

CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ
ISSN on-line 2317-4535 / ISSN impresso 1413-9022 - v. 53, p.24 - 42, 2023
DOI: 10.12957/cadest.2023.84431

COMBINANDO PRECISÃO E ESFORÇO COMPUTACIONAL NO AJUSTE DE DISTRIBUIÇÕES WEIBULL TRIPARAMÉTRICAS COM ELEVADAS ASSIMETRIAS PARA GERAÇÃO DE CENÁRIOS DE VENTOS CORRELACIONADOS COM AFLUÊNCIAS

Albert C.G. Melo

Universidade do Estado do Rio de Janeiro e Centro de Pesquisas de Energia Elétrica
albert.melo@ime.uerj.br

Maria Elvira P. Maceira

Universidade do Estado do Rio de Janeiro e Centro de Pesquisas de Energia Elétrica
melvira@ime.uerj.br

José Francisco M. Pessanha

Universidade do Estado do Rio de Janeiro e Centro de Pesquisas de Energia Elétrica
pessanha@ime.uerj.br

Resumo

Propõe-se uma modelagem das velocidades mensais de ventos através de distribuições Weibull triparamétricas, na geração de cenários de ventos mensais correlacionados com afluições às usinas hidrelétricas, em um esquema de Programação Dinâmica Dual Estocástica-PDDE. A abordagem constitui-se de duas etapas. Na primeira, os parâmetros da distribuição Weibull dos resíduos das velocidades mensais históricas do vento são estimados pelo Método dos Momentos de 3ª Ordem. Se a assimetria das velocidades sintéticas for suficientemente próxima da histórica, o algoritmo para. Caso contrário, a estimativa do parâmetro posição, obtida na primeira etapa é utilizada como valor inicial para calcular estimativas dos outros dois parâmetros (forma e escala) por meio do Método dos Momentos de 2ª Ordem. Em seguida, de forma iterativa, o algoritmo atualiza a estimativa inicial do parâmetro posição com vistas a reduzir a diferença entre as assimetrias das velocidades sintéticas e históricas do vento, até que se atinja um valor considerado aceitável. Procura-se, assim, assegurar precisão com o menor esforço computacional possível. A aplicação a parques eólicos do Nordeste e Sul do Brasil, em diferentes meses e assimetrias, aponta para a sua robustez, apresentando desempenho sempre superior a 17 outros métodos avaliados.

Palavras-chave: Energia eólica; Geração de séries sintéticas; Distribuição de Weibull; Método dos momentos; Planejamento e operação; Otimização estocástica.

1. Introdução

Em países que apresentam alta participação de energia hidrelétrica, o planejamento da expansão e, principalmente, da operação é baseado em modelos de otimização e desagregado em problemas com horizontes específicos (MACEIRA *et al.*, 2002; HELSETH & MELO, 2020, HELSETH *et al.*, 2023).

No caso do Brasil, o problema de planejamento da expansão e da operação do Sistema Interligado Nacional (SIN) é dividido em planejamento da expansão (longo prazo), planejamento da operação (médio e curto prazos), e programação da operação, sendo resolvido através de uma cadeia de modelos computacionais (MACEIRA *et al.*, 2002; HELSETH & MELO, 2020, HELSETH *et al.*, 2023). Além disso, a matriz elétrica brasileira apresenta alta participação de fontes renováveis, respondendo por mais de 80% do consumo de energia. As fontes intermitentes, em especial a eólica, têm experimentado um crescimento acelerado—e.g., na última década, a energia eólica cresceu 13 vezes, atingindo hoje 22,3 GW de capacidade instalada em cerca de 830 parques eólicos e tornou-se a segunda maior fonte na matriz elétrica (12%), atrás apenas da hidrelétrica (60%), tendência esperada nas próximas décadas.

Um dos modelos-chave desta cadeia para as atividades de planejamento da expansão e da operação energética do sistema interligado é o modelo NEWAVE, que desde 1998 vem sendo utilizado no setor elétrico brasileiro (MACEIRA *et al.*, 2018), (MACEIRA *et al.*, 2019).

Desde a sua primeira versão, o modelo NEWAVE vem aplicando e aprimorando algoritmos de programação dinâmica dual estocástica – PDDE (PEREIRA & PINTO, 1991), (MACEIRA, 1993) para resolver o problema de planejamento da operação de longo e médio prazos. No algoritmo PDDE, as variáveis que caracterizam o estado do sistema são o armazenamento dos reservatórios e a tendência hidrológica, representada pelas afluências passadas às usinas hidrelétricas ou reservatórios equivalentes de energia (REEs). Um esquema de simulação de Monte Carlo é usado para construir iterativamente funções de custo futuro multivariadas, que representam a política de operação ótima. Além disso, após a política de operação ter sido construída, uma simulação final da operação do sistema executada para obter índices de desempenho do sistema.

Na implementação atual da PDDE, um modelo autorregressivo periódico - PAR (p) (SALAS *et al.*, 1980), (MACEIRA & BEZERRA, 1997), (PENNA *et al.*, 2011) - é

utilizado para gerar cenários de energias/vazões afluentes, que são usados nos passos forward e backward do algoritmo PDDE e na simulação da operação do sistema com a política de operação calculada. Neste modelo de séries temporais, as últimas p energias/vazões afluentes compõem a equação de regressão, juntamente com um ruído aleatório.

Para representar a incerteza das fontes intermitentes, foi proposta em (MACEIRA *et al.*, 2020, 2023) e implementada no modelo NEWAVE uma metodologia para a representação das incertezas associadas às fontes eólicas, com o objetivo de manter o problema estocástico e de grande porte tratável computacionalmente. A abordagem proposta é composta por quatro etapas principais: (i) agrupamento estatístico dos regimes de ventos; (ii) avaliação de funções de transferência mensais (FTMs) entre ventos e produção eólica, a partir de curvas de potência probabilísticas mensais; (iii) geração de séries sintéticas de ventos/produção eólica correlacionadas às afluições aos reservatórios das hidrelétricas; e (iv) representação das produções eólicas mensais no algoritmo PDDE do modelo NEWAVE. O presente trabalho está inserido na etapa (iii).

Para evitar o aumento do número de variáveis de estado do algoritmo PDDE, as velocidades do vento podem ser modeladas como uma regressão das energias/vazões afluentes dos períodos atuais ou anteriores, ou da velocidade do vento no período anterior, ou ainda, da média, mais um resíduo, incluindo também a correlação cruzada entre as afluições às hidrelétricas e as velocidades do vento dos parques eólicos.

Os resíduos empregados podem resultar de uma amostragem aleatória simples ou de uma amostragem seletiva, dependendo da cardinalidade desejada. Por exemplo, os ruídos aleatórios que serão usados nas séries sintéticas de energias/vazões e ventos utilizadas na simulação da política operativa obtida pelo algoritmo de PDDE, comumente 2.000 séries sintéticas, são obtidos por amostragem aleatória simples. Já nas simulações forward e backward do algoritmo, os ruídos aleatórios, comumente 200 e 20 séries sintéticas respectivamente, provêm de uma amostragem seletiva. Nessa metodologia, são amostrados 100.000 ruídos aleatórios por amostragem aleatória simples e através de um algoritmo de agregação, por exemplo, K-Means, reduz-se a cardinalidade da amostra original (100.000) para 200 (forward) e 20 (backward).

Ao contrário dos resíduos de afluições, que apresentam assimetrias positivas, os resíduos das velocidades mensais do vento podem apresentar tanto assimetrias positivas

quanto negativas em vários meses, o que impede, neste último caso, o emprego da distribuição Lognormal 3 parâmetros, atualmente utilizada nas energias/vazões. Uma alternativa é usar a distribuição Weibull (WEILBULL, 1939, 1951), que é bastante versátil e é definida para assimetrias à direita e à esquerda. Como, por construção, os resíduos apresentam valores negativos, é necessário utilizar distribuições Weibull com 3 parâmetros.

Vários métodos estão disponíveis na literatura para se estimar os parâmetros de forma, escala e posição de distribuições Weibull tri-paramétricas, e.g., (COHEN & WHITTEN, 1982), a maioria baseados nos métodos dos momentos (MoM) ou da máxima verossimilhança (MLE), ou em modificações deles.

Contudo, como essas estimativas envolvem estatísticas de ordem superior (no caso, de terceira ordem), tanto o MoM quanto o MLE podem apresentar dificuldades e resultados insatisfatórios ao considerar os três parâmetros da distribuição de Weibull. De fato, observamos que a qualidade das estimativas desses métodos aplicados às velocidades médias mensais dos parques eólicos brasileiros varia muito em função do mês do ano e da localização dos parques eólicos, sendo que os piores desempenhos estão associados aos meses com assimetrias negativas com módulos elevados.

Assim, foi desenvolvido e implementado um algoritmo para modelagem dos resíduos de velocidades mensais do vento através de distribuições Weibull tri-paramétricas, buscando preservar a média, desvio-padrão e assimetria das velocidades históricas mensais, principalmente em situações de elevadas assimetrias (MELO, MACEIRA & PESSANHA, 2020, 2021).

A abordagem baseia-se no fato de, quando o parâmetro de posição é conhecido, as estimativas dos outros dois parâmetros podem ser calculadas de forma mais simples, i.e., usando MoM de segunda ordem, pois não há necessidade de usar a informação de assimetria: o parâmetro de forma pode ser estimado com base no coeficiente de variação, porém substituindo a média e a variância populacionais pelos respectivos valores amostrais; em seguida, o parâmetro de escala é obtido pela expressão do valor esperado. No algoritmo desenvolvido, a determinação da estimativa inicial do parâmetro de posição foi realizada por meio de regressões lineares; e, iterativamente, esta estimativa inicial é atualizada para reduzir a diferença entre a assimetria das velocidades de vento mensais sintéticas e amostrais (históricas).

Esta proposta foi aplicada a vários parques eólicos equivalentes (PEEs), considerando diferentes meses e assimetrias (positivas e negativas), apresentando, em todos os casos, desempenhos melhores que outros métodos disponíveis na literatura (MELO, MACEIRA & PESSANHA, 2020, 2021). Por outro lado, por se tratar de um algoritmo iterativo, a abordagem implica aumento do custo computacional na etapa de geração de cenários sintéticos.

Para contornar este ponto, este trabalho propõe dividir o algoritmo anterior em duas etapas. Na primeira, para um determinado mês e parque eólico (ou PEE), a estimativa inicial do parâmetro de posição (regressões lineares) é substituída pelo MoM de terceira ordem (que também permite o cálculo dos outros dois parâmetros); se a assimetria das velocidades de ventos sintéticas, em relação à assimetria das velocidades de ventos históricas for aceitável, o algoritmo para, não necessitando executar iterações. Caso contrário, a estimativa pelo MoM de terceira ordem é utilizada como estimativa inicial do parâmetro de posição, e o algoritmo entra na segunda etapa, i.e., no procedimento iterativo. Procura-se, assim, assegurar precisão com o menor esforço computacional possível.

São apresentados e discutidos estudos de caso com a abordagem proposta em parques eólicos brasileiros, com distintos regimes de vento, comparando-se o desempenho da abordagem descrita com outros métodos disponíveis na literatura. São também discutidos o impacto da cardinalidade dos resíduos correlacionados no desempenho da metodologia proposta.

2. A Distribuição Weibull

O modelo probabilístico de Weibull foi proposto em 1939 pelo físico sueco Waloddi Weibull em estudos de resistência de materiais e ruptura de sólidos (WEIBULL, 1939). Embora a sua proposição carecesse de uma fundamentação teórica (WEIBULL, 1951) a simplicidade da definição de sua função de distribuição acumulada, aliada à flexibilidade na utilização de três parâmetros – forma, escala e posição - conferiu-lhe grande versatilidade em aplicações pragmáticas em diversas áreas do conhecimento, tais como: física (estudo da cristalização, resistência de materiais, ruptura de sólidos); climatologia (estudo das marés); engenharia/controle de qualidade (análises de confiabilidade e de tempos para a falha; manutenção; velocidades horárias de ventos; resistência da fibras, e.g., algodão); agronomia (caracterização da estrutura diamétrica de

florestas); biologia (comprimento de cyrtoideae), finanças (análise de sobrevivência com longa duração de carteira de clientes); psicologia cognitiva (estudo do tempo para completar uma tarefa); etc.

A função de distribuição acumulada de uma variável aleatória x com distribuição Weibull com três parâmetros, i.e., $W(\gamma, \beta, \alpha)$, é dada por (1):

$$F(x; \gamma, \beta, \alpha) = 1 - \exp \left[- \left(\frac{x - \alpha}{\beta} \right)^\gamma \right] \quad (1)$$

onde $\gamma > 0$, $\beta > 0$ e $\alpha < x$ são os parâmetros de forma, escala e posição, respectivamente.

A correspondente função densidade de probabilidade é definida como (2):

$$f(x; \gamma, \beta, \alpha) = \left(\frac{\gamma}{\beta} \right) \cdot \left(\frac{x - \alpha}{\beta} \right)^{\gamma - 1} \exp \left[- \left(\frac{x - \alpha}{\beta} \right)^\gamma \right] \quad (2)$$

Quando o parâmetro de posição é nulo, a distribuição é denominada Weibull com 2 parâmetros, a qual somente está definida para valores positivos. Assim, quando a variável aleatória modelada apresenta valores nulos ou negativos, deve-se acrescentar o parâmetro posição. Dependendo dos valores de seus parâmetros, a distribuição Weibull pode assumir várias formas, conforme ilustrado em (MELO, MACEIRA & PESSANHA, 2020).

3. Principais Métodos para Ajustes de Distribuições Weibull Triparamétricas

Dentre os principais métodos para estimação dos parâmetros do modelo Weibull triparamétrico destacam-se o estimador de máxima verossimilhança - MLE e o método dos momentos - MoM (COHEN & WHITTEN, 1982). O MoM está resumido a seguir, enquanto uma descrição do MLE pode ser encontrada em (MELO *et al.*, 2020, 2021).

3.1 Estimativa pelo Método dos Momentos de 3ª Ordem (MoM-3)

As três estatísticas básicas (momentos centrais) da distribuição Weibull, i.e., a média, a variância e a assimetria populacionais são respectivamente expressas em (3), (4) e (5) por (COHEN & WHITTEN, 1982):

$$E(x) = \alpha + \beta \Gamma_1 \quad (3)$$

$$Var(x) = \beta^2 (\Gamma_2 - \Gamma_1^2) \quad (4)$$

$$Sk(x) = \frac{\Gamma_3 - 3\Gamma_2\Gamma_1 + 2\Gamma_1^3}{(\Gamma_2 - \Gamma_1^2)^{3/2}} \quad (5)$$

onde:

$$\Gamma_k(\gamma) = \Gamma\left(1 + k/\gamma\right) \quad (6)$$

e $\Gamma(z)$ é a função gama, definida como:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \quad (7)$$

No método dos momentos, as estimativas dos parâmetros são obtidas igualando-se os momentos populacionais aos correspondentes momentos amostrais. Alguns estimadores MoM foram propostos para as distribuições Weibull com dois parâmetros, e.g., o proposto em (HARTER & MOORE, 1965), contudo eles não podem ser generalizados para uma distribuição Weibull com três parâmetros, pois é necessário a inclusão de mais um momento no sistemas de equações (CRAN, 1988).

Conforme proposto em (COHEN & WHITEN, 1982), os três parâmetros podem ser obtidos de maneira sequencial. A partir de (5) é possível derivar uma estimativa MoM para a forma, $\hat{\gamma}_{MoM}$, que não requer informações dos outros dois parâmetros, substituindo $Sk(x)$ pela assimetria amostral \hat{g} e resolvendo a seguinte equação transcendental (8):

$$\frac{\Gamma_3 - 3\Gamma_2\Gamma_1 + 2\Gamma_1^3}{(\Gamma_2 - \Gamma_1^2)^{3/2}} - \hat{g} = 0 \quad (8)$$

De posse de $\hat{\gamma}_{MoM}$, a estimativa $\hat{\beta}_{MoM}$ pode ser obtida de (4) utilizando a variância amostra s^2 (9):

$$\hat{\beta}_{MoM} = s / \sqrt{\Gamma_2(\hat{\gamma}_{MoM}) - \Gamma_1^2(\hat{\gamma}_{MoM})} \quad (9)$$

Finalmente, $\hat{\alpha}_{MoM}$ pode ser computado por (3), empregando a média amostral \bar{x} :

$$\hat{\alpha}_{MoM} = \bar{x} - \hat{\beta}_{MoM} \Gamma_2(\hat{\gamma}_{MoM}) \quad (10)$$

3.2 Considerações sobre os Métodos MLE e MoM

Embora tradicionais, os estimadores de momento (MoM) e de máxima verossimilhança (MLE), em determinadas situações, podem apresentar dificuldades teóricas e resultados insatisfatórios quando se consideram os três parâmetros da Weibull.

Para que um método de estimação de parâmetros de uma distribuição seja eficiente à medida que o tamanho da amostra aumenta para o infinito, a distribuição deve satisfazer condições de regularidade usuais, tais como eficiência, consistência e normalidade assintótica. No entanto, para algumas combinações de parâmetros, as condições de regularidade não são satisfeitas para a estimativa MLE da distribuição Weibull triparamétrica, porque o domínio da variável aleatória depende da posição do limite inferior - o parâmetro α - que é desconhecido. Consequentemente, os estimadores MLE podem não existir e também dar origem a estimativas sem as propriedades assintóticas usuais e até mesmo a estimativas inconsistentes (COHEN & WHITTEN, 1982).

Já os estimadores MoM para distribuições Weibull são bem conhecidos e tendem a superar as desvantagens da abordagem MLE, porém sua existência, unicidade e consistência são questões em aberto. Adicionalmente, as estatísticas de momentos superiores são mais variáveis e têm maior probabilidade de resultar em uma estimativa menos precisa. Assim, modificações do método MoM têm sido desenvolvidas para as distribuições Weibull com três parâmetros, principalmente procurando substituir momento de terceira ordem por outra relação menos variável.

4. Abordagem Proposta

Conforme mencionado anteriormente, o objetivo final deste trabalho é a modelagem de ventos para a geração de cenários sintéticos de velocidades mensais em um esquema de PDDE. Neste trabalho, a velocidade dos ventos será modelada pela média mensal mais um resíduo multiplicado pelo desvio-padrão mensal. Uma distribuição Weibull tri-paramétrica será ajustada aos resíduos das médias mensais de velocidades de ventos correlacionados entre si e com as afluições (MACEIRA *et al.*, 2020, 2023).

Embora o ajuste das distribuições Weibull tri-paramétricas seja aplicado aos resíduos das velocidades, a geração de cenários sintéticos na PDDE é de ventos. Assim, a abordagem proposta busca ainda preservar a média, o desvio padrão e, especialmente, a assimetria das velocidades históricas mensais dos ventos, sendo particularmente adequada em situações de assimetrias elevadas.

A abordagem proposta constitui-se de duas etapas.

Na *primeira etapa*, os parâmetros da distribuição Weibull dos resíduos das velocidades mensais históricas do vento são estimados pelo Método dos Momentos de Terceira Ordem (MoM-3), conforme Seção 3.1. Se a assimetria das velocidades sintéticas

do vento for suficientemente próxima da assimetria das velocidades históricas do vento, o algoritmo para; não necessitando entrar no procedimento iterativo da segunda etapa e, portanto, apresentando menor esforço computacional.

Caso contrário, entra-se na *segunda etapa* da metodologia, constituída de um algoritmo iterativo. Nele, a estimativa do parâmetro *posição*, obtida pelo MoM-3 (primeira etapa) é utilizada como valor inicial ($\hat{\alpha}_{inic}$) para calcular estimativas dos outros dois parâmetros (*forma* e *escala*) por meio do Método dos Momentos de Segunda Ordem (MoM-2) – vide Seção 4.1. Em seguida, e de forma iterativa, o algoritmo atualiza a estimativa inicial do parâmetro *posição* com vistas a reduzir a diferença entre as assimetrias das velocidades sintéticas e históricas do vento, até que se atinja um valor considerado aceitável. Ressalta-se que, como na estimativa dos parâmetros pelo MoM-3 há a necessidade de se resolver a equação transcendental dada por (8), na hipótese de não se encontrar uma solução real, o valor inicial do parâmetro *posição* passa a ser obtido por meio de regressões lineares, conforme proposto em (MELO, MACEIRA & PESSANHA, 2020, 2021).

Procura-se, assim, assegurar precisão com o menor esforço computacional possível.

4.1 Estimativa quando o parâmetro de posição α é conhecido - MoM de 2a Ordem (MoM-2)

Quando a posição α é conhecida, as estimativas dos outros dois parâmetros podem ser calculadas pelo MoM, porém de maneira mais simples (HARTER & MOORE, 1965), visto que não há a necessidade de utilizar a equação da assimetria (5). Neste método, aqui denominado MoM-2O, consiste em estimar o parâmetro de forma $\hat{\gamma}$ com base no coeficiente de variação obtido a partir de (3) e (4), porém substituindo a média e a variância populacionais pelos respectivos valores amostrais. Para isso, resolve-se a seguinte equação transcendental (11):

$$\frac{\Gamma_2(\hat{\gamma})}{\Gamma_1^2(\hat{\gamma})} - 1 - \frac{s^2}{\bar{x} - \alpha} = 0 \quad (11)$$

Uma vez obtido $\hat{\gamma}$, a estimativa $\hat{\beta}$ pode ser obtida de (3):

$$\hat{\beta} = \frac{(\bar{x} - \alpha)}{\Gamma_1(\hat{\gamma})} \quad (12)$$

4.2 Algoritmo Proposto

De maneira compacta, o algoritmo proposto, aqui denominado *MoM-3/Iter*, é constituído dos seguintes passos:

- i. Leitura dos dados amostrais, da distribuição de resíduos correlacionados e normalmente distribuídos, e da equação de regressão para, se necessário, determinar $\hat{\alpha}_{inic}$;
- ii. Definição de parâmetros de controle e inicialização de variáveis;
- iii. Compute as estimativas $\hat{\gamma}_{MoM-3}$, $\hat{\beta}_{MoM-3}$ e $\hat{\alpha}_{MoM-3}$ pelo método MoM-3, conforme Seção 3.1. Se for encontrada solução real para a equação transcendental (8.1), faça $\hat{\alpha}_{inic} = \hat{\alpha}_{MoM-3}$ e vá para o *passo (vi)*; caso contrário, calcule $\hat{\alpha}_{inic}$ pela equação de regressão e vá para o *passo (v)*;
- iv. Atualize o valor de $\hat{\alpha}$ por meio de interpolações lineares, considerando os valores prévios deste parâmetro e das assimetrias;
- v. Compute as estimativas $\hat{\gamma}_{MoM-2}$ e $\hat{\beta}_{MoM-2}$ pelo método MoM-2, conforme Seção 4.1;
- vi. Gere os resíduos sintéticos Weibull e as velocidades sintéticas de ventos associadas, usando os resíduos lidos, i.e., resíduos correlacionados e normalmente distribuídos;
- vii. Calcule a assimetria das velocidades sintéticas do vento;
- viii. Calcule o erro relativo entre as assimetrias histórica e sintética das velocidades de ventos;
- ix. Se o erro relativo > tolerância, vá para o *passo (iv)*;
- x. Fim do algoritmo.

5. Estudos de Caso

Para os testes da metodologia proposta em (MACEIRA *et al.* 2020, 2023) seria relevante que se tivesse acesso às séries horárias consistidas e pareadas de velocidades de ventos e produções eólicas, e.g., para os parques vencedores nos leilões públicos de compra de energia elétrica.

Como no Brasil não existe a disponibilização pública desses dados, com as resoluções espaciais e temporais necessárias, tornou-se forçosa a realização de avaliações

adicionais, incluindo dados oriundos de reanálises, por exemplo, os dados oriundos do MERRA-2 global (*National Aeronautics and Space Administration* – NASA) e ERA5 (*European Centre for Medium-Range Weather Forecasts* – ECMWF). As séries temporais de velocidades horárias do vento foram então agrupadas por mês para a obtenção de médias mensais das velocidades do vento.

A metodologia proposta foi aplicada a parques eólicos das regiões Nordeste e Sul do Brasil, considerando meses distintos e diversas assimetrias – positivas e negativas.

Para cada parque estudado, foi considerada uma amostra de 37 anos de velocidades médias mensais do vento (obtidas por reanálise), e foram gerados resíduos correlacionados com cardinalidade 2.000, 200 (amostragem aleatória simples), 50 e 20 (amostragem seletiva).

O desempenho da metodologia proposta foi comparado com outros 17 métodos disponíveis na literatura, os quais foram implementados pelos autores ou se encontram disponíveis no software R (R CORE TEAM 2019).

Devido a limitação de espaço, apenas os resultados relativos a dois parques eólicos equivalentes (PEEs) são aqui apresentados, para meses específicos.

5.1 Comparação com Métodos Existentes – 2000 Cenários

O primeiro caso estudado (caso NE_Int_Nov) refere-se a um parque eólico situado no interior do Nordeste. O histograma dos 37 valores históricos das médias mensais das velocidades de vento para o mês de novembro está apresentado na Figura 1(a). As estatísticas média, desvio-padrão e assimetria amostrais são, respectivamente, 4,872; 0,752 e -0,412.

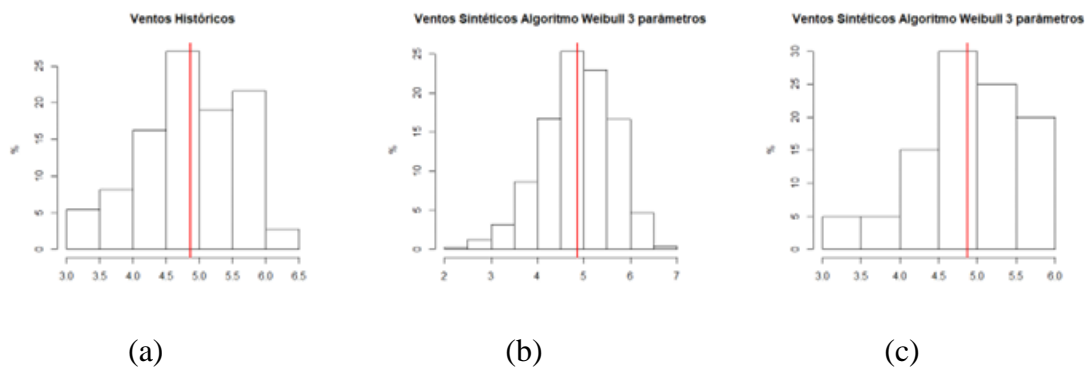
Os resultados da aplicação do método proposto (*MoM-3/Iter*), considerando as duas etapas do algoritmo, e dos demais 17 utilizados na estimação dos três parâmetros da distribuição Weibull estão exibidos na Tabela 1, assim como as estatísticas dos valores sintéticos das velocidades mensais do vento, considerando 2.000 resíduos correlacionados e normalmente distribuídos. Nesta tabela, os métodos estão ordenados do melhor para o pior desempenho, adotando o erro relativo entre as assimetrias histórica e sintética das velocidades do vento como critério de parada. Como critérios adicionais são adotados os erros relativos dos desvios-padrão e das médias, nesta ordem.

Verifica-se na Tabela 1 que a assimetria obtida com os 2.000 valores sintéticos das velocidades mensais de ventos produzidos pelo algoritmo proposto ficou igual à

histórica (37 valores), classificando-o como o método de melhor desempenho (erro relativo = 0%). O método M13 foi o segundo melhor método (erro relativo assimetria = 12%) Nota-se também que 3 métodos apresentaram erro relativo inferior a 10%; 15 acima de 20%; 15 acima de 50%, sendo que 7 deles apresentaram erros acima de 150% e, inclusive, com assimetrias positivas. Há ainda métodos que produziram desempenhos não satisfatórios ou inaceitáveis, e que apresentaram desempenhos razoáveis em outros casos analisados.

A boa aderência dos valores sintéticos das velocidades de ventos produzidos pelo método *MoM-3/Iter* aos valores históricos pode ser constatada por meio da Figura 1 (b).

Figura 1 – Histogramas de velocidades de vento para o NE_Int_Nov ($\hat{g} = -0,412$): (a) 37 valores históricos; 2.000 valores sintéticos (algoritmo MoM-3/Iter proposto); e 20 valores sintéticos (algoritmo MoM-3/Iter proposto).



Fonte: Autores

Tabela 1 – NE_Int_Nov: parâmetros estimados da distribuição Weibull; estatísticas dos 2.000 valores de velocidades sintéticas mensais de ventos; erro relativo entre assimetrias histórica e sintética.

Parâmetros Weibull			Estatísticas - Valores Sintéticos				Método
Forma γ	Escala β	Posição α	Média	D.P.	Assimetria	Erro Rel (%)	
6.584	6.031	-5.623	4.872	0.753	-0.412	0	M1 - Algoritmo MoM-3/Iterativo Proposto
6.032	5.335	-9.926	4.872	0.719	-0.362	12	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
7.298	6.690	-6.272	4.872	0.764	-0.467	13	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
7.847	7.805	-7.344	4.872	0.836	-0.503	22	M18 - Momentos T-L - Pacote R ForestFit
8.047	7.100	-6.679	4.872	0.743	-0.515	25	M11 - MLE (Green, 1994) - Pacote R ForestFit
8.048	7.122	-6.709	4.872	0.746	-0.515	25	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
5.179	4.718	-4.330	4.872	0.725	-0.267	35	M6 - MLE Ajustado - Pacote R Weibullness
9.284	2.108	0.713	4.872	0.194	-0.580	41	M17 - MLE Ponderado - (Cousineau, 2009) - Pacote R ForestFit
9.813	9.312	-8.863	4.872	0.817	-0.604	47	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit
4.597	3.695	-4.044	4.872	0.628	-0.185	55	M2 - MoM (Cohen, 1982)
4.359	4.289	-3.907	4.872	0.763	-0.147	64	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
2.590	2.588	-2.288	4.872	0.715	0.313	176	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
2.487	2.622	-2.326	4.872	0.750	0.356	186	M5 - Recálculo do MoM com α dado por E[Min de uma amostra]
1.238	1.318	-2.326	4.872	0.743	1.391	438	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)
1.238	2.492	-2.326	4.872	1.404	1.391	438	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)
1.232	0.710	-0.664	4.872	0.402	1.400	440	M7 - MoM (Cran, 1988) - Pacote R ForestFit
0.683	3.806	-2.200	4.872	5.384	3.460	940	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit
0.250	0.005	-0.122	4.872	0.567	17.718	4401	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit

Fonte: Autores

5.2 Comparação com Métodos Existentes – 20 Cenários

Em seguida foi realizada uma análise de sensibilidade para o caso NE_Int_Nov, com amostras de cardinalidades iguais a 200, 50 e 20, dos resíduos correlacionados e normalmente distribuídos, obtidos empregando-se técnicas de análise de agrupamentos. Devido a limitação de espaço, apenas os resultados relativos à cardinalidade 20 é apresentado.

Assim, a Tabela 2 exibe o resultado considerando agora 20 resíduos correlacionados e normalmente distribuídos. Observa-se, novamente, que o algoritmo *MoM-3/Iter* proposto obteve o mesmo valor da assimetria histórica, alcançando o melhor desempenho em relação a todos os métodos estudados. Já o segundo e terceiro melhores métodos foram o M5 (erro relativo = 13%) e M12 (erro relativo = 25%); observe que ambos ocupavam respectivamente a 13^a e 12^a posição no caso anterior, e produziram desempenhos inaceitáveis (assimetrias positivas e com erros no entorno de 180%). Também, dos 18 métodos, 17 apresentam erros relativos superiores a 10%, e apenas 2 abaixo de 20%. Há ainda 15 métodos que produziram desempenhos não satisfatórios ou inaceitáveis.

A boa aderência dos valores sintéticos das velocidades de ventos produzidos pelo método *MoM-3/Iter* aos valores históricos, mesmo com cardinalidade reduzida (20) pode também ser constatada por meio da Figura 1 (c).

Tabela 2 – NE_Int_Nov: parâmetros estimados da distribuição Weibull; estatísticas dos 20 valores de velocidades sintéticas mensais de ventos; erro relativo entre assimetrias histórica e sintética.

Parâmetros Weibull			Estatísticas - Valores Sintéticos				Método
Forma γ	Escala β	Posição α	Média	D.P.	Assimetria	Erro Rel (%)	
2.380	2.523	-2.236	4.871	0.680	-0.412	0	M1 - Algoritmo MoM-3/iterativo Proposto
2.487	2.622	-2.326	4.872	0.685	-0.466	13	M5 - Recálculo do MoM com α dado por E[Min de uma amostra]
2.590	2.588	-2.288	4.880	0.658	-0.515	25	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
4.359	4.289	-3.907	4.875	0.757	-1.096	166	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
4.597	3.695	-4.044	4.372	0.628	-1.147	178	M2 - MoM (Cohen, 1982)
1.238	1.318	-2.326	4.017	0.581	0.381	192	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)
1.238	2.492	-2.326	4.814	1.098	0.381	192	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)
1.232	0.710	-0.664	4.855	0.314	0.386	194	M7 - MoM (Cran, 1988) - Pacote R ForestFit
5.179	4.718	-4.330	4.883	0.735	-1.256	205	M6 - MLE Ajustado - Pacote R Weibullness
6.032	5.335	-9.926	1.133	0.741	-1.382	235	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
7.298	6.690	-6.272	4.872	0.801	-1.520	269	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
7.847	7.805	-7.344	4.871	0.883	-1.567	280	M18 - Momentos T-L - Pacote R ForestFit
8.047	7.100	-6.679	4.879	0.787	-1.583	284	M11 - MLE (Green, 1994) - Pacote R ForestFit
8.048	7.122	-6.709	4.871	0.789	-1.583	284	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
9.284	2.108	0.713	6.910	0.208	-1.666	304	M17 - MLE Ponderado - (Cousineau, 2009) - Pacote R ForestFit
9.813	9.312	-8.863	4.861	0.878	-1.695	311	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit
0.683	3.806	-2.200	6.296	3.072	1.038	352	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit
0.250	0.005	-0.122	4.797	0.032	2.218	638	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit

Fonte: Autores

Tabela 3 – Estimativas das assimetrias sintéticas – MoM-3, MoM-3/Iter e RL/Iter) – Caso NE_Int_Nov.

No. de Cenários	Assimetria Histórica = -0.412		
	Vel. Ventos Sintéticos		
	MoM-3	MoM-3/Iter	RL/Iter
2000	-0,396	-0,412	-0,412
200	-0,464	-0,412	-0,412
50	-0,753	-0,412	-0,412
20	-1,427	-0,412	-0,412

Fonte: Autores

Tabela 4 – Estimativas das assimetrias sintéticas – MoM-3, MoM-3/Iter e RL/Iter) – Caso Sul_Int_Nov.

No. de Cenários	Assimetria Histórica = -0.366		
	Vel. Ventos Sintéticos		
	MoM-3	MoM-3/Iter	RL/Iter
2000	-0,379	-0,366	-0,366
200	-0,143	-0,366	-0,366
50	-0,320	-0,366	-0,366
20	-1,175	-0,366	-0,366

Fonte: Autores

5.3 Comparação entre os Desempenho das Etapas 1 (MoM-3) e 2 (Iterativa) do Método Proposto

Com o intuito de ilustrar a precisão nas duas etapas do algoritmo proposto (MoM-3 e MoM-3/Iter), as Tabelas 3 e 4 apresentam, respectivamente, as estimativas das assimetrias para dois parques (NE_Int_Nov e Sul_Int_Nov). As assimetrias das velocidades históricas foram calculadas com uma amostra de 37 valores, enquanto as assimetrias sintéticas foram computadas com número de cenários iguais a 2.000, 200, 50 e 20.

Observa-se na Tabela 3 que, para 2.000 cenários, o MoM-3 apresenta erro relativo para a assimetria igual a 3,9%, o que é bastante aceitável. Contudo, quando se reduz o número de cenários para 200, 50 e 20, este erro aumenta para 12,6%, 82,9% e 246,4%. Por outro lado, quando o procedimento iterativo é acionado, tendo como estimativa inicial $\hat{\alpha}_{inic} = \hat{\alpha}_{MoM-3}$, o método MoM-3/Iter consegue obter erro nulo, i.e., as assimetrias sintéticas se igualam às amostrais, para todos os números de cenários considerados. A Tabela 3 ainda apresenta as estimativas das assimetrias sintéticas, usando o procedimento iterativo, porém substituindo o $\hat{\alpha}_{inic}$ pelo valor obtido por regressões lineares método RL/Iter, conforme proposto em (MELO, MACEIRA E PESSANHA, 2020). Nota-se que

o RL/Iter apresenta, em termos de assimetria, o mesmo resultado que o MoM-3/Iter, com a diferença que nos casos de 20 e 200 cenários fez uma iteração a mais, passando, respectivamente, de 8 e 4 para 9 e 5 iterações.

Comportamento semelhante também ocorreu com o caso Sul_Int_Nov, conforme Tabela 4. Já no caso de 20 cenários, ambos os métodos MoM-3/Iter e RL/Iter realizaram 11 iterações, ao passo que com 200 cenários, o RL/Iter realizou 1 iteração a menos (5) que o MoM-3/Iter (6).

Por sua vez, as Tabelas 5 e 6 apresentam as estimativas da média e do desvio-padrão em função do número de cenários, para ambos os parques eólicos. Verifica-se que há uma pequena degradação nas estimativas dos desvios-padrão à medida que se diminui o número de cenários, sendo que a do MoM-3/Iter é um pouco superior ao do MoM-3. Por exemplo, considerando 20 cenários, os coeficientes de variação para o caso NE_Int_Nov são 16% (MoM-3) e 14% (MoM-3/Iter); já para o Sul_Int_Nov, os valores são 8% (MoM-3) e 6% (MoM-3/Iter).

Tabela 5 – Estimativas das médias e desvios-padrão sintéticos – MoM-3 e MoM-3/Iter) – NE_Int_Nov.

	Média = 4,872		D. Pad. = 0,752 (Históricos)	
	Vel. Ventos Sintéticos			
	MoM-3		MoM-3/Iter	
No. de Cenários	Média	Desv. Pad.	Média	Desv. Pad.
2000	4,852	0,753	4,859	0,753
200	4,895	0,777	4,896	0,777
50	4,833	0,755	4,832	0,730
20	4,873	0,781	4,871	0,680

Fonte: Autores

Tabela 6 – Estimativas das médias e desvios-padrão sintéticos – MoM-3 e MoM-3/Iter) – Sul_Int_Nov.

	Média = 6,203		D. Pad. = 0,444 (Históricos)	
	Vel. Ventos Sintéticos			
	MoM-3		MoM-3/Iter	
No. de Cenários	Média	Desv. Pad.	Média	Desv. Pad.
2000	6,213	0,442	6,213	0,442
200	6,207	0,473	6,205	0,469
50	6,067	0,429	6,067	0,430
20	6,210	0,470	6,203	0,346

Fonte: Autores

6. Conclusões

Este trabalho apresentou uma proposta de modelagem das velocidades mensais de ventos por meio de distribuições Weibull triparamétricas, na geração de cenários de ventos mensais correlacionados com afluências às usinas hidrelétricas, em um esquema de PDDE, sem aumentar o número associado de variáveis de estado.

A metodologia proposta foi aplicada a parques eólicos das regiões Nordeste e Sul do Brasil, considerando meses distintos e diversas assimetrias – positivas e negativas, e o seu desempenho foi comparado a 17 métodos disponíveis na literatura. Por limitação de espaço, apenas parte dos resultados foram aqui apresentados.

Em todos os casos estudados, a abordagem proposta (MoM-3/Iter) mostrou-se robusta e apresentou desempenho superior aos outros 17 métodos. Verificou-se que quando a cardinalidade da amostra de resíduos correlacionados diminuiu, o desempenho dos outros métodos também diminuiu, o que não ocorreu com a abordagem proposta. Observou-se ainda que, com exceção do método proposto, os métodos que exibiram os melhores desempenhos variaram de caso para caso e, também, que métodos que apresentam bom desempenho em um caso podem ter desempenho insatisfatório, ou mesmo inaceitável, em outros casos.

Com relação às etapas 1 (MoM-3) e 2 (MoM-3/Iter) do método proposto, verificou-se que para amostras dos resíduos correlacionados com cardinalidade elevada (e.g., 2.000 cenários), o método MoM-3 produziu boas estimativas para as assimetrias das velocidades sintéticas em relação às históricas. Contudo, quando a cardinalidade diminuiu, a precisão das estimativas se reduz, podendo deixar de ficar aceitável (e.g., 20 cenários). Por outro lado, o algoritmo MoM-3/Iter conseguiu produzir estimativas sintéticas idênticas às do histórico, para todos os números de cenários utilizados.

Por fim, a abordagem proposta é geral, e pode ser utilizada em outras aplicações.

Agradecimentos

Este trabalho foi parcialmente financiado pelo CNPq/MCTI/FNDCT por meio do Projeto de Pesquisa 409715/2021-2.

Referências

COHEN, A.C., WHITTEN, B.J. Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution. **Communications in Statistics—Theory and Methods** 11, 2631–2656, 1982.

CRAN, G.W. Moment estimators for the 3-parameter Weibull distribution. **IEEE Transactions on Reliability** 37, 360–363, 1988.

HARTER, H.L., MOORE, H., “Maximum likelihood estimation of the parameters of Gamma and Weibull populations from complete and from censored samples”. **Technometrics**, Vol. 7, pp. 639-643, 1965.

HELSETH, A., MELO, A.C.G. Scheduling Toolchains in Hydro-Dominated Systems - Evolution, Current Status and Future Challenges for Norway and Brazil, **SINTEF Energy Research Technical Report**, 2020-08-10, May 2020, disponível em <https://sintef.brage.unit.no/sintef-xmlui/handle/11250/2672581?locale-attribute=en>.

HELSETH, A., MELO, A.C.G., PLOUSSARD, Q.M., MO, B., MACEIRA, M.E.P., BOTTERUD, VOISIN, A., N. Hydropower Scheduling Toolchains: Comparing Experiences in Brazil, Norway, and USA and Implications for Synergistic Research, **Journal of Water Resources Planning and Management**, Vol. 149, Issue 7 (July 2023). Published online: May 13, 2023. Vol 149, No 7, 2023, DOI: <http://dx.doi.org/10.1061/JWRMD5.WRENG-5911>, ISSN: 0733-9496 resposta do Lattes, pgs 1-17, 2023.

MACEIRA, M. E. P. Programação Dinâmica Dual Estocástica Aplicada ao Planejamento da Operação Energética de Sistemas Hidrotérmicos com Representação do Processo Estocástico de Afluências por Modelos Auto-Regressivos Periódicos, **Relatório Técnico CEPEL** 237/93, Junho 1993.

MACEIRA, M.E.P.; BEZERRA, C.V. Stochastic Streamflow Model for Hydroelectric Systems, **5th Probabilistic Methods Applied to Power Systems** - PMAPS, Vancouver, Canada, September 1997.

MACEIRA, M.E.P., TERRY, L.A., COSTA, F.S., DAMÁZIO, J.M., MELO, A.C.G. Chain of optimization models for setting the energy dispatch and spot price in the Brazilian system, **14th Power Systems Computation Conference** – PSCC 2002, Seville, Spain, 2002.

MACEIRA, M.E.P., PENNA, D.D.J., DINIZ, A.L., PINTO, R.J., MELO, A.C.G., VASCONCELLOS, C.V., CRUZ, C.B. Twenty Years of Application of Stochastic Dual Dynamic Programming in Official and Agent Studies in Brazil – Main Features and Improvements on the NEWAVE Model, **20th Power System Computation Conference** – PSCC 2018, Dublin, Ireland, 2018.

MACEIRA, M. E. P., CRUZ, C. B., PENNA, D. D. J., DINIZ, A. L., TREISTMAN, F., MELO, A. C. G., JUSTINO, T. C. Representação de Usinas Hidrelétricas Individualmente e de Forma Agregada na Programação Dinâmica Dual Estocástica - A Abordagem Híbrida. **XXV Seminário Nacional de Produção e Transmissão de Energia Elétrica**, Belo Horizonte, 2019.

MACEIRA, M.E.P., MELO, A.C.G., PESSANHA, J.F.M., CRUZ, C.B., ALMEIDA, V.A., JUSTINO, T. C. Uma Abordagem para a Representação das Incertezas da Fonte de Geração Eólica no Modelo NEWAVE, **Cadernos do IME – Série Estatística**, v. 48, p.1 – 36, Junho 2020. ISSN on-line 2317-4535 / ISSN impresso 1413-9022. DOI: 10.12957/cadest.2020.55395.

MACEIRA, M.E.P., MELO, A.C.G., PESSANHA, J.F.M., CRUZ, C.B., ALMEIDA, V.A., JUSTINO, T. C. Combining monthly wind and inflow uncertainties in the stochastic dual dynamic programming. **Energy Syst**, 2023. <https://doi.org/10.1007/s12667-023-00580-5>.

MELO, A.C.G., MACEIRA, M.E.P., PESSANHA, J. Ajuste de Distribuições Weibull Tri-Paramétricas com Elevadas Assimetrias na Geração de Velocidades Mensais de Ventos Correlacionadas com Vazões na Programação Dinâmica Dual Estocástica, **Cadernos do IME – Série Estatística**, v. 49, Dezembro 2020. ISSN on-line 2317-4535 / ISSN impresso 1413-9022.

MELO, A.C.G., MACEIRA, M.E.P., PESSANHA, J.F.M. Fitting three-parameter Weibull distributions with high negative skewness to generate monthly wind scenarios correlated with inflows to the Brazilian hydropower reservoirs, **41st International Symposium on Forecasting** – ISF, 2021. On-line.

PEREIRA, M. V. F., PINTO, L. M. V. G. Multi Stage Stochastic Optimization Applied to Energy Planning,

Mathematical Programming, v. 52, pp. 359-375, 1991.

PENNA, D.D.J., MACEIRA, M.E.P., DAMÁZIO, J.M. Selective Sampling Applied to Long-Term Hydrothermal Generation Planning, **17th PSCC**, Stockholm, Sweden, 2011.

PESSANHA, J.F.M, ALMEIDA, V.A., MELO, A.C.G. Fator de capacidade da geração eólica na região Nordeste: um estudo com dados de reanálises oriundos do MERRA-2 (global). **VIII SBSE**, Santo André-SP, 2020.

PESSANHA, J.F.M, ALMEIDA, V.A., MELO, A.C.G. Fator de capacidade da geração eólica na região Nordeste: um estudo com dados de reanálises oriundos do MERRA-2 (global). **VIII SBSE**, Santo André, 2020.

R CORE TEAM. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. 2019. Disponível em URL <https://www.R-project.org/>.

SALAS, J.D.; DELLEUR, J.W.; YEVJEVICH, V.; LANE, W.L Applied Modeling of Hydrologic Time Series, **Water Resources Publications**, 1980.

WEIBULL, W. The Phenomenon of Rupture in Solids. In: **Ingeniors Vetenskaps Akademiens Handlingar**, vol. 153, 1939.

WEIBULL, W. A Statistical Distribution Function of Wide Applicability, **Journal of Applied Mechanics**, September, 1951.

FITTING THREE-PARAMETER WEIBULL DISTRIBUTIONS WITH HIGH SKEWNESS IN THE GENERATION OF MONTHLY WIND SPEED SCENARIOS CORRELATED WITH INFLOWS IN STOCHASTIC DUAL DYNAMIC PROGRAMMING

Abstract

This work proposes an approach for modeling monthly wind speeds using tri-parametric Weibull distributions, in the generation of monthly wind scenarios correlated with inflows to hydropower plants, in a Stochastic Dual Dynamic Programming (SDDP) scheme, and without increase the number of SDDP state variables. The approach consists of two stages. In the first, the parameters of the Weibull distribution of the residuals of historical monthly wind speeds are estimated by the Third Order Method of Moments. If the skewness of the synthetic speeds is sufficiently close to the historical one, the algorithm stops. Otherwise, the estimate of the position parameter, obtained in the first step, is used as the initial value to calculate estimates of the other two parameters (shape and scale) using the Second Order Method of Moments. Then, iteratively, the algorithm updates the initial estimate of the position parameter targeting to reduce the difference between the skewnesses of synthetic and historical wind speeds, until a value considered acceptable is reached. The aim is, therefore, to ensure accuracy with the least possible computational effort. The application to wind farms in the Northeast and South of Brazil, considering different months and distinct skewnesses, points to its robustness, always performing better than 17 other methods evaluated. Finally, the proposed approach is general, and can be used in other applications.

Key-words: Wind power; Generation of synthetic series; Weibull distribution; Method of moments; Renewable energy; Planning and operation; Stochastic optimization.