hora $h+1$. Este procedimento deve ser repetido para simular as séries horárias de geração eólica e solar conjuntamente para todo o período desejado e pode ser representado segundo a Equação (4).

$$SWS_{h,d,m,y,g,c} = M_{gk} \quad se \quad p_{acum_{k-1}} < p_{unif[0,1]} \leq p_{acum_k} \tag{4}$$

onde:

- $SWS_{h,d,m,y,g,c}$ é a variável que representa a geração simulada para hora h, dia d, mês m, ano y, tipo de geração g, e cenário c;
- M_{gk} é o valor do estado simulado, e corresponde ao atributo do centróide do *cluster* k para o tipo de geração g;
- p_{acum_k} é o valor da probabilidade de transição, na matriz de Markov acumulada, para o estado k, a partir da linha do último estado simulado;
- $p_{unif[0,1]}$ é o número aleatório da distribuição uniforme $U \sim [0,1]$, obtido por simulação de Monte Carlo.

Destaca-se que a matriz de transição é alterada a cada iteração do processo de simulação, buscando representar nos cenários gerados tanto a variabilidade horária intramensal, quanto intermensal das fontes de energia consideradas.

3. Base de dados – Usinas Assú V e Santa Clara

Este capítulo apresenta a base de dados da pesquisa, detalhando as etapas de obtenção, tratamento e análise descritiva dos dados. Cabe mencionar que todas as etapas de pré-processamento, modelagem e simulação foram realizadas a partir de códigos elaborados em linguagem de programação R [2022].

Para compor a base de dados, visando explorar os efeitos de complementaridade espaço-temporal entre as ERVs nos cenários simulados, foram selecionadas duas séries históricas de geração de usinas de locais distintos do estado do Rio Grande do Norte (RN). A Usina Fotovoltaica Assú V, de 34 MW de capacidade instalada, situa-se no município de Assú e iniciou sua operação em dezembro de 2017. Já o Complexo Eólico Santa Clara está localizado no município de Parazinho, iniciou sua operação comercial em abril de 2014 e totaliza 188 MW de capacidade de geração. A distância entre as usinas é de aproximadamente 130 km.

As séries históricas de gerações horárias (MWh) foram obtidas do site do nos [2022] para o período de dezembro de 2017 a junho de 2022. Como o mês de dezembro de 2017 foi o mês de início da operação de Assú V, decidiu-se por não o utilizar para não incorporar à modelagem as instabilidades inerentes ao período operacional inicial. Além disso, mesmo Santa Clara apresentando uma série histórica mais longa do que Assú V, foi considerado apenas o período a partir de janeiro de 2018, visto que as séries de geração eólica e solar devem possuir a mesma janela temporal de dados para possibilitar a modelagem simultânea das ERVs.

Embora não tenham sido encontrados dados de geração negativos ou acima da capacidade instalada das usinas, foram identificados alguns dados faltantes (0,06% da base) ou dias inteiros de geração nula (1,65% do histórico). Dados faltantes são usualmente ocasionados por erros de registro. Já dias inteiros de geração nula ocorrem, geralmente, devido à manutenção e paralisação

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

das usinas. Em todos esses casos, para cada registro horário nulo ou em branco foi atribuída a geração horária média calculada a partir dos demais dias do mês de ocorrência da lacuna.

Na Figura 2 da Introdução deste trabalho, é possível observar a geração horária p.u. das fontes eólica e solar das duas usinas, destacando a complementaridade diária entre os recursos. Para analisar o comportamento das energias renováveis ao longo do ano, são apresentados, na Figura 4, os *boxplots* mensais das gerações das duas usinas.

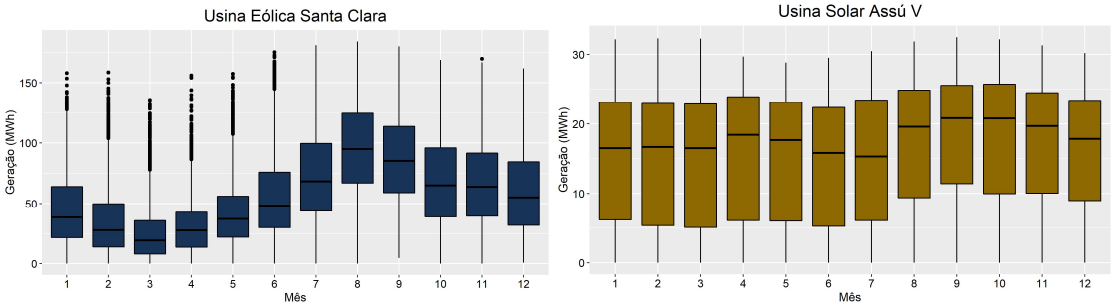


Figura 4 - *Boxplots* mensais das gerações de Santa Clara e Assú V.
Fonte: Elaborado a partir de dados do ONS [2022].

Para a fonte eólica, o mês de março apresenta os menores níveis de geração, que passam, então, a crescer até atingir os maiores patamares em agosto/setembro. Já o comportamento mensal da fonte solar não varia substancialmente ao longo do ano, mas é possível notar que a geração mantém-se em torno do mesmo nível entre janeiro e julho, a partir do qual a produção solar cresce até atingir as medianas máximas entre setembro e novembro. A partir de então, retoma-se o nível de geração dos meses iniciais do ano. Assim, é possível notar uma inegável complementaridade entre as fontes, tanto no nível horário quanto no nível mensal.

4. Resultados

No primeiro passo do estágio de modelagem, obteve-se um número finito de estados por "hora-mês" aplicando a técnica *k-means*. Assim, foram realizados 24 agrupamentos por mês, totalizando 288 discretizações. Para cada série, foram analisados diversos agrupamentos, variando o número de *clusters* entre 1 e 20. Escolheu-se, para cada conjunto, o número de *clusters* n cuja distância total intracluster fosse no mínimo igual a 98% da distância total intracluster de $n-1$ grupos, com base em Maçaira et al. [2019]. O número de *clusters* por hora variou entre 13 e 19.

No segundo passo da modelagem, foram geradas as matrizes de transição individuais e acumuladas. No total, foram obtidas 24 matrizes de transição para cada mês (intramensais) e 11 entre os meses (intermensais), totalizando 299 matrizes de transição. A Tabela 1 ilustra, como exemplo, a matriz acumulada entre as horas 05 e 06 da manhã do mês de junho. Note que o número de estados identificados para a hora 05 foi igual a 19 (linhas da matriz) e, para a hora 06, igual a 18 (colunas da matriz). Os centroides de cada estado para cada hora estão descritos na Tabela 2 em MWh, sendo "GE" geração eólica e "GS", solar.

Tabela 1 - Matriz de transição de probabilidade acumulada entre as horas 05 e 06 de junho.

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

Estado a	Estado b																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,36	0,55	0,55	0,91	0,91	0,91	0,91	1,00
2	0,00	0,50	0,50	0,50	0,63	0,63	0,63	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,88	0,88	0,88	1,00
3	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,25	0,38	0,63	0,63	0,63	0,63	0,63	0,63	0,63	0,75	1,00
4	0,00	0,07	0,07	0,36	0,36	0,43	0,43	0,43	0,43	0,43	0,43	0,43	0,57	0,57	0,57	0,57	1,00	1,00
5	0,00	0,00	0,00	0,00	0,17	0,17	0,17	0,50	0,67	0,67	0,67	0,67	0,67	0,67	1,00	1,00	1,00	1,00
6	0,00	0,00	0,00	0,00	0,00	0,20	0,20	0,80	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
7	0,00	0,00	0,00	0,00	0,20	0,20	0,40	0,40	0,40	0,40	0,40	0,40	0,40	0,40	1,00	1,00	1,00	1,00
8	0,00	0,00	0,00	0,00	0,00	0,00	0,60	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
9	0,00	0,00	0,25	0,25	0,25	0,25	0,25	0,38	0,38	0,50	0,50	0,50	0,88	0,88	0,88	1,00	1,00	1,00
10	0,00	0,00	0,08	0,08	0,17	0,25	0,42	0,42	0,50	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,83	1,00
11	0,20	0,20	0,20	0,20	0,20	0,20	0,20	0,20	0,20	0,20	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
12	0,45	0,45	0,45	0,45	0,45	0,45	0,45	0,45	0,45	0,45	0,64	0,91	0,91	1,00	1,00	1,00	1,00	1,00
13	0,00	0,00	0,00	0,00	0,00	0,40	0,40	0,40	0,40	0,40	0,40	0,40	0,60	0,60	0,60	0,60	1,00	1,00
14	0,00	0,00	0,17	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	1,00	1,00
15	0,00	0,00	0,00	0,00	0,00	0,13	0,13	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,88	1,00	1,00	1,00
16	0,00	0,00	0,00	0,13	0,13	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,50	1,00	1,00
17	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,83	0,83	0,83	0,83	0,83	0,83	1,00	1,00
18	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,38	0,50	0,63	0,63	0,63	0,63	0,63	0,75	1,00
19	0,09	0,09	0,18	0,18	0,18	0,27	0,27	0,45	0,73	0,73	0,73	0,73	0,73	0,73	0,73	0,73	0,73	1,00

Tabela 2 - Valores (MWh) dos centroides dos estados das horas 05 e 06 de junho.

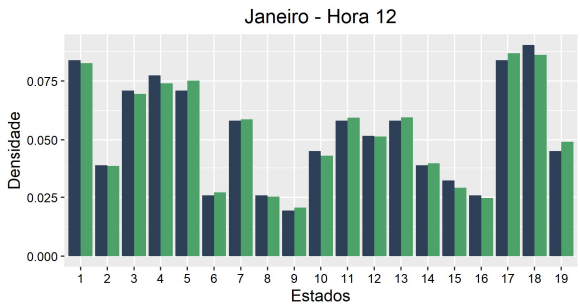
Painel A: Estados Hora 05/Junho																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
GE (MWh)	115,0	17,58	57,68	38,64	11,17	28,85	23,38	34,89	34,60	48,39	140,8	95,58	40,76	65,23	10,07	40,03	91,45	80,44	43,69
GS (MWh)	0,049	0,129	0,016	0,229	0,004	0,002	0,035	0,037	0,126	0,051	0,260	0,217	0,457	0,143	0,053	0,302	0,416	0,036	0,005
Painel B: Estados Hora 06/Junho																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	-
GE (MWh)	89,13	14,73	48,12	54,00	14,03	27,85	32,39	25,24	39,34	55,54	128,2	95,72	26,32	109,8	11,15	9,94	53,59	56,20	-
GS (MWh)	9,512	10,68	6,590	10,50	8,159	9,277	4,208	1,788	0,226	4,277	9,766	7,575	10,73	2,846	3,991	9,716	8,936	2,218	-

Em seguida, obtém-se o vetor de probabilidade incondicional para se sortear o primeiro valor de geração de cada cenário. Cada cenário deste estudo foi construído para um horizonte de tempo de um ano completo, iniciando em 01 de janeiro e finalizando em 31 dezembro. Logo, o vetor de probabilidade de estado foi calculado para os estados da hora zero do mês de janeiro. Após obtidas as matrizes de transição e o vetor de probabilidade de estado inicial, foram simulados os cenários conjuntos. Por motivo de convergência, todo o processo de simulação foi repetido 200 vezes, gerando, portanto, o mesmo número de cenários anuais para o parque híbrido.

Para validar as séries de geração simuladas, as propriedades estatísticas mais relevantes dos dados históricos devem ser reproduzidas pelos cenários sintéticos, dentro de um limite numérico razoável. Dessa forma, foram avaliados os seguintes fatores estatísticos: (1) funções de distribuição de probabilidade; (2) estatísticas descritivas (média e desvio padrão das ERVs); (3) funções de autocorrelação (FACs). Os resultados são apresentados nas subseções a seguir.

4.1 Funções de distribuição de probabilidade:

Foram geradas 288 funções de distribuição históricas e o mesmo número para as séries simuladas, 24 por mês, representando-se, assim, todas as horas do ano. Constata-se a similaridade das frequências dos estados nas séries simuladas e nas séries históricas para todos os casos, apontando que o método de geração de cenários sintéticos proposto reproduz satisfatoriamente o comportamento dos dados reais do parque híbrido em relação às frequências dos estados de geração. Para manter o trabalho conciso, apresenta-se, na Figura 5, a distribuição dos estados da série histórica e simulada para o horário de meio-dia (12h) de janeiro. Os centroides de cada estado podem ser vistos na Tabela 3.



DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834



Figura 5 - Densidade histórica e simulada da hora 12 de janeiro.

Tabela 3 - Valores (MWh) dos centroides dos estados da hora 12 de janeiro.

Fonte	Estado (Hora 12/Janeiro)																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
GE (MWh)	101,0	70,08	38,30	31,25	51,55	5,55	73,76	18,37	23,37	14,43	49,35	50,91	19,05	33,87	52,71	43,63	29,78	70,59	102,0
GS (MWh)	24,18	18,22	22,64	21,04	19,94	21,78	27,97	10,53	22,11	29,19	26,61	23,63	24,92	24,84	12,25	29,89	17,05	23,75	15,89

4.2 Estatísticas descritivas - média e desvio padrão das ERVs

A Tabela 4 apresenta as médias e desvios padrão mensais para os dados históricos e as séries sintéticas, sendo "GE" geração eólica e "GS", geração solar. Note a proximidade entre os valores das séries simuladas e observadas: para a média eólica, os erros percentuais variam entre 0,40% e 0,91%, enquanto para solar, entre 0,34% e 5,91%; já em relação ao desvio-padrão, a fonte eólica apresenta uma discrepância entre 0,76% e 2,44%, enquanto a solar, entre 0,10% e 3,42%. Isso indica que o método proposto replica as características estatísticas mensais avaliadas das gerações eólica e solar.

Tabela 41 - Estatísticas mensais históricas e simuladas.

Mês	Média GE (MWh)		Desvio padrão GE (MWh)		Média GS (MWh)		Desvio padrão GS (MWh)	
	Histórico	Simulação	Histórico	Simulação	Histórico	Simulação	Histórico	Simulação
Jan	45,271	44,987	30,198	29,755	15,031	14,142	9,336	9,644
Fev	35,514	35,671	28,467	28,128	14,828	14,004	9,487	9,737
Mar	25,394	25,379	23,258	22,815	14,559	14,232	9,481	9,558
Abr	31,113	31,016	22,904	22,361	15,486	15,100	9,670	9,799
Mai	41,074	40,742	26,317	25,675	14,810	13,941	9,316	9,635
Jun	55,462	55,425	34,462	33,977	14,033	13,593	9,242	9,371
Jul	73,049	73,156	37,145	36,779	14,488	14,142	9,507	9,627
Ago	96,494	95,617	37,762	37,143	16,621	16,333	9,596	9,675
Set	86,789	86,911	35,129	34,715	17,863	17,802	9,611	9,602
Out	69,274	69,075	37,049	36,594	17,873	17,811	9,593	9,538
Nov	66,868	66,524	34,016	33,757	17,177	16,600	9,262	9,542
Dez	60,242	60,525	34,664	34,401	15,900	15,328	9,089	9,334

A seguir, são apresentados, na Figura 6, os gráficos por fonte com a geração média mensal da série histórica e de cada um dos 200 cenários simulados, bem como a média total das simulações. Também são apresentados o 5° e o 95° percentis das médias mensais dos cenários, denominados nos gráficos como limites e representando, portanto, um intervalo de confiança empírico de 90% para as gerações renováveis. Os 10% fora do intervalo representam casos extremos, de baixa probabilidade de ocorrência com base no histórico.

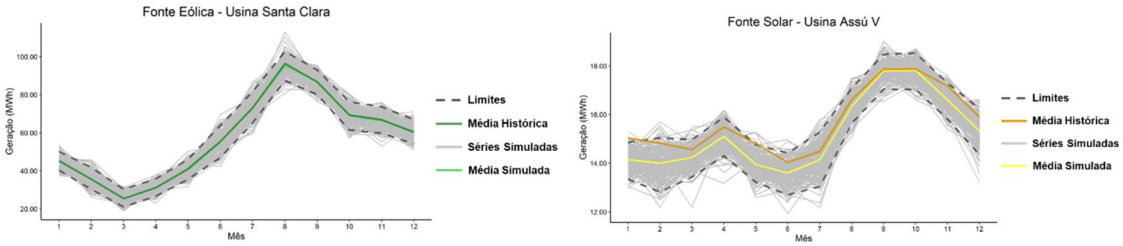


Figura 6 – Médias histórica e simulada das gerações de Santa Clara e Assú V.

Em ambas as renováveis, a média simulada acompanhou bem a média histórica, sobretudo para a fonte eólica, visto que, para o caso fotovoltaico, a média simulada se deslocou levemente da histórica em alguns meses, especificamente nos dois primeiros meses do ano e em maio. Destaca-se que, ainda assim, as discrepâncias entre as médias nesses três meses, embora sejam as maiores do ano, não ultrapassam o valor de 6%.

4.3 Funções de autocorrelação

A função de autocorrelação (FAC) mede o grau de correlação de uma variável, em um dado instante, consigo mesma, em um instante de tempo posterior. A FAC é, dessa forma, uma

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

ferramenta importante para encontrar padrões de repetição/sazonalidade em séries temporais. É possível observar que as FACs dos cenários simulados reproduzem adequadamente as FACs das séries temporais históricas. Logo, pode-se dizer que os padrões de repetição de nível horário foram replicados de forma satisfatória pelas séries sintéticas. A Figura 7 ilustra as FACs históricas e simuladas para as fontes eólica e solar.

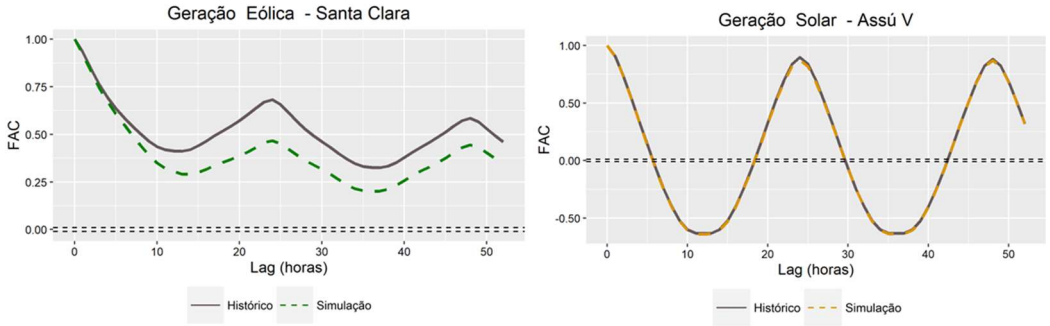


Figura 7 - FACs histórica e simulada das gerações de Santa Clara e Assú V.

Note a similaridade entre os *shapes* das FACs reais e sintéticas, com destaque para a fonte fotovoltaica. Outra constatação relevante relaciona-se às defasagens (*lags*) nas quais as autocorrelações se destacam em comparação aos demais. Em ambos os casos, eólico e solar, os *lags* múltiplos de 24 se destacam com autocorrelações positivas, visto que as gerações tendem a se comportar de forma semelhante nos mesmos horários ao longo dos dias. Além disso, no caso solar, os *lags* múltiplos de 12 também se destacam, no entanto, alternando o sinal da autocorrelação a cada 12 *lags*. Isto se deve à alternância entre o dia e a noite, que tendem a durar em média 12h cada em regiões de latitudes baixas, como é o caso do estudo.

5. Conclusões e implicações

No Sistema Elétrico Brasileiro (SEB), complexo e de grande porte, o planejamento e a programação da operação envolvem avaliar as condições futuras de fornecimento de energia de fontes hídricas e térmicas ao longo dos horizontes de planejamento, buscando-se minimizar o valor esperado do custo de operação no período considerado e priorizando a redução do consumo de combustível devido ao eventual acionamento de usinas térmicas para atender à demanda da carga de energia.

Destaca-se que o SEB é pautado pela interdependência operativa entre as plantas e deve considerar os efeitos de complementaridade entre as diferentes fontes, tarefa de significativa complexidade. Para o caso específico da geração hidrelétrica, à exceção de usinas a fio d'água, é possível minimizar a volatilidade da fonte por meio do gerenciamento de reservatórios, o que não é aplicável para os casos eólico e solar, cujas usinas não são despacháveis. Assim, a simulação de processos estocásticos de geração de energia, tanto em escalas temporais quanto espaciais, consiste em uma poderosa ferramenta de apoio aos processos de planejamento, operação e monitoramento de sistemas elétricos complexos, como o brasileiro, podendo contribuir, dessa forma, para aumentar sua robustez e confiabilidade.

Nos modelos contemporâneos de otimização de despacho de energia, a produção a partir das fontes intermitentes é descontada da demanda de energia de forma determinística no SEB. Em outras palavras, tais modelos não consideram o comportamento estocástico das ERVs. Dessa forma, emerge a necessidade de também se considerar cenários estocásticos de geração por fontes eólica e solar, buscando-se planejar a operação do sistema de forma mais eficiente ao se mitigar possíveis desperdícios e se reduzir a probabilidade de *blackouts* e racionamentos.

Nesse contexto, este trabalho identificou uma oportunidade de contribuição científica relacionada à simulação conjunta de recursos com características complementares. A metodologia proposta foi aplicada aos dados de duas usinas de diferentes fontes de geração, uma eólica e outra, fotovoltaica, e situadas em locais distintos, explorando-se, assim, os efeitos de complementaridade espaço-temporal entre as fontes. Destaca-se que a metodologia gerou cenários que reproduziram

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

adequadamente múltiplos fatos estilizados das séries históricas em ambas as aplicações. Por conseguinte, o objetivo desta pesquisa de gerar cenários conjuntos de energias renováveis que replicam as características principais dos processos estocásticos reais foi atingido satisfatoriamente.

Conclui-se, dessa forma, que a metodologia proposta pode ser utilizada para se gerar cenários probabilísticos de geração de energia, passíveis de serem utilizados no processo de tomada de decisões do SEB. Os cenários estocásticos podem ser simulados para serem integrados no modelo computacional de despacho hidrotérmico brasileiro, contribuindo para se mitigar erros na tomada de decisão pelo ONS e otimizar o uso nacional de recursos de energia.

No contexto da iniciativa privada, as empresas que investem em produção de energia devem reunir o máximo de informações acuradas sobre as características da geração em suas regiões de interesse para tomar decisões apropriadas acerca de investimentos, comércio de eletricidade, gestão de risco e formação de portfólio. Assim, a simulação de cenários probabilísticos também configura uma importante ferramenta para o setor privado.

Destaca-se, ainda, que pesquisas no contexto deste trabalho contribuem para a construção de conhecimento a respeito dos processos estocásticos de geração a partir das fontes intermitentes de energia, propiciando ferramentas que fomentam o setor e auxiliam o gerenciamento dos sistemas elétricos. A temática abordada é, assim, de suma importância para fomentar a transição para uma matriz energética global limpa e sustentável, essencial para o alcance dos Objetivos do Desenvolvimento Sustentável da ONU, sobretudo o Objetivo 13, que visa à adaptação e à limitação das mudanças climáticas.

Por fim, para estudos futuros, sugerem-se: (i) a simulação de cenários conjuntos de geração considerando-se outras fontes de energia além da eólica e solar, como hidráulica e biomassa; (ii) a aplicação da metodologia proposta em outros locais do Brasil e do mundo, buscando-se obter resultados mais generalizados; (iii) a inserção de cenários probabilísticos das fontes eólica e solar no modelo de otimização de despacho do SEB, avaliando seus impactos; e, por fim, (iv) a inclusão de variáveis exógenas na modelagem, como indicadores dos fenômenos climáticos El Niño e La Niña.

Agradecimentos

Este trabalho foi apoiado por: Coordenação Brasileira de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) [número 001]; o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) com Bolsas [números 422470/2021-0, 307084/2022-1 e 311519/2022-9], a Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) com Bolsas [números 210.618/2019, 211.086/2019, 211.645/2021, 201.243/2022, 201.348/2022 e 210.041/2023] e Programa de P&D da ANEEL e Grupo CTG BRASIL sob outorga nº 10381-0322/2022 (*Portfolio Diversification with a View to Optimizing Revenue from Electricity Sales*).

Referências

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

Baldioti HR, Souza RC. Using a Markov chain Monte Carlo technique to simulate synthetic natural inflow energy scenarios. 2018 International Conference on Probabilistic Methods Applied to Power Systems 2018;8440277.

EPE – Empresa de Pesquisa Energética, Plano Decenal de Expansão de Energia, https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/plano-decenal-de-expansao-de-energia-2029; 2020 [accessed 20 March 2023].

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834

GESEL – Grupo de Estudos do Setor Elétrico. A energia renovável variável, https://gesel.ie.ufrj.br/app/webroot/files/IFES/BV/sales90.pdf; 2017 [accessed 14 July 2022].

IRENA – International Renewable Energy Agency, https://www.irena.org/publications; 2023 [accessed 10 March 2023].

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

Li H, Ren Z, Fan M, Li W, Xu W, Jiang Y, Xia W. A review of scenario analysis methods in planning and operation of modern power systems: Methodologies, applications, and challenges. Electric Power Systems Research 2022;205:1-20.

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

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

Maçaira PM, Cyrillo YM, Oliveira FLC, Souza RC. Including wind power generation in Brazil's long-term optimization model for energy planning. Energies 2019;12(826).

Macqueen J. Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1967:14.

Melo GA, Maçaira PM, Oliveira FLC, Silva Junior AS. Simulação de geração multivariada de renováveis via cadeias de Markov. Anais do Simpósio Brasileiro de Pesquisa Operacional 2021, João Pessoa; 1:1.

Miao S, Xie K, Yang H, Tai HM, Hu B. A Markovian wind farm generation model and its application to adequacy assessment. Renewable Energy 2017;113:1447-1461.

Morales JM, Mínguez R, Conejo AJ. A methodology to generate statistically dependent wind speed scenarios. Applied Energy 2010;87(3):843–855.

ONS – Operador Nacional do Sistema Elétrico, https://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/geracao_energia.aspx; 2022 [accessed 05 December 2022].

Pinson P. Wind energy: Forecasting challenges for its operational management. Statistical Science 2013;28:564–585.

R Core Team 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sakki GK, Tsoukalas I, Kossieris P, Makropoulos C, Efstratiadis A. Stochastic simulation-optimization framework for the design and assessment of renewable energy systems under uncertainty. Renewable and Sustainable Energy Reviews 2022;168:112886.

Tawn R, Browell J. A review of very short-term wind and solar power forecasting. Renewable and Sustainable Energy Reviews 2022;153:111758.

DOI: 10.59254/sbpo-2023-174834
http://live-proceedings.pantheonsite.io/p/174834