LVI SBPO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
DE 04 ATÉ 07 NOVEMBRO 2024
FORTALEZA | CEARÁ

Simulação de Energias Renováveis Complementares sob o Efeito de El Niño e La Niña

Gustavo de Andrade Melo[1], Fernando Luiz Cyrino Oliveira[1], Paula Medina Maçaira[1], Rafaela Moreira de Azevedo Ribeiro[1], Tuany Esthefany Barcellos de Carvalho Silva[1], Rafael Araújo Couto[1], Thiago dos Santos de Lima Vieira[1], Bruno Fânzeres dos Santos[1], Reinaldo Castro Souza[1], Olavo Bet[2], Jônatas Rafael Hernandes[2]

[1]Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900
[2]China Three Gorges Corporation (CTG Brasil)
Rua Funchal, 418 - 3º andar, Vila Olímpia - São Paulo, SP - Brasil - 04551-060
gustavo.melo.rio@gmail.com

RESUMO

Este trabalho propõe uma metodologia baseada em Cadeias de Markov para simulação dos processos estocásticos de geração de fontes de Energias Renováveis Variáveis (ERVs). Objetiva-se representar, nos cenários, tanto a influência de fenômenos climáticos, quanto a complementaridade entre diferentes fontes. Para validação da metodologia, são apresentados estudos de caso cobrindo todo o território brasileiro a partir de dados de três ERVs distintas. Os resultados são promissores, visto que a metodologia proposta supera o modelo *benchmark* aplicado ao mercado brasileiro sob a ótica dos critérios considerados na avaliação dos cenários.

PALAVRAS CHAVE. Simulação de Energias Renováveis. Fenômenos Climáticos. Complementaridade.

Tópicos. EN&PG - PO na área de Energia, Petróleo e Gás. EST&MP - Estatística e Modelos Probabilísticos.

ABSTRACT

This work proposes a methodology based on Markov Chains for simulating stochastic processes of Variable Renewable Energy (VRE) sources production. The main objective is to capture, in the scenarios, both the influence of climate phenomena and the complementarity between different sources. Case studies are developed based on data from three renewable energies covering the Brazilian territory. Regarding the evaluation criteria, the results are satisfactory, and the proposed methodology performs better than the benchmark model applied in the Brazilian market.

KEYWORDS. Renewable Energy Simulation. Climatic Phenomena. Complementarity.

Paper topics. EN&PG - OR in the Energy, Oil, and Gas Sector. EST&MP - Statistics and Probabilistic Models.

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

1. Introdução

O aumento da participação de fontes de energias renováveis nas matrizes elétricas dos países é essencial para a descentralização energética, a descarbonização e o alcance de diversos Objetivos de Desenvolvimento Sustentável (ODS) estabelecidos pela Organização das Nações Unidas [Çelik et al., 2022]. Entretanto, de acordo com [Morales et al., 2010], tal aumento traz desafios cruciais para o planejamento e a operação dos sistemas de energia, devido, sobretudo, à intermitência das renováveis e ao fato de, geralmente, não serem despacháveis. Neste sentido, [Pinson, 2013] destaca que um dos principais desafios do setor envolve capturar a estocasticidade das ERVs, visando modelar e simular cenários realistas das gerações renováveis, constituindo uma ferramenta poderosa no apoio à tomada de decisão, tanto no setor público quanto no privado.

Nos últimos anos, muitas pesquisas têm sido direcionadas para modelar, prever ou simular ERVs. Nesse contexto, a avaliação dos efeitos complementares entre diferentes recursos renováveis é um tópico de relevância no setor, dada a obtenção de potências combinadas com menor variabilidade e intermitência [Jurasz et al., 2021]. Outro tema relevante envolve examinar a influência de fenômenos climáticos, como o El Niño Oscilação do Sul (ENOS) e suas fases distintas (El Niño, La Niña e Neutra) nos recursos primários renováveis ([NCEP, 2024; Iizumi et al., 2014]). No Brasil, por exemplo, durante o El Niño, as regiões Sul e Sudeste, em geral, experimentam aumento no volume de chuvas, enquanto o Norte e o Nordeste costumam sofrer com mais secas [Murari et al., 2020]. Durante o La Niña, observam-se padrões opostos [Murari et al., 2020]. Além disso, nos períodos secos, prevalecem condições de céu limpo, resultando em maior disponibilidade de radiação solar para a produção fotovoltaica (FV), e vice-versa. Já a influência dos fenômenos no comportamento dos ventos tende a variar significativamente dentro da mesma região. No entanto, no NE brasileiro, por exemplo, esperam-se ventos mais fortes durante o El Niño em comparação com as demais fases [dos Santos e Santos e Silva, 2013]. Ainda em relação ao ENOS, observa-se que, desde a década de 1960, os índices de temperatura do Pacífico usados para detectar a ocorrência das diferentes fases do fenômeno têm mostrado oscilações mais extremas [NCEP, 2024], devido, sobretudo, às mudanças climáticas, intensificando os impactos do ENOS nos recursos naturais.

Dessa forma, torna-se essencial capturar e representar, nos cenários probabilísticos de ERVs, tanto as propriedades complementares entre fontes renováveis diferentes, quanto a influência dos fenômenos climáticos, como o ENOS. Nesse contexto, uma das possíveis classificações para os métodos de simulação ou previsão de fontes renováveis é a divisão em métodos paramétricos e não-paramétricos [Li et al., 2022]. A maioria dos modelos de séries temporais, por exemplo, é classificada como paramétrica, pois assume pressupostos estatísticos como a normalidade dos resíduos [Baldioti e Souza, 2018]. Contudo, uma possível desvantagem é que, na maior parte dos casos reais de fontes renováveis, os ruídos não podem ser tratados desse modo por apresentarem comportamentos de cauda intrinsicamente assimétricos [Baldioti e Souza, 2018].

Consequentemente, é possível observar um alto volume de publicações recentes nessa área de pesquisa envolvendo métodos não-paramétricos, além de técnicas híbridas ou combinadas. Nesse meio, se destacam modelos de *machine learning*, especialmente redes neurais e suas variações (consulte [Farah et al., 2022; Yuan et al., 2022; Alipour et al., 2020]). Sublinha-se que, embora possuam vantagens na representação da aleatoriedade das ERVs em comparação a outros métodos, modelos de aprendizado de máquina tendem a exigir grande esforço computacional e são geralmente difíceis de interpretar e controlar. Nessas circunstâncias, diversos trabalhos modelam os processos estocásticos de fontes renováveis a partir de Cadeias de Markov ([Baldioti e Souza, 2018; Almutairi et al., 2016; Maçaira et al., 2019]), os quais são mais simples de interpretar. Com base nessa abordagem, identificamos diversas oportunidades de contribuição científica e propomos um modelo inovador para a simulação estocástica de fontes renováveis.

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

Em suma, o método proposto neste artigo pode ser dividido em dois estágios principais: (i) modelagem de renováveis e (ii) simulação de cenários. As contribuições principais são resumidas a seguir, com mais detalhes presentes na Seção 2. No estágio de modelagem (i), são calculados estados de ERVs visando considerar três aspectos ao mesmo tempo: (a) diferentes efeitos complementares (espacial e temporal) entre as fontes, com os estados representando simultaneamente múltiplas ERVs; (b) a influência das fases do ENOS nos recursos primários de energia; (c) e a sazonalidade das renováveis. No estágio de simulação (ii), os cenários são obtidos a partir de três tipos de sorteios concatenados: (a) o primeiro é referente à fase do ENOS; (b) o segundo, ao estado/*cluster* de recursos renováveis baseados na fase ENOS definida anteriormente; (c) por fim, o terceiro passo envolve sortear um dos elementos de recursos primários pertencentes ao *cluster* definido no segundo passo.

Para validar a metodologia proposta, foram gerados cenários para dados de recursos hídricos, eólicos e solares cobrindo todo o território brasileiro. As séries sintéticas foram avaliadas a partir de técnicas estatísticas para verificar a aderência às características das séries históricas. Além disso, os resultados do método proposto foram comparados com cenários obtidos pelo modelo *benchmark* aplicado ao setor elétrico brasileiro. Assim, este artigo é estruturado da seguinte maneira: a Seção 2 apresenta o *framework* proposto para simular cenários de renováveis; ns Seção 3, os estudos de caso para teste da metodologia são apresentados e avaliados; por fim, a Seção 4 traz as conclusões do trabalho.

2. Framework de Simulação de Cenários de Renováveis

A Figura 1 resume o *framework* proposto, composto por dois estágios principais: modelagem de renováveis e simulação de cenários, cada um constituído por três passos, detalhados nas Seções 2.1 e 2.2, respectivamente.

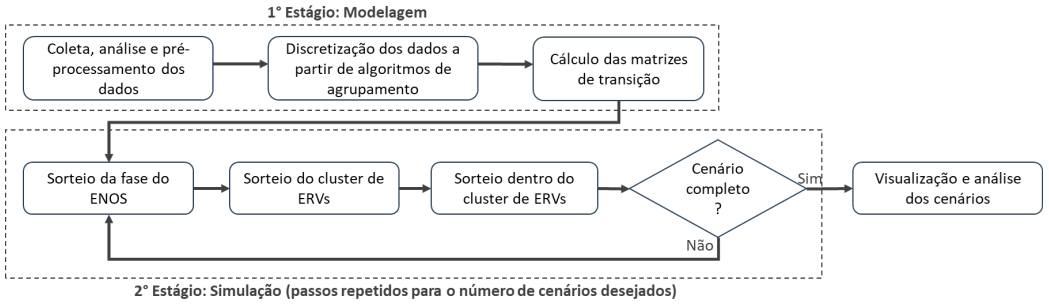


Figura 1: *Framework* de simulação de cenários de renováveis.

2.1. Modelagem de Renováveis

O primeiro passo da modelagem das renováveis consiste em obter, analisar e tratar os dados. Dependendo do interesse e disponibilidade dos mesmos, podem abranger tanto dados de recursos primários (por exemplo, velocidade do vento e radiação solar), quanto de geração de energia. Esse passo também envolve a análise descritíva dos dados e a aplicação de testes estatísticos, como o teste de Dickey-Fuller [Dickey e Fuller, 1979] para aferir acerca da estacionariedade dos dados.

Além disso, durante a etapa de pré-processamento, os dados devem ser divididos em classes, denominadas $c_{l,e,t}$. Cada $c_{l,e,t}$ compreende o conjunto de dados de recursos de uma localização l, na fase do ENOS e no tempo t. Considerando um banco de dados mensal, por exemplo, para compor todas as classes, os dados originais devem ser divididos entre todas as combinações possíveis de "fase do ENOS - mês". Logo, a classe $c_{l,neutra,janeiro}$ deve ser composta pelas observações históricas de recursos renováveis (tantos quantos forem considerados) do local l nos meses de janeiro cuja fase do ENOS foi neutra.

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

O segundo passo consiste em discretizar os dados e agrupá-los em estados finitos. Neste contexto, cada estado representa a disponibilidade coincidente de diferentes tipos de recursos em um momento específico. Assim, uma sequência de estados compreende a série temporal discretizada. Nesta etapa, algoritmos de *clustering*, como o k-means [MacQueen et al., 1967], podem ser utilizados devido à simplicidade e precisão [Almutairi et al., 2016]. Algoritmos de *clustering* têm como objetivo classificar objetos em vários *clusters* de forma com que os elementos do mesmo *cluster* sejam o mais semelhantes possível (alta similaridade intraclasse) [MacQueen et al., 1967]. Em contraste, objetos de *clusters* diferentes devem ser o mais diferentes possível (baixa similaridade interclasses) [MacQueen et al., 1967].

Assim, os dados multivariados (onde cada fonte de energia constitui uma variável) de cada classe ($c_{l,e,t}$) são discretizados em diferentes *clusters*. Uma vez que cada $c_{l,e,t}$ consiste em diferentes recursos, os estados obtidos visam capturar os efeitos de complementaridade entre as ERVs. Portanto, cada *cluster* calculado terá um número de dimensões ou atributos correspondente ao número de fontes renováveis. Além disso, o agrupamento é realizado de forma independente para cada classe, mantendo as características de cada localização l de interesse, no tempo t, para a fase do ENOS e. No final do processo de discretização, cada vetor de ERVs no tempo t deve pertencer a um *cluster* específico.

Na terceira etapa da modelagem, dois tipos de matrizes de transição de estados são calculadas: (i) entre as ERVs e (ii) entre as fases do ENOS. Detalhando as primeiras, com o número finito de estados definidos na etapa anterior, as matrizes entre estados de ERVs devem ser calculadas na mesma periodicidade, mantendo a consistência na modelagem. A probabilidade de transição estacionária ($p_{a,b}$) do estado a para o estado b, para todos os índices $1 \leq (a, b) \leq k$, é calculada de acordo com: $p_{a,b} = (n_{a,b} / \sum_{i=1}^{k} n_{a,i})$, onde $n_{a,b}$ é o número de transições históricas de a para b, e o denominador é o total de transições partindo de a para todos os estados de destino possíveis. Após obter as probabilidades de transição, calculam-se as matrizes a partir de (1). Na equação, l representa um determinado local, t é o tempo de origem e $(t + 1)$, de destino. Além disso, considera-se que, no tempo t, a fase do ENOS é e_i, e no tempo $(t + 1)$, e_j. Assim, $(i, j) \in$ {Neutro, El Niño, La Niña}.

$$P^{l}_{t_{e_i} \to (t+1)_{e_j}} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,k} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{w,1} & p_{w,2} & \cdots & p_{w,k} \end{bmatrix} \tag{1}$$

Na matriz de transição, as linhas correspondem aos possíveis estados de origem (w no total), enquanto as colunas se referem aos potenciais estados de destino (k no total). Assim, $p_{1,1}$, por exemplo, representa a probabilidade de transição do estado 1 no período t na fase do ENOS e_i para o estado 1 no período $(t + 1)$ na fase do ENOS e_j. Lembrando que são estados distintos, dado que, como detalhado na etapa de agrupamento, os *clusters* são calculados independentemente para cada combinação de "fase do ENOS - período". Além disso, cada estado tem uma dimensão para cada ERV considerada. Assim, se duas fontes estão sendo modeladas (por exemplo, eólica e fotovoltaica), $p_{1,1}$ será a probabilidade de transição de um estado bivariado (eólico-FV) para outro estado bivariado (eólico-FV). No final, o número de matrizes de transição para um determinado local deve corresponder ao número total de combinações possíveis para o par "fase do ENOS - período".

A partir de cada matriz obtida por (1), calcula-se sua versão acumulada, na qual as probabilidades de cada linha em (1) são somadas uma a uma até a última coluna. Consequentemente, todos os valores na última coluna da matriz acumulada devem ser iguais a 1, pois representam a

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

soma de todas as probabilidades de transição possíveis partindo do estado da linha correspondente. As matrizes acumuladas serão utilizadas oportunamente no estágio de simulação.

Quanto às matrizes entre as fases do ENOS, a estratégia para calculá-las é semelhante ao caso anterior, com base nas transições históricas. Todavia, desta vez, não há necessidade de aplicar algoritmos de *clustering* para identificar os estados do ENOS, os quais são três por definição do fenômeno (El Niño, La Niña e Neutro). Tais matrizes também devem ter sua versão acumulada ($P_{t \to t+1}^{ENOS,acum}$). Além disso, são independentes de qualquer localização, pois as fases do ENOS não se alteram de acordo com a região. Finalmente, são usadas na primeira etapa da simulação de cenários, detalhada na subseção a seguir.

2.2. Simulação de Cenários

O primeiro passo do estágio de simulação envolve o sorteio da fase ENOS com base na matriz de transição acumulada ($P_{t \to t+1}^{ENOS,accum}$) entre as fases do fenômeno calculada na etapa anterior. Para sortear o primeiro valor de cada cenário, a última fase do ENOS do histórico é considerada como o estado de origem. Em seguida, um número aleatório de uma distribuição uniforme contínua $[0, 1]$ é sorteado e comparado com os elementos da linha correspondente ao estado de origem na matriz acumulada. Suponha que o número aleatório seja maior que a probabilidade acumulada do estado anterior e menor ou igual à probabilidade acumulada do estado subsequente. Nesse caso, este estado é escolhido para representar a próxima fase ENOS e será usado na próxima etapa da simulação. O pseudo-código da Fig. 2 representa esse processo.

```
Begin
  Unif ~ U[0,1]
  Line = State(t)
  Column=1
  Repeat
    If (matrix(Line,Column)<Unif and Unif<=matrix(Line,Column+1)) then
      State(t+1) = Column+1
      Column = (No. columns matrix) + 1
    Else
      Column = Column + 1
  While (Column <= (No. columns matrix))
End
```

Figura 2: Pseudo-código para simulação baseada em matrizes de transição de estados.

O segundo passo consiste em sortear um dos *clusters* a partir da matriz de transição acumulada entre as ERVs ($P_{t_{e_i} \to (t+1)_{e_j}}^{l,acum}$). Para tal, deve-se, primeiro, selecionar a matriz correta com base na fase do ENOS no tempo t e na fase sorteada e_j no passo anterior para o tempo $(t + 1)$. O último estado do histórico é considerado como estado de origem caso ainda estejamos na primeira iteração da simulação. Assim, um novo estado é sorteado para o período $(t + 1)$, a partir do mesmo mecanismo mostrado na Fig. 2.

O terceiro e último passo consiste em sortear aleatoriamente um dos elementos pertencentes ao *cluster* sorteado no passo anterior. Os elementos possíveis foram definidos na etapa de agrupamento da série temporal e cada um possui o número de dimensões correspondente ao número de fontes renováveis do estudo. Dessa forma, teremos, finalmente, as ERVs simuladas para o período $(t + 1)$.

Os três passos de simulação devem ser repetidos até que o horizonte de cada cenário seja alcançado. Finalmente, a metodologia proposta deve ser avaliada, verificando o desempenho e a aderência às características históricas. Para isso, comparamos: (i) as métricas históricas e simuladas, incluindo a aplicação de testes de hipóteses; e (iii) o desempenho do modelo proposto com o

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

benchmark. A Subseção 2.3 apresenta o modelo considerado *benchmark* no contexto da aplicação desenvolvida neste trabalho.

2.3. Modelo *Benchmark*

Para selecionar o *benchmark* correto, é essencial observar a granularidade temporal da série e o horizonte de simulação. Nesse sentido, a metodologia proposta foi aplicada a séries mensais brasileiras para gerar cenários de médio e longo prazo (Seção 3). A partir disso, o modelo de referência usado para o planejamento do Sistema Elétrico Brasileiro (SEB) para simulações mensais são os autoregressivos periódicos, ou PAR(p) [Ferreira et al., 2015; de Castro et al., 2015; Maçaira et al., 2017].

Em geral, séries mensais de ERVs são caracterizadas pelo comportamento periódico de suas propriedades probabilísticas, como média, variância, assimetria e estrutura de autocorrelação. Assim, o modelo PAR(p) ajusta um modelo AR(p) para cada período da série, onde p é a ordem do modelo, ou seja, o número de termos autorregressivos. Em geral, p é um vetor onde cada elemento indica a ordem de cada período. Por exemplo, $p = [p_1,\ p_2,\ p_3, \ldots,\ p_{12}]$, no caso de séries mensais. Uma série Z com S períodos que se repetem por N anos pode ser representada por $Z = [z_{(1,1)}, z_{(1,2)}, \ldots, z_{(1,S)}, \ldots, z_{(N,S)}]$. O modelo PAR dessa série Z no período m é dado por 2:

$$\left(\frac{z_{(t,m)} - \mu_m}{\sigma_m}\right) = \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t,m-i)} - \mu_{m-i}}{\sigma_{m-i}}\right) + u_{t,m} \tag{2}$$

Onde μ_m é a média do período m; σ_m é o desvio padrão do período m; p_m é a ordem autorregressiva do período m; $\varphi_i^{(m)}$ representa o i-ésimo coeficiente autorregressivo no período m; e $u_{t,m}$ é a série de resíduos independentes (média zero e desvio padrão = σ_m^u).

3. Estudos de Caso

3.1. Base de Dados

Para validar a metodologia proposta, foram gerados cenários para diferentes regiões do território brasileiro visando obter resultados robustos e generalizados, considerando três recursos renováveis - hídrico, eólico e FV, os quais representam mais de 80% da matriz elétrica do Brasil [ONS, 2024]. Os dados foram obtidos no nível espacial relativo às bacias hidrográficas do país, sendo que 17 bacias cobrem todo o território nacional. Além disso, correspondem a séries temporais mensais de recursos primários, ou seja, vazão incremental, velocidade do vento e radiação solar, com o objetivo de realizar simulações de médio prazo. É importante destacar que o objetivo dos estudos de caso não consiste em transformar recursos primários em energia, mas avaliar se a metodologia proposta captura a aleatoriedade das ERVs e suas características complementares para gerar cenários representativos no Brasil.

Os dados hídricos das bacias foram obtidos do NEWAVE [CCEE, 2024], ferramenta utilizada para planejamento da operação do SEB. Nota-se que dados de duas das 17 bacias não estão disponíveis no NEWAVE, portanto, não foram consideradas nesta aplicação. As demais variáveis, velocidade do vento a 100 m de altura e radiação solar superficial, foram extraídas da base de reanálise Merra-2 da [NASA, 2024], uma alternativa amplamente utilizada pela academia e por profissionais do mercado quando há falta ou limitação de dados medidos para recursos renováveis [Jurasz et al., 2021]. Cada bacia cobre uma região com vários pontos da rede Merra-2, portanto, foi obtida uma média mensal de cada recurso eólico e FV para obter uma única série por bacia. Embora as séries de vazão incremental estejam disponíveis a partir de 1931 no NEWAVE, as demais estão disponíveis apenas a partir de janeiro de 1980 no Merra-2. Como todas devem ter a mesma janela temporal, os dados hídricos anteriores a 1980 foram descartados. A última observação das séries corresponde a dezembro de 2022, totalizando 42 anos completos, ou 504 observações mensais.

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

A Fig. 3 mostra um exemplo dos três recursos da bacia Grande e suas respectivas funções de autocorrelação (FACs). Os gráficos das outras bacias podem ser acessados em um repositório do GitHub através deste link, bem como os resultados completos de todos os estudos de caso desenvolvidos neste trabalho. As séries temporais, em geral, apresentam um comportamento sazonal bem definido. Outra observação relevante está relacionada aos *lags* das FACs em que os valores de autocorrelação se destacam em comparação aos demais. Em todos os casos, os *lags* múltiplos de 12 se destacam com autocorrelações positivas, já que as ERVs tendem a se comportar de maneira semelhante nos mesmos meses devido às estações do ano. Por outro lado, a cada 6 *lags*, os sinais das autocorrelações se invertem, já que os comportamentos tendem a ser contrários nas estações opostas. Além disso, as séries temporais das ERVs devem ser estacionárias para que a modelagem possa ser realizada usando matrizes de transição de estados. Para isso, o teste Augmented Dickey-Fuller (ADF) [Dickey e Fuller, 1979] foi aplicado, e, para um nível de significância de 5%, a hipótese nula de não estacionariedade foi rejeitada para todos os casos. Por fim, para identifição das fases do ENOS mês a mês, foi utilizado o índice ONI [NOAA, 2024].

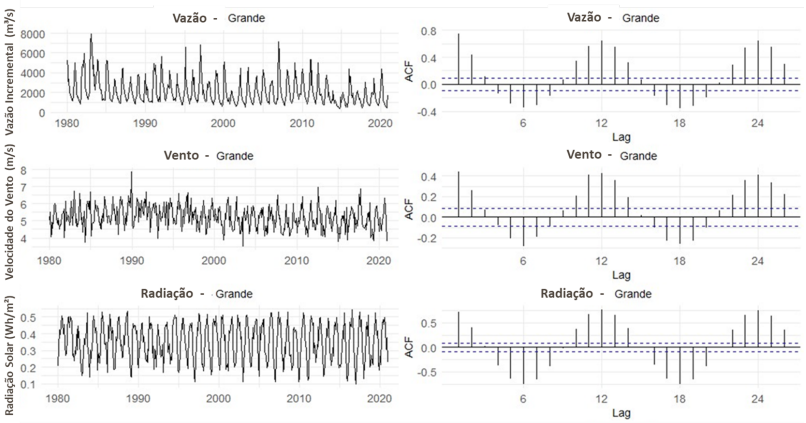


Figura 3: Séries temporais e FACs - Estudo de Caso '*Grande*'.

3.2. Modelagem de Renováveis

Todas as etapas do desenvolvimento dos estudos de caso foram realizadas na linguagem de programação R [R Core Team, 2021]. Devido às limitações de espaço, apresentamos um exemplo de cada etapa. Os outros resultados estão disponíveis no mesmo link para o repositório do GitHub apresentado na seção anterior. Além disso, os dados dos últimos dois anos do histórico (2021 e 2022) foram separados como período *out-of-sample* para verificação da capacidade de simulação de ambas os modelos, proposto e *benchmark*.

Iniciando com as matrizes de transição de estados do fenômeno climático calculadas a partir do histórico do índice ONI, a Tabela 1 apresenta, como exempplo, a matriz de transição de fevereiro para março. Existem 12 matrizes desse tipo, correspondendo ao número total de transições possíveis entre os meses. Observa-se que não há probabilidades de transição direta entre as fases El Niño e La Niña, em conformidade com a definição do índice do fenômeno, que deve sempre passar por um período neutro entre as duas fases extremas.

Prosseguindo com a modelagem das ERVs, os *clusters* e as matrizes de transição também foram calculados por mês. Todas as séries foram normalizadas antes da etapa de agrupamento para eliminar a influência da dimensionalidade dos dados. Para cada bacia, aplicamos o algoritmo k-means 36 vezes, cobrindo todas as combinações dos 12 meses do ano com as três fases do ENOS, e o número de *clusters* para cada uma das 36 classes variou de dois a cinco. Esse número de

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

Tabela 1: Matriz de Transição de Fases do ENOS de Fevereiro para Março - $P^{ENOS}_{Fevereiro \rightarrow Maro}$.

Estado	Neutro	El Niño	La Niña
Neutro	0.941	0.059	0.000
El Niño	0.333	0.667	0.000
La Niña	0.083	0.000	0.917

clusters possibilitou, em todos os casos, a estimação das matrizes de transição a partir dos dados históricos disponíveis. Quanto ao número de matrizes de transição mensais por bacia, há 84 no total, 12 para cada possível transição de fase, ou seja: (i) entre os meses da mesma fase do ENOS: Neutro-Neutro, El Niño-El Niño e La Niña-La Niña, totalizando 36 matrizes; (ii) entre os meses de diferentes fases do ENOS: Neutro-El Niño, Neutro-La Niña, El Niño-Neutro e La Niña-Neutro, totalizando 48 matrizes. A Tabela 2 mostra a matriz de transição entre as ERVs de "Janeiro/Fase Neutra" para "Fevereiro/Fase Neutra" para o caso da bacia '*Grande*'.

Tabela 2: Matriz de Transição de ERVs ($P^{Grande}_{Janeiro_{(Neutro)} \rightarrow Fevereiro_{(Neutro)}}$) - Estudo de Caso '*Grande*'.

Janeiro - Neutro	**Fevereiro - Neutro**			
	$1_{Fev,Neut}$	$2_{Fev,Neut}$	$3_{Fev,Neut}$	$4_{Fev,Neut}$
$1_{Jan,Neut}$	0.20	0.00	0.60	0.20
$2_{Jan,Neut}$	0.25	0.00	0.25	0.50
$3_{Jan,Neut}$	0.00	0.50	0.50	0.00
$4_{Jan,Neut}$	0.00	0.00	0.00	1.00

Embora tanto "Janeiro-Neutro" quanto "Fevereiro-Neutro" apresentem quatro estados de ERVs cada, é importante observar que não são os mesmos estados, uma vez que o agrupamento de cada conjunto de dados "mês-fase" foi realizado independentemente. Além disso, cada estado pode ser representado pelo centroide do *cluster* ao qual se refere, refletindo, assim, as três fontes renováveis simultaneamente, já que cada centroide é um vetor com três dimensões ou atributos, um para cada ERV. Dessa forma, exploram-se os efeitos complementares entre as fontes renováveis. Como exemplo, o estado 1 de "Janeiro-Neutro" é representado pelo seguinte centroide: [6827 m^3/s, 5.01 m/s, 0.11 Wh/m^2], sendo o primeiro atributo referente à vazão incremental, o segundo, à velocidade do vento e o terceiro, à radiação solar. Já o estado 1 de "Fevereiro-Neutro" tem o seguinte centroide: [4723 m^3/s, 5.29 m/s, 0.22 Wh/m^2]. Todas as matrizes de transição, quer entre as fases do ENOS ou entre as ERVs, também possuem sua versão de probabilidades acumuladas, efetivamente utilizadas nos passos de simulação, detalhados na próxima subseção.

3.3. Simulação de Cenários

Para cada bacia, foram gerados 2000 cenários de 60 meses, equivalente a cinco anos completos, com base no horizonte de simulação do NEWAVE. Para simular o primeiro estado de cada cenário, a última observação histórica foi utilizada como estado predecessor ou de origem. De acordo com a Fig. 1, sorteia-se uma fase para o ENOS no primeiro passo. O último mês dos dados *in-sample* foi dezembro de 2020, cuja fase do ENOS foi La Niña. Portanto, os cenários começam em janeiro e, para tal, uma fase do ENOS é sorteada a partir da matriz acumulada de dezembro para janeiro ($P^{ENOS,acum}_{Dez \rightarrow Jan}$). Assim, um valor aleatório de uma uniforme [0,1] é gerado via Monte Carlo. Com o número aleatório, a próxima fase do fenômeno é identificada a partir da linha correspondente ao estado predecessor (La Niña), conforme a Fig. 2.

O segundo passo é o sorteio do *cluster* de ERVs. Com a fase anterior do ENOS sorteada, o *cluster* de recursos renováveis para janeiro é sorteado usando a matriz de transição correta entre

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

ERVs. Para cada bacia, essa matriz deve corresponder à transição de recursos de "Dezembro-La Niña" para "Janeiro-Fase sorteada primeiramente". O processo de sorteio é semelhante ao primeiro passo, com o estado predecessor correspondendo ao último estado de ERVs do histórico. No terceiro e último passo, com o *cluster* de ERVs para janeiro definido, um dos vetores de recursos pertencentes ao *cluster* é sorteado aleatoriamente. Seria possível usar o centroide do *cluster* para representar os valores finais de ERVs, mas para introduzir maior variabilidade aos cenários, este terceiro passo foi adicionado, aumentando as possibilidades de valores de ERVs na série sintética. Os três passos são, então, repetidos, até completar o horizonte de cada cenário.

Seguem duas observações relevantes: (i) as fases do ENOS sorteadas no primeiro passo são as mesmas para todas as bacias em todos os cenários, mantendo a correlação entre as diferentes regiões em relação ao fenômeno - as diferenças ocorrem a partir do segundo passo com as matrizes de transição de ERVs e (ii) para que as fases extremas do ENOS ocorram, deve haver pelo menos cinco meses consecutivos de aquecimento (El Niño) ou resfriamento (La Niña) das águas do Oceano Pacífico na região de medição do índice ONI. Portanto, para representar essa especificidade do fenômeno nos cenários, a mesma fase é mantida por, ao menos, cinco meses ao entrar em uma das duas fases, interrompendo a primeira etapa da simulação durante o mesmo período. Então, após cinco meses na mesma fase, retorna-se com o sorteio do primeiro passo, possibilitando a continuação ou não do estado atual do ENOS. Se ocorrer uma saída e um subsequente retorno a uma das duas fases (El Niño ou La Niña), o prazo de cinco meses deve ser cumprido novamente.

3.4. Análise de Cenários

Iniciando pela comparação mensal entre as frequências históricas e simuladas para as fases do ENOS, observamos, na Figura 4, que as fases do fenômeno apresentam sazonalidade, a qual foi bem reproduzida pelas simulações.

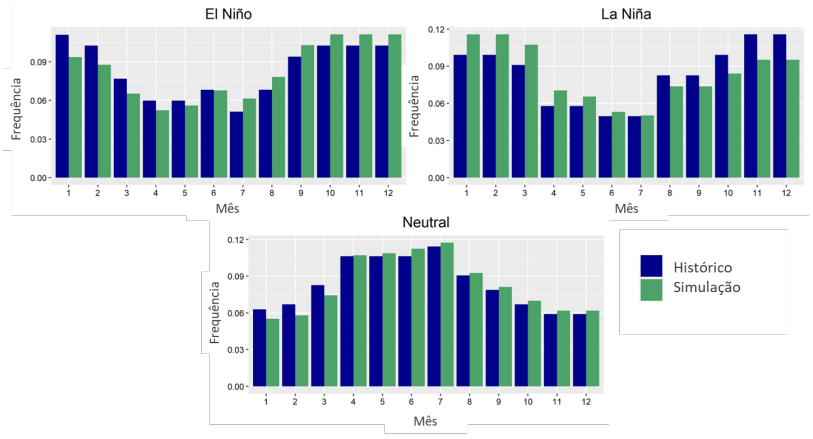


Figura 4: Histogramas das fases do ENOS.

Prosseguindo para a avaliação das ERVs, o teste de Kolmogorov-Smirnov [Massey Jr, 1951] foi aplicado para comparar, com robustez estatística, se as distribuições dos dados simulados seguem as distribuições históricas por fase do fenômeno. A um nível de significância de 10%, os resultados indicam que, para as três ERVs, nas três fases do fenômeno, considerando todas as bacias, os cenários simulados apresentam a mesma distribuição histórica em 98% dos casos.

Também é importante validar as simulações comparando com os resultados gerados pelo *benchmark* do setor. Assim, considerando as 15 bacias e os 12 meses do ano, foram estimados 180 modelos PAR(p). A ordem de cada modelo foi selecionada com base na função de autocorrelação

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

Tabela 3: Comparação de métricas estatísticas – Estudo de Caso'*Grande*'.

ERV	Dado	Média	Mediana	DP	Assimetria	Curtose
Hidro	Hist	2098	1702	1298	1.36	4.79
	Bench	1971 (1407, 2606)	1607 (1134, 2193)	1232 (825, 1659)	1.46 (0.70, 2.41)	5.60 (2.39, 11.3)
	Prop	2105 (1824, 2482)	1698 (1416, 1981)	1300 (939, 1749)	1.36 (0.66, 2.04)	4.76 (2.30, 8.08)
	Bench Disc	6.1%	5.6%	5.1%	7.4%	16.9%
	Prop Disc	0.3%	0.3%	0.2%	0.4%	0.7%
Eol	Hist	5.22	5.18	0.64	0.21	3.01
	Bench	5.23 (5.01, 5.47)	5.22 (4.94, 5.49)	0.63 (0.51, 0.76)	0.12 (-0.50, 0.83)	2.76 (1.91, 4.29)
	Prop	5.21 (5.06, 5.36)	5.17 (4.96, 5.41)	0.64 (0.52, 0.79)	0.23 (-0.36, 1.11)	3.02 (1.96, 5.81)
	Bench Disc	0.3%	0.8%	0.5%	41.6%	8.0%
	Prop Disc	0.2%	0.1%	0.4%	13.8%	0.4%
FV	Hist	0.35	0.36	0.11	-0.23	1.91
	Bench	0.35 (0.32, 0.38)	0.36 (0.32, 0.40)	0.11 (0.10, 0.12)	-0.17 (-0.51, 0.16)	1.99 (1.60, 2.61)
	Prop	0.35 (0.33, 0.37)	0.36 (0.32, 0.39)	0.11 (0.10, 0.13)	-0.23 (-0.52, 0.05)	1.88 (1.56, 2.34)
	Bench Disc	0.5%	1.6%	0.7%	24.2%	4.2%
	Prop Disc	0.3%	0.1%	1.7%	0.7%	1.4%

parcial das séries históricas, variando de 1 a 6. Para cada bacia, geramos o mesmo número de cenários obtidos a partir da metodologia proposta.

Usando a bacia '*Grande*' novamente como exemplo, a Tabela 3 apresenta diversas medidas estatísticas dos dados históricos *in-sample* ($Hist$) e dos cenários obtidos, tanto pela metodologia proposta ($Prop$), quanto pelo *benchmark* ($Bench$). Para as métricas simuladas, também apresenta, entre parênteses, o 1º e o 99º percentis dos cenários. Por fim, a tabela mostra as discrepâncias ($disc$) entre as simulações e os dados históricos. A discrepância é definida como a diferença percentual absoluta entre o valor da métrica histórica e a média simulada para a mesma métrica. Assim, *disc* é calculada por $(|(\textit{Histórico} - \textit{Simulado})/\textit{Histórico}|) \times 100$. Analisando todas as bacias, em 80% das métricas, os resultados da metodologia proposta são mais precisos, ou seja, apresentam discrepâncias menores (*Prop Disc*) em comparação com o *benchmark* (*Bench Disc*).

Analisando ainda a Tabela 3, observamos que os percentis do *benchmark* apresentam valores mais extremos em relação ao modelo proposto na maioria dos casos. Neste sentido, o PAR(p) é criticado por não representar bem a assimetria das ERVs, ao passo que a proposta apresenta um desempenho superior. Quando mal estimada, a assimetria prejudica a representação da cauda das distribuições, podendo superestimar ou subestimar os casos extremos de ERVs. Isso resulta em uma representação inadequada do risco dos ativos de energia renovável.

De acordo com os critérios pré-definidos, os últimos dois anos de dados foram separados como período *out-of-sample*. A Figura 5 apresenta, para cada ERV da bacia '*Grande*', os primeiros dois anos dos cenários, tanto das simulações do *benchmark*, quanto da proposta, além dos dados observados em 2021 e 2022. Notavelmente, embora ambas as metodologias reproduzam o padrão sazonal observado, o PAR(p) apresenta cenários mais extremos, corroborando a análise *in-sample*, visto que alonga as caudas das renováveis, frequentemente simulando valores irreais. A metodologia proposta, geralmente, mantém os cenários dentro de uma faixa mais estreita e realista. Observa-se que padrões semelhantes são apresentados para as demais bacias (link do GitHub).

Destaca-se que, nos anos 2021 e 2022, houve prevalência da fase La Niña, de acordo com o índice ONI - 83% dos meses. Na região da bacia '*Grande*', o La Niña tende a reduzir os volumes de chuva, aumentar a incidência solar e pode trazer os padrões de vento para níveis mais baixos. Assim, observa-se que, em comparação com a média *in-sample*, 2021 e 2022 apresentaram níveis

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

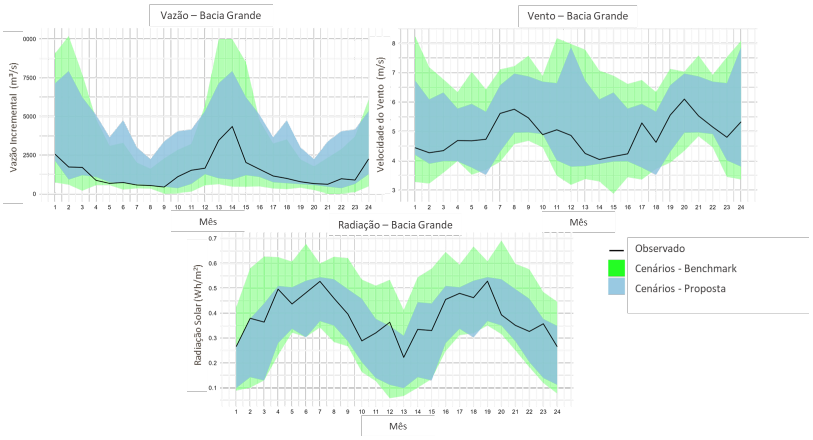


Figura 5: Cenários e dados observados (2021-2022) - Estudo de Caso '*Grande*'.

mais baixos para vazão incremental e para velocidade do vento, além do oposto para a radiação solar. Ao mesmo tempo, os dois primeiros anos dos cenários da metodologia proposta apresentaram 42% dos meses em La Niña, 41% na fase Neutra e 17% em El Niño. Apesar de uma parte significativa dos meses ter sido simulada em La Niña, a frequência foi menor do que a observada, justificando os níveis médios sintéticos mais altos para os recursos hídricos e eólicos, em comparação com os dados observados, além do padrão oposto para a fonte FV. Por fim, destaca-se que simulamos com precisão a frequência das fases do ENOS no período *in-sample* (Fig. 4), mas, visando aumentar a precisão *out-of-sample* em trabalhos futuros, recomenda-se usar previsões das fases do ENOS disponibilizadas pelo NOAA [NOAA, 2024]. No entanto, sublinhamos que o horizonte de previsão se restringe a nove meses.

4. Conclusões

Este trabalho propõe uma metodologia para simulação de cenários de ERVs, visando capturar a complementaridade entre diferentes fontes, além da influência de fenômenos climáticos, como o ENOS, nos recursos renováveis. A metodologia é baseada em cadeias de Markov e inova ao propor estados multidimensionais, representando mais de uma fonte simultaneamente, além de utilizar informações exógenas relevantes, como as diferentes fases do ENOS e seus impactos nas ERVs.

Vale ressaltar que os índices mensais de temperatura do Pacífico utilizados para identificar as fases do ENOS têm apresentado níveis de oscilações mais extremos nas últimas seis décadas. Portanto, como utilizamos dados desse período, consideramos esse novo comportamento nos cenários. Tais padrões são pouco prováveis de mudar em cenários de médio prazo, como no caso deste estudo. No entanto, adaptações na metodologia são necessárias se o interesse estiver em cenários representativos para décadas à frente, ou ainda, em simulações que englobem mudanças de curtíssimo prazo.

Finalmente, para avaliar a proposta, foram utilizados dados mensais de três ERVs abrangendo todo o território brasileiro. Os resultados obtidos foram satisfatórios e, em comparação ao modelo considerado *benchmark* no Brasil, apresentaram maior acurácia. Como sugestões para trabalhos futuros, destacamos: (i) aplicar os cenários em modelos de otimização de planejamento energético; (ii) validar a metodologia com dados horários, avaliando se os cenários reproduzem os padrões de outras granularidades temporais; (iii) adaptar a metodologia para representar tendências climáticas de longo prazo ou oscilações de curtíssimo prazo nos estados de ERVs; (iv) testar outras técnicas de agrupamento na fase de modelagem; (v) utilizar previsões oficiais do ENOS nos meses

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

iniciais dos cenários.

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, 210.041/2023 e 210.015/2024] 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

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*, 205:118106.

Almutairi, A., Ahmed, M. H., e Salama, M. (2016). Use of mcmc to incorporate a wind power model for the evaluation of generating capacity adequacy. *Electric Power Systems Research*, 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. In *2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, p. 1–6. IEEE.

CCEE (2024). Available: `https://www.ccee.org.br/acervo-ccee?especie=44884&periodo=365`.

Çelik, D., Meral, M. E., e Waseem, M. (2022). Investigation and analysis of effective approaches, opportunities, bottlenecks and future potential capabilities for digitalization of energy systems and sustainable development goals. *Electric Power Systems Research*, 211:108251.

de Castro, C. M. B., Marcato, A. L. M., Souza, R. C., Junior, I. C. S., Oliveira, F. L. C., e Pulinho, T. (2015). The generation of synthetic inflows via bootstrap to increase the energy efficiency of long-term hydrothermal dispatches. *Electric Power Systems Research*, 124:33–46.

Dickey, D. A. e Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a):427–431.

dos Santos, A. T. S. e Santos e Silva, C. M. (2013). Seasonality, interannual variability, and linear tendency of wind speeds in the northeast brazil from 1986 to 2011. *The Scientific World Journal*, 2013.

Farah, S., Humaira, N., Aneela, Z., Steffen, E., et al. (2022). Short-term multi-hour ahead country-wide wind power prediction for germany using gated recurrent unit deep learning. *Renewable and Sustainable Energy Reviews*, 167:112700.

Ferreira, P. G. C., Oliveira, F. L. C., e Souza, R. C. (2015). The stochastic effects on the brazilian electrical sector. *Energy Economics*, 49:328–335.

Iizumi, T., Luo, J.-J., Challinor, A. J., Sakurai, G., Yokozawa, M., Sakuma, H., Brown, M. E., e Yamagata, T. (2014). Impacts of el niño southern oscillation on the global yields of major crops. *Nature communications*, 5(1):3712.

Jurasz, J., Mikulik, J., Dabek, P. B., Guezgouz, M., e Kaźmierczak, B. (2021). Complementarity and 'resource droughts' of solar and wind energy in poland: an era5-based analysis. *Energies*, 14(4):1118.

Li, H., Ren, Z., Fan, M., Li, W., Xu, Y., Jiang, Y., e Xia, W. (2022). A review of scenario analysis methods in planning and operation of modern power systems: Methodologies, applications, and challenges. *Electric Power Systems Research*, 205:107722.

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br

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*, 12(5):826.

Maçaira, P. M., Oliveira, F. L. C., Ferreira, P. G. C., Almeida, F. V. N. d., e Souza, R. C. (2017). Introducing a causal par (p) model to evaluate the influence of climate variables in reservoir inflows: a brazilian case. *Pesquisa Operacional*, 37:107–128.

MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, p. 281–297. Oakland, CA, USA.

Massey Jr, F. J. (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78.

Morales, J. M., Minguez, R., e Conejo, A. J. (2010). A methodology to generate statistically dependent wind speed scenarios. *Applied Energy*, 87(3):843–855.

Murari, T. B., Filho, A. S. N., Moret, M. A., Pitombo, S., e Santos, A. A. (2020). Self-affine analysis of enso in solar radiation. *Energies*, 13(18):4816.

NASA (2024). Available: `https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/`.

NCEP (2024). Available: `https://origin.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/enso.shtml`.

NOAA (2024). Available: `https://www.ncei.noaa.gov/access/monitoring/enso/sst`.

ONS (2024). Available: `https://www.ons.org.br/paginas/resultados-da-operacao/historico-da-operacao/dados-gerais`.

Pinson, P. (2013). Wind energy: Forecasting challenges for its operational management. *Statistical Science*, 28(4):564–585.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL `https://www.R-project.org/`.

Yuan, R., Wang, B., Sun, Y., Song, X., e Watada, J. (2022). Conditional style-based generative adversarial networks for renewable scenario generation. *IEEE Transactions on Power Systems*, 38(2):1281–1296.

DOI: 10.59254/sbpo-2024-193840
https://proceedings.science/p/193840?lang=pt-br