

## CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ  
ISSN on-line 2317-4535 / ISSN impresso 1413-9022 - v. 49, p.35 - 65, 2020  
DOI: 10.12957/cadest.2020.61793

# AJUSTE DE DISTRIBUIÇÕES WEIBULL TRI-PARAMÉTRICAS COM ELEVADAS ASSIMETRIAS NA GERAÇÃO DE VELOCIDADES MENSAS DE VENTOS CORRELACIONADAS COM VAZÕES NA PROGRAMAÇÃO DINÂMICA DUAL ESTOCÁSTICA

Albert C.G. Melo

Universidade do Estado do Rio de Janeiro e Centro de Pesquisas de Energia Elétrica  
[albert.melo@ime.uerj.br](mailto: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](mailto:melvira@ime.uerj.br)

José F. M. Pessanha

Universidade do Estado do Rio de Janeiro e Centro de Pesquisas de Energia Elétrica  
[pessanha@ime.uerj.br](mailto:pessanha@ime.uerj.br)

## Resumo

*Propõe-se uma abordagem para a modelagem de velocidades mensais de ventos por meio de distribuições Weibull tri-paramétricas, na geração de cenários de ventos mensais correlacionados com vazões às usinas hidroelétricas, em um esquema de Programação Dinâmica Dual Estocástica. A velocidade dos ventos é modelada como uma regressão das afluências mais um resíduo. Assim, a distribuição Weibull é ajustada aos resíduos das velocidades mensais de ventos, correlacionados e normalmente distribuídos, que também considera as correlações cruzadas entre as vazões às hidroelétricas e as velocidades do vento dos parques eólicos. O algoritmo proposto parte de um valor inicial para a estimativa do parâmetro posição, o qual pode ser obtido através de regressões lineares, calcula estimativas dos outros parâmetros por meio do método dos momentos e, de forma iterativa, atualiza a estimativa inicial com vistas a reduzir a diferença entre as assimetrias das velocidades de ventos sintéticas (geradas) e históricas. A metodologia proposta foi aplicada a parques eólicos do Nordeste e Sul do Brasil em diversas situações, incluindo assimetrias elevadas, e o seu desempenho foi superior a 17 outros métodos. A abordagem proposta é geral, e pode ser utilizada em outras aplicações.*

**Palavras-chave:** Distribuição de Weibull; Estimação de Parâmetros; Métodos dos momentos; Máxima verossimilhança; Geração eólica; Otimização estocástica.

## 1. Introdução

Em fevereiro de 2021 a capacidade instalada em geração eólica no Sistema Interligado Nacional (SIN) alcançou 18.000 MW, distribuídos em 695 usinas. Em termos relativos, a geração eólica corresponde hoje a cerca de 10% da capacidade instalada no SIN, e há a perspectiva de maior participação desta fonte na matriz elétrica brasileira no futuro, conforme indicado no Plano Decenal de Expansão de Energia 2020-2029 (MME & EPE 2020), segundo o qual estima-se que em 2029 a capacidade instalada da fonte eólica aumente mais de 2,5 vezes, atingindo 39.500 MW, i.e., 17,3% de toda a capacidade instalada do Brasil.

Apesar das vantagens da tecnologia eólica (e solar), a intermitência da geração eólica horária, dada a natureza intrínseca da variabilidade dos ventos, constitui um desafio em termos de sua integração a sistemas de energia elétrica. Os valores de geração apresentam ampla dispersão e são sujeitos a variações horárias, mensais e anuais, e, portanto, podem ser de difícil previsão. Uma capacidade inadequada de prever a produção de energia eólica nos diversos horizontes de tomada de decisão afeta adversamente a estabilidade, a confiabilidade, o planejamento e a programação da operação do sistema de energia, e também o seu benefício econômico. Assim, torna-se fundamental o desenvolvimento e o aprimoramento de metodologias para representar as incertezas das fontes renováveis intermitentes – eólica e solar, incluindo-se os modelos de planejamento da expansão e da operação, em seus diversos horizontes de tomada de decisão (MELO *et al.*, 2020; MACEIRA *et al.* 2020; PESSANHA *et al.*, 2020).

Em muitos países, o planejamento da expansão e da operação em sistemas com predominância hidroelétrica tem sido realizado desagregando os problemas em horizontes específicos (MACEIRA *et al.* 2002; HELSETH & MELO 2020). No caso do Brasil, o problema de planejamento da expansão e da operação do 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; MACEIRA *et al.* 2016; HELSETH & MELO 2020).

Um dos modelos-chave desta cadeia para as atividades de planejamento da expansão e da operação energética do sistema interligado é o modelo NEWAVE (MACEIRA *et al.* 2008; MACEIRA *et al.* 2018) que, desde 1998, vem sendo utilizado em estudos oficiais pelas entidades setoriais (Operador Nacional do Sistema – ONS, Câmara de Comercialização de Energia Elétrica – CCEE, Empresa de Pesquisa

Energética – EPE, Agência Nacional de Energia Elétrica – ANEEL, e Ministério de Minas e Energia – MME), bem como por agentes e concessionárias de energia elétrica.

No modelo NEWAVE, o problema de planejamento da operação é representado como um problema de programação linear estocástico multi-estágio, cuja função objetivo é minimizar o valor esperado do custo total de operação ao longo de todo o período de planejamento, considerando mecanismos de aversão a risco, dado um estado inicial do sistema (armazenamento e tendência hidrológica). Os custos de combustível das usinas termoeletricas e penalidades para o não atendimento ao mercado consumidor compõem o custo de operação. A solução deste problema resulta em uma estratégia operativa que, para cada estágio do período de planejamento, dado o estado inicial do sistema no estágio, produz metas de geração de energia e de intercâmbios entre subsistemas.

No NEWAVE, os vários reservatórios do sistema podem ser agregados em reservatórios equivalentes de energia (REEs) (ARVANTIDIS & ROSING, 1970, TERRY *et al.*, 2004) ou representados por meio de uma modelagem híbrida, permitindo que se represente as usinas hidroelétricas (UHEs) individualmente, em todo ou em parte do seu horizonte de planejamento (MACEIRA *et al.*, 2019a, 2019b). Por sua vez, o estado do sistema inclui a energia armazenada dos reservatórios equivalentes e informações da "tendência hidrológica", e.g., as últimas  $p$  energias ou vazões naturais afluentes a cada reservatório, individual ou equivalente.

A representação das afluições aos reservatórios hidroelétricos no modelo NEWAVE é estocástica, por meio de uma árvore de cenários, onde cada caminho na árvore é denominado cenário hidrológico, e cada nó representa uma possível realização da afluição. Essas sequências seguem um processo estocástico multivariado, temporal e espacialmente correlacionado, com propriedades estatísticas semelhantes ao registro histórico, que são preservadas durante a construção da árvore. Para obter as afluições para o problema de otimização, é utilizado um modelo auto-regressivo periódico, ou seja, o valor obtido para a variável aleatória em um determinado período é função das afluições dos períodos anteriores.

Dado que não é possível percorrer a árvore completa de cenários de afluições, o algoritmo de Programação Dinâmica Dual Estocástica – PDDE (PEREIRA & PINTO 1991; MACEIRA, 1993) é aplicado para resolver o problema. A estratégia de solução consiste em percorrer iterativamente uma amostra (sub-árvore) de cenários de afluições, por meio de dois passos: (i) uma simulação *forward*, do estágio  $t=1$  até  $t=T$ , onde os cenários hidrológicos são amostrados com a finalidade de gerar novos estados, para os

quais a função de custo futuro será avaliada e novos cortes de Benders serão construídos na próxima recursão *backward*/ (ii) uma recursão *backward*, do estágio  $t=T$  até  $t=1$ , onde os cortes de Benders que compõem a função de custo futuro são construídos para todos os nós da sub-árvore resultantes da última simulação *forward*, representando a estratégia de operação.

Após o cálculo da estratégia de operação, é realizada uma simulação final adotando outros 2000 cenários hidrológicos. A partir dos resultados desta simulação final são produzidas estatísticas (valores esperados, desvio padrão, distribuições de frequências, etc) associadas a diversos indicadores de desempenho do sistema, entre os quais, custo total de operação, custo marginal de operação, risco de déficit, déficit de energia, geração hidrelétrica e térmica, vertimentos, etc (MACEIRA *et al.*, 2002, 2016).

Atualmente, em consonância com a Resolução ANEEL N° 843/2019, a representação da geração eólica no modelo NEWAVE é realizada de forma simplificada, com base na média mensal do histórico dos últimos cinco anos de geração líquida disponibilizada ao SIN de cada usina, agregada por subsistema, por mês e por patamar de carga, para todo o horizonte de planejamento. Conforme já mencionado, a perspectiva de maior participação da geração eólica na matriz elétrica brasileira aponta para a necessidade de aprimorar a representação atual na cadeia de modelos computacionais, incluindo o NEWAVE.

Com este intuito, encontra-se em desenvolvimento uma metodologia para a representação das incertezas associadas às fontes eólicas no modelo NEWAVE (MACEIRA *et al.* 2020), aderentes às características intrínsecas das modelagens adotadas, com o objetivo de manter o problema estocástico e de grande porte tratável computacionalmente.

A abordagem proposta é composta por quatro etapas interligadas, com desenvolvimentos simultâneos (MACEIRA *et al.* 2020): (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 (vide ANDRADE *et al.*, 2021); (iii) geração de séries sintéticas de ventos/produção eólica correlacionadas às afluências aos reservatórios das hidroelétricas; e (iv) representação das produções eólicas mensais no algoritmo PDDE do modelo NEWAVE. O presente trabalho está inserido na etapa (iii).

Na implementação atual do algoritmo PDDE, um modelo auto-regressivo periódico - PAR(p) (SALAS *et al.*, 1980; MACEIRA & BEZERRA, 1997; FRANSES & PAAP, 2006; PENNA *et al.*, 2011) - é usado para gerar os cenários de afluências de

energia/vazões que são usados nos passos *forward* e *backward* do algoritmo PDDE e também na simulação da operação do sistema com a política de operação calculada. Neste modelo de séries temporais, as últimas  $p$  afluições compõem a equação de regressão. Uma amostra de cardinalidade elevada de resíduos normais não correlacionados espacialmente é gerada usando amostragem aleatória simples, e o método *K-means* é aplicado para reduzir a cardinalidade da amostra original. As correlações espaciais entre as afluições aos REEs/UHEs são utilizadas para tornar os resíduos correlacionados e uma distribuição Lognormal com três parâmetros (CHARBENEAU, 1978) é ajustada aos resíduos espacialmente correlacionados. Assim, os cenários de afluições gerados são multivariados e apresentam uma distribuição assimétrica semelhante ao processo de afluições original.

Como um dos condicionantes da metodologia proposta é não estender o número de variáveis de estado do algoritmo PDDE, a fim de manter o problema ainda tratável computacionalmente, a velocidade dos ventos será modelada como uma regressão nas afluições do período atual ou anteriores mais um resíduo. Os resíduos de velocidade dos ventos recebem o mesmo tratamento que os resíduos das afluições, mas agora a estrutura de correlação espacial também compreende, além da correlação espacial entre usinas hidrelétricas, as correlações entre parques eólicos e a correlação cruzada entre as afluições às hidrelétricas e as velocidades do vento dos parques eólicos.

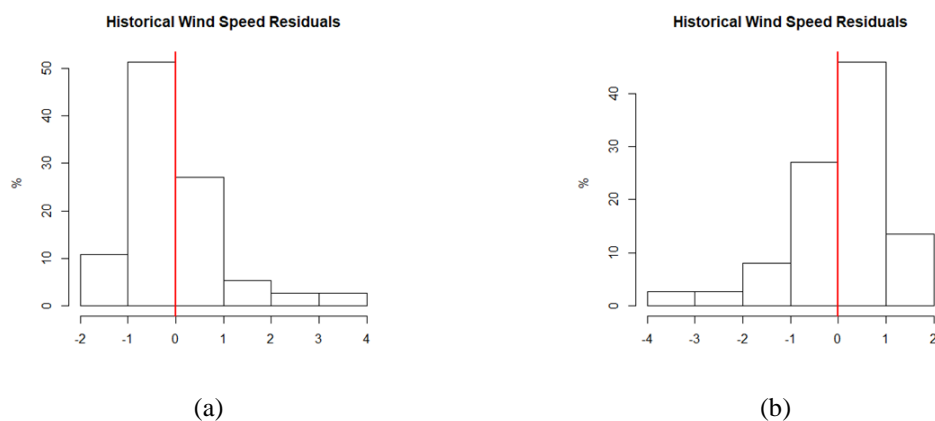
Porém, diferentemente dos resíduos de afluição, os resíduos das velocidades do vento em nível mensal podem apresentar assimetrias positivas (Figura 1a) como também assimetrias negativas (Figura 1b) em vários meses, o que impede, neste último caso, o emprego da distribuição Lognormal. Neste caso, uma alternativa é usar a distribuição Weibull (WEIBULL, 1939a,b; 1951), que é bastante versátil e é definida para assimetrias à direita e à esquerda.

Adicionalmente, os resíduos das velocidades mensais dos ventos apresentam, por construção, valores negativos, o que implica na necessidade de se 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 (ROCKETTE *et al.*, 1974; COHEN & WHITTEN, 1982; COHEN *et al.*, 1984; CHENG & AMIN, 1983; SMITH, 1985; GREEN *et al.*, 1994; COUSINEAU, 2019a,b), a maioria baseados em modificações dos métodos dos momentos ou da máxima verossimilhança. Porém, conforme verificado e também apresentado na Seção 5, a qualidade das estimativas destes métodos aplicados aos parques eólicos brasileiros varia muito em função do mês do ano e da localização dos

parques eólicos, sendo os piores desempenhos associados aos meses com assimetrias negativas elevadas.

Figura 1.1 – Exemplos de resíduos das médias mensais das velocidades de ventos em parques eólicos no Brasil: (a) agosto - litoral de Pernambuco; (b) junho - interior da Bahia



Fonte: Autores

O objetivo deste trabalho é apresentar uma abordagem para a modelagem de resíduos de velocidades mensais de ventos por meio de distribuições Weibull tri-paramétricas, na geração de cenários de ventos mensais correlacionados com afluências às usinas hidroelétricas, quando da utilização da técnica de Programação Dinâmica Dual Estocástica (PDDE). A abordagem proposta busca preservar a média, o desvio padrão e a assimetria das velocidades históricas de ventos mensais, sendo especialmente adequada em situações de assimetrias elevadas.

A metodologia proposta é aplicada a um conjunto de parques eólicos situados nas regiões Nordeste e Sul do Brasil, e o seu desempenho é avaliado em relação a um conjunto de 17 métodos disponíveis na literatura.

## 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, 1939a,b). 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/control 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:

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

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:

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

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 a seguir.

## 2.1 O Parâmetro de Forma $\gamma$

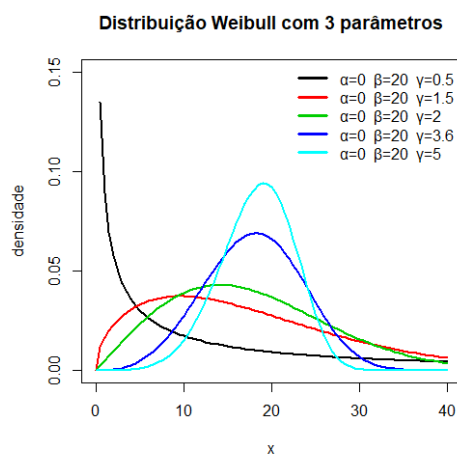
O parâmetro de forma ( $\gamma$ ) descreve como os dados estão distribuídos. Para  $\gamma < 1$ , a distribuição assume a forma hiperexponencial ou “J-invertido”, e com  $\gamma = 1$ , assume o formato da distribuição exponencial. Já para  $\gamma > 1$  passa a ter o formato de sino, sendo que com  $\gamma = 3,6$  se aproxima de uma curva normal, i.e., simétrica. Quando  $\gamma < 3,6$ , a distribuição Weibull apresenta assimetria positiva, e para  $\gamma > 3,6$  a assimetria passa a ser negativa. A Figura 2.1 ilustra o efeito deste parâmetro.

## 2.2 O Parâmetro de Escala $\beta$

Para uma dada forma da distribuição Weibull, o parâmetro de escala ( $\beta$ ) está relacionado com a amplitude e dispersão da curva, sendo que à medida que  $\beta$  aumenta a curva tende a se tornar mais dispersa, conforme ilustrado na Figura 2.2. Adicionalmente, por corresponder ao percentil de 63,21% dos dados, a escala define a posição da curva

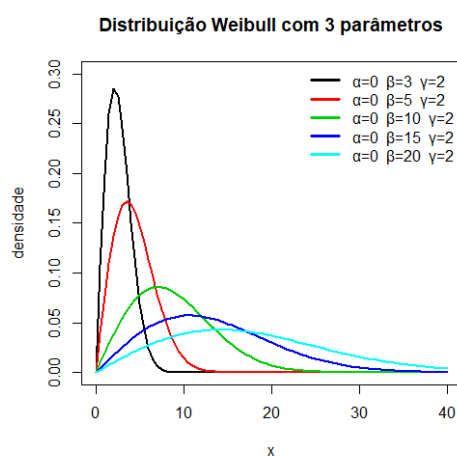
Weibull em relação ao parâmetro posição, sendo também denominado de vida característica, na análise de confiabilidade.

Figura 2.1. Efeito do parâmetro de *forma* da distribuição Weibull



Fonte: Autores

Figura 2.2. Efeito do parâmetro de *escala* da distribuição Weibull



Fonte: Autores

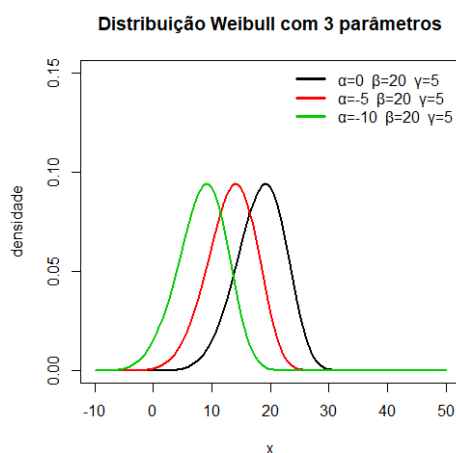
### 2.3 O Parâmetro de Posição $\alpha$

O parâmetro de posição ( $\alpha$ ) - também denominado de deslocamento ou limite - descreve o deslocamento da distribuição com relação ao valor zero. Como todos os dados devem ser maiores que este limite, na ocorrência de dados com valores negativos, este parâmetro também assume um valor negativo.



Conforme exemplificado na Figura 2.3, um limite negativo desloca a distribuição para a esquerda e um limite positivo desloca a distribuição para a direita. Todos os dados devem ser maiores que o limite. A distribuição Weibull de 2 parâmetros é a mesma que o Weibull de 3 parâmetros com um limite igual a zero. Por exemplo, a distribuição Weibull de 3 parâmetros  $W(3;100;50)$  tem a mesma forma e propagação que a distribuição Weibull de 2 parâmetros  $W(3;100)$ , mas seu limite é deslocado em 50 unidades para a direita

Figura 2.3. Efeito do parâmetro de *posição* da distribuição Weibull



Fonte: Autores

### 3. Principais Métodos para Ajustes de Distribuições Weibull com 3 Parâmetros

Dentre os principais métodos para estimação dos parâmetros destacam-se o estimador de máxima verossimilhança (MLE) e o método dos momentos (MoM), os quais são discutidos nesta Seção.

#### 3.1 Estimador de Máxima Verossimilhança (MLE)

Seja  $X = \{x_i\}$ ,  $i = 1, 2, \dots, n$ , uma amostra aleatória ordenada de uma distribuição com função densidade de probabilidade dada por (2). A função de verossimilhança dos parâmetros da distribuição Weibull para esta amostra aleatória  $X$  de tamanho  $n$  é dada por (COHEN & WHITTEN, 1982; COUSINEAU, 2019a):

$$L(\gamma, \beta, \alpha | X) = \prod_{i=1}^n f(x_i | \gamma, \beta, \alpha) \quad (3)$$

As soluções MLE não estão disponíveis de forma direta, mesmo para o caso de dois parâmetros. De uma maneira geral, aplica-se o logaritmo à função de verossimilhança (3), obtendo-se

$$\log(L(\gamma, \beta, \alpha|X)) = -n\gamma \log(\beta) + n \log(\gamma) + (\gamma - 1) \sum_{i=1}^n \log(x_i - \alpha) - \beta^{-\gamma} \sum_{i=1}^n (x_i - \alpha)^{\gamma} \quad (4)$$

Os estimadores de máxima verossimilhança podem ser encontrados resolvendo simultaneamente o sistema de equações não lineares, obtido igualando à zero as derivadas parciais da função log-verossimilhança (4) em relação aos parâmetros.

Por exemplo, no caso da derivada em relação ao parâmetro de escala  $\beta$ , a equação resultante pode ser reorganizada para isolar este parâmetro, conforme (5). Entretanto, a equação (5) ainda requer o conhecimento dos outros dois parâmetros a serem calculados.

$$\beta = \sqrt[\gamma]{\frac{1}{n} \sum_{i=1}^n (x_i - \alpha)^{\gamma}} \quad (5)$$

Derivando-se (4), agora em relação a  $\gamma$ , e substituindo  $\beta$  por (5), chega-se a:

$$\frac{1}{\gamma} + \frac{1}{n} \sum_{i=1}^n \log(x_i - \alpha) - \frac{\sum_{i=1}^n \log(x_i - \alpha)(x_i - \alpha)^{\gamma}}{\sum_{i=1}^n (x_i - \alpha)^{\gamma}} \quad (6)$$

De maneira similar para  $\alpha$ , obtém-se:

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i - \alpha} \cdot \frac{\sum_{i=1}^n (x_i - \alpha)^{\gamma}}{\sum_{i=1}^n (x_i - \alpha)^{\gamma-1}} - \frac{\gamma}{\gamma-1} = 0 \quad (7)$$

Os estimadores  $\hat{\gamma}_{MLE}$  e  $\hat{\alpha}_{MLE}$  podem ser obtidos resolvendo simultaneamente as equações transcendentais (6) e (7), e a partir deles calcular  $\{\hat{\beta}\}_{MLE}$  utilizando (5).

### 3.2 Método dos Momentos (MoM)

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

$$E(x) = \alpha + \beta \Gamma_1 \quad (8)$$

$$Var(x) = \beta^2 (\Gamma_2 - \Gamma_1^2) \quad (9)$$

$$Sk(x) = \frac{\Gamma_3 - 3\Gamma_2\Gamma_1 + 2\Gamma_1^3}{(\Gamma_2 - \Gamma_1^2)^{3/2}} \quad (10)$$

onde:

$$\Gamma_k(\gamma) = \Gamma(1 + k/\gamma) \quad (11)$$

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

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \quad (12)$$

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), mas 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 (10) 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:

$$\frac{\Gamma_3 - 3\Gamma_2\Gamma_1 + 2\Gamma_1^3}{(\Gamma_2 - \Gamma_1^2)^{3/2}} - \hat{g} = 0 \quad (13)$$

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

$$\hat{\beta}_{MoM} = s / \sqrt{\Gamma_2(\hat{\gamma}_{MoM}) - \Gamma_1^2(\hat{\gamma}_{MoM})} \quad (14)$$

Finalmente,  $\hat{\alpha}_{MoM}$  pode ser computado por (8), empregando a média amostral  $\bar{x}$ :

$$\hat{\alpha}_{MoM} = \bar{x} - \hat{\beta}_{MoM} \Gamma_2(\hat{\gamma}_{MoM}) \quad (15)$$

### 3.3 Considerações sobre os Métodos MLE e MoM e Abordagens Derivadas

Estimadores de momento (MoM) e estimadores de máxima verossimilhança (MLE) têm sido considerados os estimadores tradicionais para parâmetros de distribuições de probabilidade, incluindo a Weibull.

Entretanto, em determinadas situações, estes métodos podem apresentar dificuldades teóricas e resultados insatisfatórios quando se consideram os três parâmetros da Weibull, conforme discutido a seguir.

#### 3.3.1 Considerações sobre Métodos MLE

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 (SMITH, 1985). No entanto, para algumas combinações de parâmetros, as condições de regularidade não são satisfeitas para a estimativa de máxima verossimilhança da distribuição Weibull com três parâmetros, porque o domínio da variável aleatória depende da posição do limite inferior - o parâmetro  $\gamma$  - que é desconhecido (ROCKETTE *et al.*, 1974).

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).

Por estas razões, várias abordagens, baseadas em modificações do MLE e também em métodos alternativos, têm sido buscadas na literatura para as distribuições Weibull com três parâmetros, e frequentemente exibem vieses e variâncias menores do que as estimativas de máxima verossimilhança.

No primeiro conjunto estão os cinco métodos MLE Modificados propostos em (COHEN & WHITTEN, 1982), que substituem a derivada em relação a  $\gamma$  (6) por uma relação funcional alternativa, e o Estimador de Máxima Verossimilhança Ponderada (*Weighted Maximum Likelihood Estimation* - w-MLE) (COUSINEAU, 2009b). Como exemplo do segundo conjunto, pode-se citar o Máximo Produto de Espaçamento (*Maximum Product of Spacing* – MPS), onde a função de verossimilhança é substituída por uma função de espaçamento (CHENG & AMIN, 1983).

### 3.3.2 Considerações sobre Métodos MoM

Os estimadores de momento para distribuições Weibull são estimadores bem conhecidos que tendem a superar as desvantagens da abordagem MLE (NEWBY, 1984), mas sua existência, unicidade e consistência são, mais uma vez, questões em aberto.

Adicionalmente, as estatísticas de momentos superiores (por exemplo, assimetria em oposição à média) 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 (assimetria) por outra relação menos variável. Exemplos destas abordagens são os três métodos MoM Modificados propostos em (COHEN & WHITTEN, 1982) e o desenvolvido em (COUSINEAU, 2019a), que incluem, por exemplo, a noção de mínimo esperado de uma amostra,  $E(\min(X))$ .

## 4. Abordagem Proposta para a Modelagem de Distribuições Weibull Tri-Paramétricas com Elevadas Assimétrias

Conforme apresentado na Seção 3, vários métodos estão disponíveis na literatura para estimar os parâmetros de forma, escala e posição de distribuições Weibull, a maioria baseados em modificações dos métodos dos momentos (MoM) ou da máxima verossimilhança (MLE).

Entretanto, quando estes métodos foram aplicados a dados de diversos parques eólicos brasileiros, em escala mensal, os desempenhos associados foram irregulares, variando bastante em função do mês do ano e da localização dos parques eólicos, e muitas vezes insatisfatórios, sendo os piores desempenhos associados aos meses com assimétrias negativas elevadas (vide Seção 5). Este comportamento suscitou a necessidade do desenvolvimento de uma abordagem eficiente para o caso brasileiro.

Como o objetivo final deste trabalho é a modelagem de ventos para a geração de cenários sintéticos de velocidades mensais de ventos em um esquema de Programação Dinâmica Dual Estocástica (PDDE), e sem aumentar o número de variáveis de estado da PDDE, a velocidade dos ventos será modelada como uma regressão das afluências às usinas hidroelétricas (do período atual ou anteriores) mais um resíduo. Em consequência, a distribuição Weibull tri-paramétrica será ajustada aos resíduos das médias mensais de velocidades de ventos, correlacionados e normalmente distribuídos. Portanto, os resíduos de velocidades mensais dos ventos recebem o mesmo tratamento que os resíduos das afluências, contudo a estrutura de correlação espacial também compreende, além da

correlação espacial entre usinas hidroelétricas, as correlações entre parques eólicos e as correlações cruzadas entre as afluências às hidroelétricas e as velocidades do vento dos parques eólicos (MACEIRA *et al.* 2020).

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.

Procurou-se, ainda, evitar que a modelagem das distribuições Weibull resultasse em estimativas inconsistentes ou imprecisas. Para isso, optou-se pelo MoM, porém considerando os momentos até a segunda ordem, i.e., média (8) e variância (9).

Considerando essas condições de contorno, a idéia básica da abordagem proposta é partir de um valor inicial para a estimativa do parâmetro de posição  $\hat{\alpha}$ , calcular as estimativas  $\hat{\gamma}$  e  $\hat{\beta}$  por meio de um MoM mais simples (vide Seção 4.1) e, de forma iterativa, atualizar a estimativa  $\hat{\alpha}$  com vistas a reduzir a diferença entre a assimetria das velocidades de ventos sintéticas (geradas) e a assimetria das velocidades de ventos amostrais (históricas).

#### 4.1 MoM quando o Parâmetro de Posição $\alpha$ é conhecido (MoM-2p)

Quando a posição  $\alpha$  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 (10). Neste método, aqui denominado MoM-2p, consiste em estimar o parâmetro de forma  $\hat{\gamma}$  com base no coeficiente de variação obtido a partir de (8) e (9), porém substituindo a média e a variância populacionais pelos respectivos valores amostrais. Para isso, resolve-se a seguinte equação transcendental:

$$\frac{\Gamma_2(\hat{\gamma})}{\Gamma_1^2(\hat{\gamma})} - 1 - \frac{s^2}{\bar{x} - \alpha} = 0 \quad (16)$$

Uma vez obtido  $\hat{\gamma}$ , a estimativa  $\hat{\beta}$  pode ser obtida de (8):

$$\hat{\beta} = \frac{(\bar{x} - \alpha)}{\Gamma_1(\hat{\gamma})} \quad (17)$$

## 4.2 Algoritmo Proposto

De maneira resumida, o algoritmo proposto é constituído dos seguintes passos:

- i. Leitura dos dados amostrais e da distribuição de resíduos correlacionados e normalmente distribuídos. Definição de parâmetros de controle e inicialização de variáveis;
- ii. Defina um valor inicial para  $\hat{\alpha}$ ,  $\hat{\alpha}_{inic}$ ;
- iii. Atualize o valor de  $\hat{\alpha}$  por meio de interpolações lineares, considerando os valores prévios deste parâmetro e das assimetrias;
- iv. Compute as estimativas de  $\hat{\gamma}$  e  $\hat{\beta}$  pelo método MoM -2p;
- v. Gere os resíduos sintéticos Weibull e as velocidades de ventos associadas, usando os resíduos lidos, i.e., resíduos correlacionados e normalmente distribuídos;
- vi. Calcule a assimetria das velocidades de ventos sintéticas
- vii. Calcule o erro relativo entre as assimetrias histórica e sintética das velocidades de ventos;
- viii. Se o erro relativo < tolerância, vá para o passo (iii);
- ix. Fim do algoritmo.

## 4.3 Considerações sobre o Valor Inicial para o Parâmetro Posição - $\hat{\alpha}_{inic}$

A aplicação do algoritmo proposto a diversos parques eólicos brasileiros, e em meses distintos, evidenciou que, apesar de se tratar de um algoritmo heurístico, na maioria das vezes a convergência ocorre rapidamente sem uma seleção cuidadosa do ponto de partida no processo iterativo (valor para  $\hat{\alpha}_{inic}$ ).

Todavia, em situações de assimetrias históricas das velocidades mensais dos ventos negativas e com módulos elevados, a convergência foi influenciada pelo ponto de partida. Assim, procurou-se identificar uma abordagem para estabelecer um valor inicial para  $\hat{\alpha}$  que fosse adequado para estas situações e, ao mesmo tempo, robusto para as demais.

Com este fim, coletou-se um conjunto de valores finais de  $\hat{\alpha}$ , obtidos a partir da modelagem de diversos parques eólicos brasileiros, aplicando o algoritmo proposto, para estabelecer uma ou mais equações de regressão linear.

Para definir as variáveis explicativas, levou-se em conta que: (i) uma variável subjacente aos resíduos é a velocidade dos ventos, cuja média histórica é um componente

do cálculo dos resíduos; (ii) no caso da distribuição Weibull tri-paramétrica, um dado relevante, que inclusive influencia o  $\alpha$ , é o valor mínimo observado; e (iii) um dos objetivos básicos da abordagem proposta, constituindo-se em critério de convergência, é aproximar as assimetrias das velocidades de ventos sintéticas e históricas. Com estas considerações, três variáveis explicativas emergiram: média, valor mínimo e assimetria históricas das velocidades mensais dos ventos.

Assim, para cada parque eólico analisado, sugere-se ajustar a seguinte equação de regressão linear para a definição de  $\hat{\alpha}_{inic}$ :

$$\hat{\alpha}_{inic} = a [(\bar{x}_v - x_v^{min})g_v] + b \quad (18)$$

onde:

$\bar{x}_v$  = média histórica das velocidades mensais dos ventos;

$x_v^{min}$  = valor mínimo histórico das velocidades mensais dos ventos;

$g_v$  = assimetria histórica das velocidades mensais dos ventos;

$a$  e  $b$  = coeficientes angular e linear da equação de regressão linear.

A partir dos dados utilizados, verificou-se que seria adequado estabelecer regressões lineares distintas para valores positivos e negativos de assimetrias históricas, assim como segmentar as equações de regressão linear em duas partes, de acordo com os valores das assimetrias históricas dos ventos. Por exemplo, no estudo de casos apresentado na Seção 5 foram utilizadas as equações de regressão apresentadas na Tabela 4.1.

Tabela 4.1 – Valores dos coeficientes angular ( $a$ ) e linear ( $b$ ) das equações de regressão lineares para determinação de  $\hat{\alpha}_{inic}$

Coeficientes	Assimetrias históricas dos ventos ( $g_v$ ) negativas		Assimetrias históricas dos ventos ( $g_v$ ) positivas	
	$-1 \leq g_v < 0$	$g_v < -1$	$0 \leq g_v \leq 1$	$g_v > 1$
$a$	5,51296	245.81143	1.67971	0.02822
$b$	-2,83252	217.16514	-2.94423	-1.41903

Fonte: Autores

## 5. Aplicação da Metodologia Proposta e Comparação com Métodos Existentes

A metodologia apresentada em (MACEIRA *et al.* 2020) propõe um modelo integrado de geração de séries sintéticas de afluições às usinas hidroelétricas e ventos nos parques eólicos, baseado em modelos auto-regressivos periódicos - PAR(p). Este



modelo integrado é utilizado nos passos *forward* e *backward* do algoritmo PDDE e também na simulação da operação do sistema com a política de operação calculada. Para isso, entre outros, faz-se necessário o ajuste de distribuições Weibull tri-paramétricas aos dados mensais de velocidades de ventos.

Uma questão importante no desenvolvimento e testes da metodologia proposta em (MACEIRA *et al.*, 2020), diz respeito à disponibilização de dados verificados (medidos). Idealmente, seria relevante que se tivesse acesso às séries horárias consistidas e pareadas de velocidades de ventos e produções eólicas, para todos os parques existentes, ou pelo menos para aqueles vencedores nos leilões públicos de compra de energia elétrica. No entanto, as instituições depositárias desses dados verificados (Empresa de Pesquisa Energética - EPE e Câmara de Comercialização de Energia Elétrica - CCEE) têm declarado que, por questões de confidencialidade com os proprietários dos dados (os parques eólicos), a sua disponibilização não é possível.

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 (*National Aeronautics and Space Administration global* – NASA) e ERA5 (*European Centre for Medium-Range Weather Forecasts* – ECMWF).

No caso específico deste trabalho, as séries temporais de velocidades horárias de ventos são agrupadas por mês para a obtenção de médias mensais das velocidades de vento. A partir delas, são construídos modelos probabilísticos mensais Weibull com três parâmetros, conforme metodologia proposta na Seção 4.

Para a avaliação do desempenho da abordagem proposta foram utilizados tres parques eólicos, selecionados com o intuito de considerar as regiões brasileiras com maior potencial e desenvolvimento de energia eólica (nordeste e sul); distintos regimes de ventos (na costa e no interior); a sazonalidade dos regimes de ventos (meses abril, junho, julho e agosto); assimetrias amostrais positivas e, especialmente, negativas, incluindo módulos elevados.

Para cada parque estudado, foram considerados 37 anos de velocidades médias mensais de vento (obtidas por reanálise) e resíduos correspondentes, estes calculados utilizando como componente explicativa a própria média das velocidades. Foram considerados 2.000 resíduos normalmente distribuídos e correlacionados, através de uma amostragem aleatória simples, sendo que em um dos casos foi também analisado o impacto de utilizar amostras de resíduos com cardinalidades iguais a 200, 50 e 20, obtidos empregando-se técnicas de análise de agrupamentos, conforme (PENNA *et al.*, 2011).

Finalmente, o desempenho da metodologia proposta foi comparada com outros 17 métodos disponíveis na literatura e listados na Tabela 5.1, os quais foram implementados pelos autores ou se encontram disponíveis no *software R* (R CORE TEAM, 2019).

Tabela 5.1 – Métodos utilizados na estimação das distribuições Weibull tri-paramétricas.

MÉTODOS APLICADOS
M1 - Algoritmo Proposto
M2 - MoM (Cohen, 1982)
M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)
M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)
M5 - Recálculo do MoM com $\alpha$ dado por $E[\text{Min de uma amostra}]$
M6 - MLE Ajustado - Pacote R Weibullness
M7 - MoM (Cran, 1988) - Pacote R ForestFit
M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
M11 - MLE (Green, 1994) - Pacote R ForestFit
M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit
M16 - Máximo Produto de Espaçamento - Pacote R ForestFit
M17 - MLE Ponderado (Cousineau, 2009) - Pacote R ForestFit
M18 - Momentos T-L - Pacote R ForestFit

Fonte: Autores

### 5.1 Caso NE\_PE – Agosto ( $\hat{g} = +1.227$ )

O primeiro caso estudado refere-se a um parque eólico situado no litoral do nordeste, mais especificamente, de Pernambuco. Os histogramas dos 37 valores das médias mensais das velocidades de vento para o mês de agosto, e dos resíduos históricos associados estão apresentados nas Figuras 5.1 (a) e (c), respectivamente. As estatísticas média, desvio-padrão e assimetria amostrais são, respectivamente, 5,736; 0,288 e 1,227. Verifica-se, portanto, que a assimetria é positiva e elevada.

Os resultados da aplicação do método proposto e dos demais 17 considerados na estimação dos três parâmetros da distribuição Weibull estão exibidos na Tabela 5.2, assim como as estatísticas dos valores sintéticos das velocidades mensais de ventos, considerando 2.000 resíduos correlacionados e normalmente distribuídos. Nesta tabela (e nas demais tabelas de resultados) os métodos estão ordenados do melhor para o pior desempenho, adotando o erro relativo entre as assimetrias histórica e sintética das

velocidades de ventos como critério de parada. Como critérios adicionais também são adotados os erros relativos dos desvios- padrão e das médias, neste ordem.

A seguinte codificação de cores é ainda utilizada: verde - método com melhor desempenho; azul - método com segundo melhor desempenho; vermelho - métodos que apresentam o melhor ou segundo melhor desempenho em outros casos estudados, mas que no caso em análise apresentam desempenhos não satisfatórios ou inaceitáveis.

Verifica-se na Tabela 5.2 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%), inclusive com relação ao desvio-padrão e à média. O método M9 foi o segundo melhor método pois, embora tenha apresentado erro relativo = 0% em relação à assimetria, o erro no desvio-padrão (2,8%) foi superior ao do método proposto (1,4%). Nota-se também que 9 métodos apresentaram erro relativo inferior a 10%; 9 acima de 20% e 7 acima de 75%, sendo que dois deles apresentaram assimetrias próximas de 7,000, valor este bem acima do valor histórico. Há ainda 4 métodos (M10, M14, M15 e M18) que produziram desempenhos não satisfatórios ou inaceitáveis, e que ocupam a segunda posição em desempenho em outros casos analisados.

A boa aderência dos valores sintéticos produzidos pelo método proposto aos valores históricos pode também ser constatada por meio da Figura 5.1, tanto para os resíduos – (c) e (d), quanto para as médias mensais das velocidades de ventos – (a) e (b).

Tabela 5.2 – Resultados para o Caso NE\_PE – Agosto: parâmetros estimados da distribuição Weibull; estatísticas dos 2000 valores sintéticos das velocidades mensais de ventos; erro relativo entre as assimetrias histórica e sintética das velocidades de ventos.

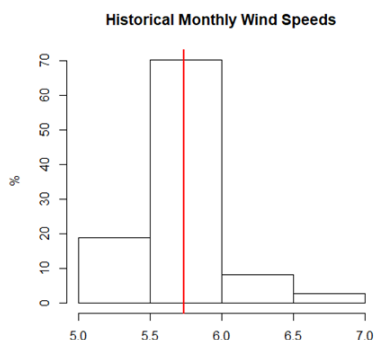
Parâmetros Weibull			Estatísticas - Valores Sintéticos				
Forma $\gamma$	Escala $\beta$	Posição $\alpha$	Média	D.P.	Assimetria	Erro Rel (%)	Método
1.410	1.528	-1.391	5.737	0.292	1.227	0	M1 - Algoritmo Proposto
1.418	1.558	-1.417	5.737	0.296	1.216	0	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
1.396	1.517	-0.609	5.961	0.293	1.247	2	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
1.434	1.577	-1.432	5.737	0.296	1.194	3	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
1.445	1.291	-1.423	5.664	0.240	1.179	4	M2 - MoM (Cohen, 1982)
1.450	1.548	-1.401	5.738	0.287	1.173	4	M6 - MLE Ajustado - Pacote R Weibullness
1.352	1.459	-1.338	5.737	0.292	1.311	7	M5 - Recálculo do MoM com $\alpha$ dado por $E[\text{Min de uma amostra}]$
1.350	1.419	-1.294	5.739	0.285	1.315	7	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
1.335	1.577	-1.408	5.749	0.320	1.338	9	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit
1.322	1.444	-1.332	5.737	0.297	1.358	11	M11 - MLE (Green, 1994) - Pacote R ForestFit
1.228	1.332	-1.246	5.737	0.298	1.525	24	M18 - Momentos T-L - Pacote R ForestFit
0.978	0.571	-0.577	5.737	0.173	2.187	78	M7 - MoM (Cran, 1988) - Pacote R ForestFit
0.962	0.946	-1.338	5.63	0.294	2.248	83	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)
0.962	1.315	-1.338	5.738	0.409	2.248	83	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)
0.850	0.499	-0.543	5.738	0.189	2.784	127	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
8.220	2.900	-1.271	6.158	0.115	-0.567	146	M17 - MLE Ponderado (Cousineau, 2009) - Pacote R ForestFit
0.555	0.310	-0.521	5.739	0.310	6.958	467	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
0.553	1.884	-1.306	6.293	1.896	6.995	470	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit

Fonte: Autores

Figura 5.1 – Histogramas para o Caso NE\_PE – Agosto ( $\hat{\gamma} = +1.227$ ): 37 valores históricos de velocidade (a) e resíduos (c) de velocidades de vento; 2.000 valores sintéticos (algoritmo proposto) de velocidade (b)

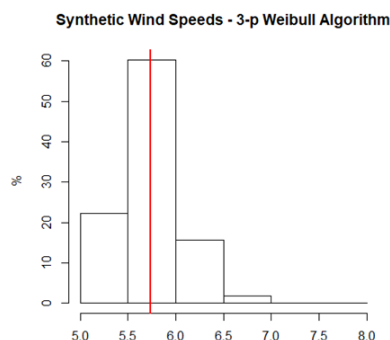
e resíduos (d) de velocidades de vento;

### Valores Históricos



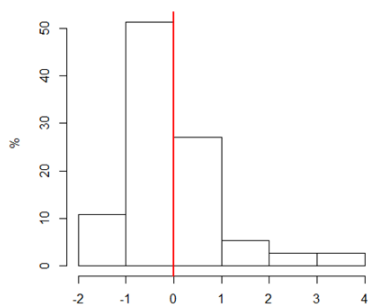
(a)

### Valores Sintéticos – Algoritmo Proposto



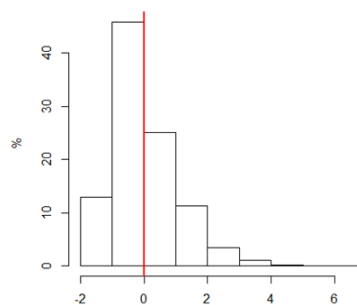
(b)

### Historical Wind Speed Residuals



(c)

### Synthetic Speed Residuals - 3-p Weibull Algorithm



(d)

Fonte: Autores

Em seguida foi realizada uma análise de sensibilidade com amostras de cardinalidades iguais a 200, 50 e 20, dos resíduos correlacionados e normalmente distribuídos, obtidos empregando-se técnicas de análise de agrupamentos.

Assim, a Tabela 5.3 exibe o resultado considerando agora 200 resíduos correlacionados e normalmente distribuídos. Observa-se, novamente, que o algoritmo proposto alcançou o melhor desempenho em relação às três estatísticas calculadas, e obteve o mesmo valor da assimetria histórica. Já o segundo melhor método foi o M18 (erro relativo = 8%), que ocupava a 11<sup>a</sup> posição no caso anterior; sendo que o M9 (erro relativo = 24%) passou para a 9<sup>a</sup> posição. Dos 18 métodos, 16 apresentam erros relativos superiores a 10%, e 6 abaixo de 20%. Há ainda 5 métodos (M8, M9, M10, M14 e M15) que produziram desempenhos não satisfatórios ou inaceitáveis, e que ocupam a segunda posição em desempenho em outros casos analisados.

Tabela 5.3 – Resultados para o Caso NE\_PE – Agosto: parâmetros estimados da distribuição Weibull;

estatísticas dos 200 valores sintéticos das velocidades mensais de ventos; erro relativo entre as assimetrias histórica e sintética das velocidades de ventos.

Parâmetros Weibull			Estatísticas - Valores Sintéticos					
Forma $\gamma$	Escala $\beta$	Posição $\alpha$	Média	D.P.	Assimetria	Erro Rel (%)	Método	
1.152	1.208	-1.149	5.763	0.299	1.226	0	M1 - Algoritmo Proposto	
0.555	1.332	-1.246	5.763	0.305	1.133	8	M18 - Momentos T-L - Pacote R ForestFit	
1.322	1.444	-1.332	5.763	0.303	1.030	16	M11 - MLE (Green, 1994) - Pacote R ForestFit	
1.35	1.577	-1.408	5.778	0.328	1.017	17	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit	
1.396	1.419	-1.294	5.765	0.291	1.002	18	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit	
0.85	1.459	-1.338	5.763	0.299	0.999	19	M5 - Recálculo do MoM com $\alpha$ dado por E[Min de uma amostra]	
0.553	0.571	-0.577	5.751	0.176	1.489	21	M7 - MoM (Cran, 1988) - Pacote R ForestFit	
1.335	1.517	-0.609	5.987	0.300	0.958	22	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit	
8.220	1.558	-1.417	5.764	0.302	0.937	24	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit	
1.228	0.946	-1.338	5.654	0.298	1.518	24	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)	
1.445	1.315	-1.338	5.772	0.414	1.518	24	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)	
0.962	1.577	-1.432	5.764	0.302	0.923	25	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit	
0.962	1.291	-1.423	5.686	0.245	0.913	26	M2 - MoM (Cohen, 1982)	
1.352	1.548	-1.401	5.764	0.293	0.909	26	M6 - MLE Ajustado - Pacote R Weibullness	
1.450	0.499	-0.543	5.752	0.190	1.749	42	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit	
0.978	0.310	-0.521	5.755	0.276	2.826	130	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit	
1.434	1.884	-1.306	6.391	1.686	2.833	131	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit	
1.418	2.900	-1.271	6.170	0.110	-0.410	133	M17 - MLE Ponderado (Cousineau, 2009) - Pacote R ForestFit	

Fonte: Autores

Por sua vez, a Tabela 5.4 mostra os resultados obtidos a partir dos 50 resíduos correlacionados e normalmente distribuídos. De novo, o método proposto alcançou o melhor desempenho em relação às três estatísticas calculadas, e obteve o mesmo valor da assimetria histórica. Já o segundo melhor método foi o M14 (erro relativo = 4%), que ocupava a 15ª posição nos dois casos anteriores, e com desempenhos inaceitáveis. Novamente, dos 18 métodos, 16 apresentam erros relativos superiores a 10%, e 5 abaixo de 20%. Agora, passou de 5 para 7 o número de métodos que produziram neste caso desempenhos não satisfatórios ou inaceitáveis, e que ocupam a segunda posição em outros casos analisados.

Tabela 5.4 – Resultados para o Caso NE\_PE – Agosto: parâmetros estimados da distribuição Weibull; estatísticas dos 50 valores sintéticos das velocidades mensais de ventos; erro relativo entre as assimetrias histórica e sintética das velocidades de ventos.

Parâmetros Weibull			Estatísticas - Valores Sintéticos					
Forma $\gamma$	Escala $\beta$	Posição $\alpha$	Média	D.P.	Assimetria	Erro Rel (%)	Método	
0.872	0.812	-0.869	5.728	0.207	1.227	0	M1 - Algoritmo Proposto	
0.555	0.499	-0.543	5.730	0.132	1.274	4	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit	
1.322	0.946	-1.338	5.624	0.215	1.058	14	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)	
1.35	1.315	-1.338	5.730	0.299	1.058	14	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)	
1.396	0.571	-0.577	5.734	0.127	1.031	16	M7 - MoM (Cran, 1988) - Pacote R ForestFit	
0.85	1.332	-1.246	5.741	0.234	0.685	44	M18 - Momentos T-L - Pacote R ForestFit	
0.553	1.444	-1.332	5.743	0.236	0.582	52	M11 - MLE (Green, 1994) - Pacote R ForestFit	
1.335	1.577	-1.408	5.757	0.256	0.569	54	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit	
8.22	1.419	-1.294	5.746	0.228	0.554	55	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit	
1.228	1.459	-1.338	5.744	0.234	0.552	55	M5 - Recálculo do MoM com $\alpha$ dado por E[Min de uma amostra]	
1.445	1.517	-0.609	5.969	0.236	0.509	58	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit	
0.962	1.558	-1.417	5.746	0.239	0.489	60	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit	
0.962	1.577	-1.432	5.746	0.239	0.474	61	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit	
1.352	1.291	-1.423	5.672	0.194	0.463	62	M2 - MoM (Cohen, 1982)	
1.450	1.548	-1.401	5.747	0.232	0.459	63	M6 - MLE Ajustado - Pacote R Weibullness	
0.978	0.310	-0.521	5.709	0.158	2.210	80	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit	
1.434	1.884	-1.306	6.110	0.961	2.216	80	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit	
1.418	2.900	-1.271	6.171	0.101	-1.009	182	M17 - MLE Ponderado (Cousineau, 2009) - Pacote R ForestFit	

Fonte: Autores

Finalmente, a Tabela 5.5 apresenta os resultados gerados com os 20 resíduos correlacionados e normalmente distribuídos. Como das outras vezes, o algoritmo proposto alcançou o melhor desempenho em relação às três estatísticas calculadas, e obteve o mesmo valor da assimetria histórica. Já na 2ª posição, inicialmente aparecem dois métodos empatados (M10 e M15) no critério erro relativo da assimetria, com valores de 27%, que ocupavam a última ou penúltima posição nos três casos anteriores, e com desempenhos inaceitáveis; no entanto, considerando que no desvio-padrão, principalmente, e também na média, os erros do M15 são elevados, o M10 passa a ser o segundo melhor método utilizando este critério adicional. Excetuando o método proposto (erro relativo = 0%), todos os outros 17 métodos, exibiram erros relativos superiores a 20%. E o número de métodos que produziram desempenhos não satisfatórios ou inaceitáveis, e que ocupam a segunda posição em outros casos analisados foi 6.

Tabela 5.5 – Resultados para o Caso NE\_PE – Agosto: parâmetros estimados da distribuição Weibull; estatísticas dos 20 valores sintéticos das velocidades mensais de ventos; erro relativo entre as assimetrias histórica e sintética das velocidades de ventos.

Parâmetros Weibull			Estatísticas - Valores Sintéticos			Erro Rel (%)	Método
Forma $\gamma$	Escala $\beta$	Posição $\alpha$	Média	D.P.	Assimetria		
4.767	-4.565	4.180	5.750	0.371	1.227	0	M1 - Algoritmo Proposto
0.555	0.310	-0.521	5.720	0.142	0.897	27	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
0.553	1.884	-1.306	6.173	0.865	0.899	27	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit
0.85	0.499	-0.543	5.739	0.134	0.456	63	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
0.962	0.946	-1.338	5.637	0.224	0.326	73	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)
0.962	1.315	-1.338	5.748	0.311	0.326	73	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)
0.978	0.571	-0.577	5.742	0.133	0.308	75	M7 - MoM (Cran, 1988) - Pacote R ForestFit
1.228	1.332	-1.246	5.750	0.256	0.061	95	M18 - Momentos T-L - Pacote R ForestFit
1.322	1.444	-1.332	5.750	0.262	-0.021	102	M11 - MLE (Green, 1994) - Pacote R ForestFit
1.335	1.577	-1.408	5.764	0.284	-0.032	103	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit
1.35	1.419	-1.294	5.752	0.254	-0.045	104	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
1.352	1.459	-1.338	5.750	0.261	-0.047	104	M5 - Recálculo do MoM com $\alpha$ dado por E[Min de uma amostra]
1.396	1.517	-0.609	5.974	0.265	-0.082	107	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
1.418	1.558	-1.417	5.751	0.269	-0.100	108	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
1.434	1.577	-1.432	5.751	0.270	-0.112	109	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
1.445	1.291	-1.423	5.675	0.220	-0.121	110	M2 - MoM (Cohen, 1982)
1.450	1.548	-1.401	5.751	0.263	-0.125	110	M6 - MLE Ajustado - Pacote R Weibullness
8.220	2.900	-1.271	6.148	0.161	-1.489	221	M17 - MLE Ponderado (Cousineau, 2009) - Pacote R ForestFit

Fonte: Autores

Dessas quatro análises efetuadas para o caso Caso NE\_PE – Agosto, com assimetria positiva e elevada, constata-se:

- em todos os casos, o algoritmo proposto teve um desempenho melhor do que os outros 17 métodos;
- quando a cardinalidade dos resíduos correlacionados normais diminui, o desempenho dos outros métodos também diminui;
- com exceção do algoritmo proposto, os métodos com melhores desempenhos variam de caso para caso;
- métodos que funcionam bem em um caso podem apresentar desempenho inaceitável em outros casos.

## 5.2 Caso Sul\_L – Julho ( $\hat{g} = -0,189$ )

O segundo caso estudado refere-se a um parque eólico situado no litoral da região sul, com valor negativo para a assimetria das velocidades médias mensais de ventos no mês de julho. A média, o desvio-padrão e a assimetria amostrais são, respectivamente, 6,207; 0,427 e -0.187.

Para cada um dos 18 métodos avaliados, as estimativas dos três parâmetros da distribuição Weibull, e também as estatísticas dos valores sintéticos das velocidades mensais de ventos considerando 2.000 resíduos correlacionados e normalmente distribuídos, estão expostos na Tabela 5.6.

Da mesma forma que nas quatro análises da Seção 5.1, o método proposto apresentou o melhor desempenho em relação às três estatísticas calculadas, e obteve o mesmo valor da assimetria histórica (erro relativo = 0%).

Já o segundo melhor método (M18) alcançou erro relativo igual a 19%. Com exceção do algoritmo proposto, os erros relativos de todos os demais 17 métodos foram novamente superiores a 20%; sendo que em 14 deles, o erro foi superior a 5%; em 6 deles a assimetria calculada foi positiva; e um (M8) não conseguiu estimar os parâmetros da distribuição Weibull. Neste caso, 7 métodos, incluindo o M8, produziram desempenhos não satisfatórios ou inaceitáveis, mas que ocupam a segunda posição em outros casos analisados.

Tabela 5.6 – Resultados para o Caso Sul\_L – Julho: parâmetros estimados da distribuição Weibull; estatísticas dos 2000 valores sintéticos das velocidades mensais de ventos; erro relativo entre as assimetrias histórica e sintética das velocidades de ventos.

Parâmetros Weibull			Estatísticas - Valores Sintéticos				
Forma $\gamma$	Escala $\beta$	Posição $\alpha$	Média	D.P.	Assimetria	Erro Rel (%)	Método
4.279	4.164	-3.789	6.215	0.431	-0.189	0	M1 - Algoritmo Proposto
0.364	3.932	-3.568	6.215	0.424	-0.153	19	M18 - Momentos T-L - Pacote R ForestFit
4.601	4.444	-4.062	6.214	0.433	-0.243	28	M11 - MLE (Green, 1994) - Pacote R ForestFit
2.118	4.446	-4.063	6.215	0.433	-0.244	29	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
6.605	5.624	-5.202	6.208	0.479	-0.351	86	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit
0.706	5.803	-5.392	6.213	0.444	-0.428	126	M6 - MLE Ajustado - Pacote R Weibullness
5.386	6.515	-4.028	7.090	0.465	-0.474	150	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
7.148	3.079	-2.749	6.215	0.431	0.122	164	M5 - Recálculo do MoM com $\alpha$ dado por E[Min de uma amostra]
4.082	1.889	0.452	7.158	0.126	-0.516	173	M17 - MLE Ponderado (Cousineau, 2009) - Pacote R ForestFit
14.789	10.879	-10.442	6.214	0.437	-0.757	300	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
1.201	11.643	-12.063	5.861	0.402	-0.810	328	M2 - MoM (Cohen, 1982)
1.201	2.109	-1.836	6.229	0.399	0.523	376	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
2.997	1.271	-2.749	5.552	0.432	1.562	925	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)
6.098	2.923	-2.749	6.227	0.992	1.562	925	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)
1.198	0.695	-0.654	6.212	0.236	1.567	928	M7 - MoM (Cran, 1988) - Pacote R ForestFit
0.706	4.600	-2.613	7.621	3.677	3.564	1984	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit
12.536	0.062	-0.274	6.213	0.454	9.008	4861	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
-	-	-	-	-	-	-	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit

Fonte: Autores



### 5.3 Caso NE\_I – Abril ( $\hat{g} = -0.320$ )

Neste terceiro caso, o parque eólico estudado está localizado no interior da região nordeste, e apresenta valor negativo para a assimetria das médias mensais das velocidades de ventos no mês de abril. As estatísticas média, desvio-padrão e assimetria amostrais são, respectivamente, 5,252; 0,542 e -0.320.

A Tabela 5.7 mostra, para o conjunto dos 18 métodos aplicados, as estimativas dos três parâmetros da distribuição Weibull, e também as estatísticas dos valores sintéticos das velocidades mensais de ventos considerando 2.000 resíduos correlacionados e normalmente distribuídos.

De maneira idêntica às cinco análises já realizadas (Seções 5.1 e 5.2), o algoritmo proposto apresentou o melhor desempenho em relação às três estatísticas calculadas, e obteve o mesmo valor da assimetria histórica (erro relativo = 0%).

Verifica-se também que o método M16 ocupa a segunda posição; que 4 métodos apresentaram erros relativos inferiores a 10%; e 11 acima de 20%, sendo que 7 deles apresentaram assimetrias positivas. Há ainda 5 métodos (M9, M10, M13, M14 e M15) que produziram desempenhos não satisfatórios ou inaceitáveis, e que ocupam a segunda posição em desempenho em outros casos analisados.

Tabela 5.7 – Resultados para o Caso NE\_I – Abril: parâmetros estimados da distribuição Weibull; estatísticas dos 2000 valores sintéticos das velocidades mensais de ventos; erro relativo entre as assimetrias histórica e sintética das velocidades de ventos.

Parâmetros Weibull			Estatísticas - Valores Sintéticos			Erro Rel (%)	Método
Forma $\gamma$	Escala $\beta$	Posição $\alpha$	Média	D.P.	Assimetria		
5.755	5.366	-4.966	5.257	0.535	-0.320	0	M1 - Algoritmo Proposto
5.731	5.754	-5.331	5.254	0.576	-0.318	1	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit
5.583	5.299	-4.897	5.257	0.542	-0.302	6	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
5.986	1.392	-0.208	5.840	0.134	-0.344	7	M17 - MLE Ponderado (Cousineau, 2009) - Pacote R ForestFit
5.264	4.879	-4.486	5.260	0.525	-0.264	18	M11 - MLE (Green, 1994) - Pacote R ForestFit
5.250	4.873	-4.487	5.257	0.526	-0.262	18	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
6.389	6.213	-5.783	5.257	0.565	-0.382	19	M18 - Momentos T-L - Pacote R ForestFit
4.766	4.454	-4.071	5.260	0.522	-0.197	38	M6 - MLE Ajustado - Pacote R Weibullness
4.759	4.408	-3.814	5.376	0.518	-0.196	39	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
4.056	3.274	-3.609	4.909	0.441	-0.078	76	M2 - MoM (Cohen, 1982)
3.870	3.876	-3.507	5.256	0.543	-0.040	87	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
2.577	2.602	-2.304	5.259	0.517	0.335	205	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
2.390	2.532	-2.245	5.255	0.538	0.417	230	M5 - Recálculo do MoM com $\alpha$ dado por E[Min de uma amostra]
1.223	1.299	-2.245	4.696	0.538	1.415	541	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)
1.223	2.398	-2.245	5.254	0.994	1.415	541	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)
1.218	0.704	-0.660	5.253	0.293	1.422	544	M7 - MoM (Cran, 1988) - Pacote R ForestFit
0.560	4.158	-2.128	7.815	6.858	4.259	1429	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit
0.263	0.008	-0.138	5.246	0.369	14.684	4682	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit

Fonte: Autores

### 5.4 Caso NE\_I – Junho ( $\hat{g} = -1.394$ )

Este quarto caso corresponde ao mesmo parque eólico da Seção 5.3, situado no interior da região nordeste, porém o mês analisado mudou de abril para junho. Em



consequência, a assimetria das velocidades médias mensais de ventos, apesar de continuar negativa, passou a apresentar módulo bastante elevado, i.e, variou de -0.320 para -1,394. A média e o desvio-padrão passaram a ser, respectivamente, 5,890; 0,468.

As estimativas dos três parâmetros da distribuição Weibull, assim como as estatísticas dos valores sintéticos das velocidades mensais de ventos considerando 2.000 resíduos correlacionados e normalmente distribuídos estão exibidos na Tabela 5.8, tanto para o método proposto quanto para os demais 17 métodos avaliados.

Observa-se que a ocorrência de assimetria amostral negativa e bastante elevada (-1,394) para as médias mensais das velocidades de ventos impactou o desempenho de todos os métodos: apenas 3 métodos (M1, M16 e M13) apresentaram erros relativos para a assimetria abaixo de 20% (erro relativo = 17%); dois outros (M5 e M12) ainda apresentaram assimetrias negativas, mas com erros relativos altos (89% e 97%); seis métodos apresentaram erros superiores a 100% e, ainda por cima, associados a assimetrias positivas, dos quais os métodos M14, M15 e M10 já haviam aparecido como segundo melhor método em alguma das análises efetuadas nas Seções 5.1 a 5.3; por fim, sete métodos (incluindo M8, M9 e M18) não conseguiram estimar os três parâmetros da distribuição Weibull e, por conseguinte, não foram capazes de gerar valores sintéticos de velocidades de vento.

Com relação aos três métodos que apresentaram os menores erros relativos, os quais foram idênticos e iguais a 17%, foi necessário lançar mão do critério adicional de desempenho, i.e., os erros relativos dos desvios-padrão sintéticos e amostrais. Os desvios-padrão oriundos dos métodos M1 (0,462), M16 (0,453) e M13 (0,533) correspondem respectivamente a erros relativos de 1,3%; 2,0% e 13,9%. Portanto, pode-se concluir que o método proposto (M1) apresentou desempenho ligeiramente superior ao M16, que ocupa a 2ª posição. Ressalta-se, todavia, que o método M16 apresentou desempenhos não satisfatórios em análises anteriores, com erros relativos de 54%, 86% e 104%.

Ainda chama a atenção os valores extremamente elevados obtidos para as estimativas do três parâmetros da distribuição Weibull, com o emprego dos três métodos que apresentaram os melhores desempenhos, especialmente os métodos M16 e M13. Conforme a Tabela 5.8, as estimativas do parâmetro de forma  $\gamma$  foram cerca de 10 mil os métodos proposto (M1); 2,4 milhões para M16; e 7,8 milhões para M13. Do ponto de vista conceitual, distribuições Weibull com assimetria negativa ocorrem para  $\gamma$  na faixa infinita  $[3,6 \dots \infty]$ ; entretanto, uma distribuição Weibull com  $\gamma = 10$  tem uma forma indistinguível de outra distribuição Weibull com  $\gamma = 10000$  (COUSINEAU, 2009a). Apesar das assimetrias obtidas com os métodos M1, M16 e M13 se tratarem de valores

extremamente elevados, parece ser razoável optar-se pelo menor valor, o que reforça a escolha da abordagem proposta (M1) como aquela de melhor desempenho.

A fim de evidenciar a boa aderência dos valores sintéticos produzidos pelo método proposto aos valores históricos, mesmo com valores extremamente elevados obtidos para as estimativas do três parâmetros da distribuição Weibull, a Figura 5.2 apresenta os histogramas para os 37 valores históricos de resíduos (a) e velocidades de vento (d), assim como os 2.000 valores sintéticos de resíduos (b e c) e velocidades de vento (e e f). No caso dos valores sintéticos, os histogramas foram construídos com e sem o mesmo número de classes dos histogramas históricos. De fato, há uma boa aderência entre os histogramas históricos e sintéticos.

Tabela 5.8 – Resultados para o Caso NE\_I – Junho: parâmetros estimados da distribuição Weibull; estatísticas dos 2000 valores sintéticos das velocidades mensais de ventos; erro relativo entre as assimetrias histórica e sintética das velocidades de ventos.

Parâmetros Weibull			Estatísticas - Valores Sintéticos				Método
Forma $\gamma$	Escala $\beta$	Posição $\alpha$	Média	D.P.	Assimetria	Erro Rel (%)	
10750	8383	-8382	5.883	0.462	-1.152	17	M1 - Algoritmo Proposto
2405924	1839979	-1839979	5.881	0.453	-1.152	17	M16 - Máximo Produto de Espaçamento - Pacote R ForestFit
7831992	7043811	-2208379	5.894	0.533	-1.152	17	M13 - MLE Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
4.189	4.090	-3.717	5.880	0.458	-0.152	89	M5 - Recálculo do MoM com $\alpha$ dado por E[Min de uma amostra]
3.683	3.106	-2.774	5.894	0.387	-0.040	97	M12 - MLE Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
3.419	2.896	-2.603	5.881	0.384	0.029	102	M14 - MLE Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
3.000	3.000	-2.462	5.980	0.444	0.161	112	M15 - MLE Modificado - Tipo 4 (Cohen, 1982) - Pacote R ForestFit
1.393	1.508	-3.717	4.779	0.452	1.321	195	M3 - MoM Modificado - Tipo 1 (Cousineau, 2009a)
1.393	4.075	-3.717	5.849	1.222	1.321	195	M4 - MoM Modificado - Tipo 2 (Cousineau, 2009)
1.376	0.769	-0.703	5.882	0.234	1.349	197	M7 - MoM (Cran, 1988) - Pacote R ForestFit
0.423	0.123	-0.354	5.878	0.486	9.955	814	M10 - MoM Modificado - Tipo 3 (Cohen, 1982) - Pacote R ForestFit
-	-	-	-	-	-	-	M2 - MoM (Cohen, 1982)
-	-	-	-	-	-	-	M6 - MLE Ajustado - Pacote R Weibullness
-	-	-	-	-	-	-	M8 - MoM Modificado - Tipo 1 (Cohen, 1982) - Pacote R ForestFit
-	-	-	-	-	-	-	M9 - MoM Modificado - Tipo 2 (Cohen, 1982) - Pacote R ForestFit
-	-	-	-	-	-	-	M11 - MLE (Green, 1994) - Pacote R ForestFit
-	-	-	-	-	-	-	M17 - MLE Ponderado (Cousineau, 2009) - Pacote R ForestFit
-	-	-	-	-	-	-	M18 - Momentos T-L - Pacote R ForestFit

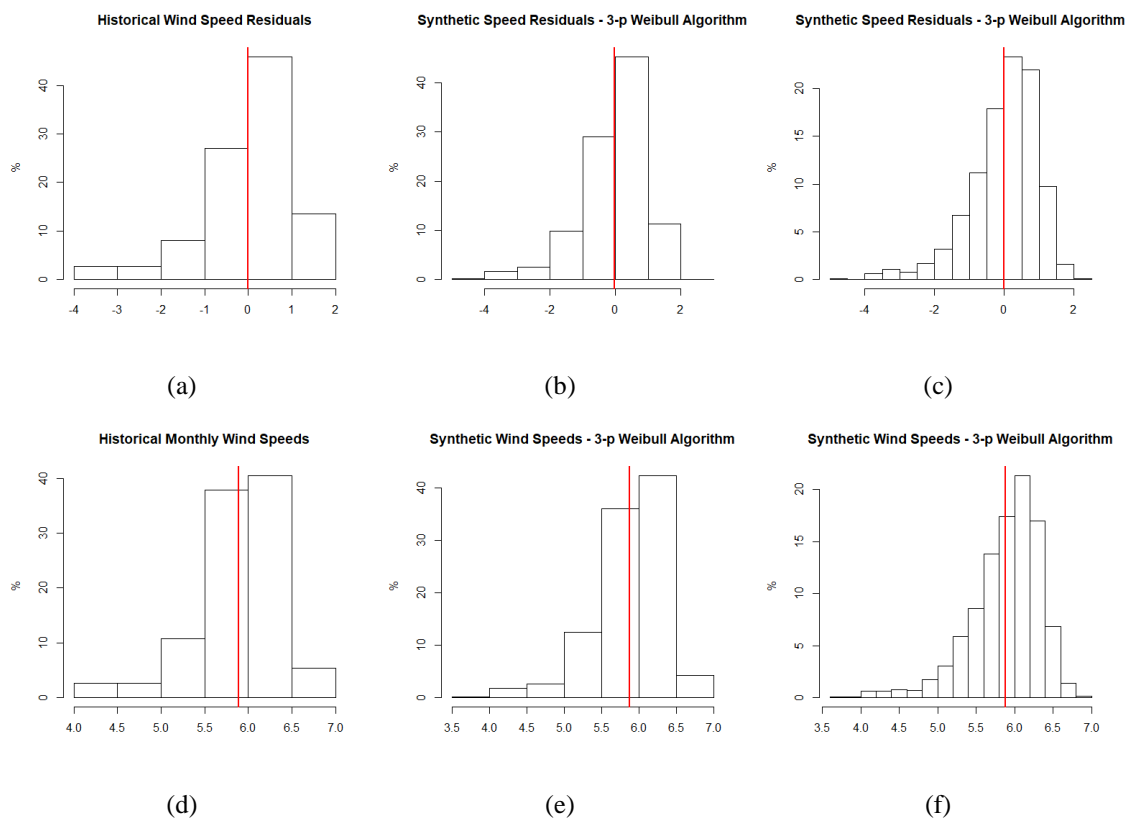
Fonte: Autores

Por fim, a Tabela 5.9 ilustra os valores dos três parâmetros da distribuição Weibull obtidos em cada iteração do algoritmo proposto. Observa-se que o algoritmo convergiu em apenas 3 iterações, a partir do valor inicial do parâmetro de posição ( $\hat{\alpha}_{inic}$ ) calculado pela equação de regressão apresentada na Tabela 4.1. Caso se considerasse como satisfatório o erro relativo de 18,1% , i.e., uma assimetria de -1.142, o algoritmo pararia na 2ª iteração, com parâmetro de forma estimado em  $\gamma = 492$ , ao invés  $\gamma = 10750$ .

Figura 5.2 – Histogramas para o Caso NE\_I – Junho ( $\hat{\gamma} = -1.394$ ): 37 valores amostrais de resíduos (a) e velocidades de vento (d); 2000 valores sintéticos (algoritmo proposto) de resíduos (b e c) e velocidades de vento (e e f).

## Valores Históricos

## Valores Sintéticos – Algoritmo Proposto



Fonte: Autores

Tabela 5.9 – Resultados para o Caso NE\_I – Junho ( $\hat{g} = -1.394$ ): processo iterativo do algoritmo proposto

Processo Iterativo - Algoritmo Proposto						
Forma $\gamma$	Escala $\beta$	Posição $\alpha$	Média	D.P.	Assimetria	Error (%)
428	335	-334	5.883	0.462	-1.140	18.2
492	385	-384	5.883	0.462	-1.142	18.1
10750	8383	-8382	5.883	0.462	-1.152	17.3

Fonte: Autores

## 6. Conclusões

Este artigo propôs uma abordagem para a modelagem de velocidades mensais de ventos por meio de distribuições Weibull tri-paramétricas, na geração de cenários de ventos mensais correlacionados com vazões às usinas hidroelétricas, em um esquema de Programação Dinâmica Dual Estocástica.

A velocidade dos ventos foi modelada como uma regressão das afluições mais um resíduo. Em consequência, a distribuição Weibull é ajustada aos resíduos das

velocidades mensais de ventos, correlacionados e normalmente distribuídos, que também considera as correlações cruzadas entre as afluências às hidroelétricas e as velocidades do vento dos parques eólicos. A abordagem proposta busca preservar a média, o desvio-padrão e, especialmente, a assimetria das velocidades históricas dos ventos mensais, sendo particularmente adequada em situações de assimetrias elevadas.

O algoritmo proposto parte de um valor inicial para a estimativa do parâmetro posição, o qual pode ser obtido através de regressões lineares, calcula estimativas dos outros parâmetros por meio do método dos momentos e, de forma iterativa, atualiza a estimativa inicial com vistas a reduzir a diferença entre as assimetrias das velocidades de ventos sintéticas (geradas) e amostrais (históricas).

A metodologia proposta foi aplicada a parques eólicos das regiões Nordeste e Sul do Brasil, considerando meses distintos e diversas assimetrias – positivas e negativas, sendo também realizada uma análise de sensibilidade com relação à cardinalidade da amostra dos ruídos correlacionados e normais. O seu desempenho foi comparado a 17 métodos disponíveis na literatura.

Em todos casos estudados, a abordagem proposta mostrou-se robusta e apresentou desempenho superior aos outros 17 métodos, considerando tanto assimetrias positivas como negativas, especialmente em situações de assimetrias negativas elevadas. Verificou-se que quando a cardinalidade da amostra de resíduos correlacionados e normalmente distribuídos 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.

Por fim, a abordagem proposta é geral, e pode ser utilizada em outras aplicações, considerando ou não resíduos correlacionados com as afluências às usinas hidrelétricas.

## Referências

ANDRADE, V.A., PESSANHA, J. F. M., MELO, A. C. G., MACEIRA, M.E.P.. Modelagem da Relação entre Velocidade de Vento e Produção Eólica no Modelo NEWAVE, **LIII Simpósio Brasileiro de Pesquisa Operacional**, João Pessoa, PB, Novembro 2021.

ARVANTIDIS, N.V., ROSING, J. Composite representation of multireservoir hydroelectric power system, **IEEE Transactions on Power Apparatus and Systems**, v. 89, n. 2, pp. 319-326, Feb. 1970.

CHARBENEAU, R.J. Comparison of the Two- and Three-Parameter Log Normal Distributions Used in Streamflow Synthesis, **Water Resources Research**, Vol. 14, N0. 1, February, 1978.

CHENG, R. C., AMIN, N.A K. Estimating parameters in continuous univariate distributions with a shifted origin”, **J. Royal Statistical Society B**, Vol. 45, pp. 394-403, 1983.

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.

COHEN, A.C., WHITTEN, B.J., DING, Y. Modified moment estimation for the three-parameter Weibull distribution. **Journal of Quality Technology** 16, 159–167, 1984.

COUSINEAU, D. Fitting the three-parameter Weibull distribution: review and evaluation of existing and new methods. **IEEE Transactions on Dielectrics Electrical Insulation** 16, 281–288, 2009.

COUSINEAU, D. Nearly unbiased estimators for the three-parameter Weibull distribution with greater efficiency than the iterative likelihood method. **British Journal of Mathematical and Statistical Psychology** 62, 167–191, 2009.

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

FRANSES, P.H.; PAAP, R. **Periodic Time Series Models**, Oxford University Press, 2006.

GREEN, E.J., ROESCH Jr, F.A., SMITH, A.F., STRAWDERMAN, W.E. Bayesian estimation for the three-parameter weibull distribution with tree diameter data. **Biometrics**, 254–269doi:10.2307/2533217, 1994.

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>.

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.. 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., 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., DUARTE, V. S., PENNA, D. D. J., MORAES, L. A. M., MELO, A. C. G.. Ten Years of Application of Stochastic Dual Dynamic Programming in Official and Agent Studies in Brazil - Description of the NEWAVE Program, **16th Power Systems Computation Conference**, Scotland, 2008.

MACEIRA, M.E.P., MELO, A.C.G., ZIMMERMANN, M.P. Application of stochastic programming and Probabilistic Analyses as Key Parameters for Real Decision Making regarding Implementing or Not Energy Rationing – A Case Study for the Brazilian Hydrothermal Interconnected System, **19th Power System Computation Conference - PSCC 2016**, Genoa, Italy, 2016.

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., MELO, A. C. G. Combined Representation of Hydropower Plants and Inflow Scenarios Re-Sampling on Stochastic Dual Dynamic Programming – Application to the Brazilian System. **15th International Conference on Stochastic Programming**, Trondheim, Norway, 2019a.

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, 2019b.

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.

MELO, A.C.G., ZIMMERMANN, M., MACEIRA, M.E.P., DAMAZIO, J.M. Challenges and Outlook for the Hydropower Development in Brazil, **International Journal of Hydropower&Dams**, London, 2020.

MME - MINISTÉRIO DE MINAS E ENERGIA/EPE - EMPRESA DE PESQUISA ENERGÉTICA, Plano Decenal de Expansão de Energia 2029. **Relatório Técnico**, Maio 2020.

NEWBY, M.J. Properties of moment estimators for the 3-parameter Weibull distribution. **IEEE Transactions on Reliability** 33, 192–195, 1984.

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

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.

PESSANHA, J. F. M.; MELO, A. C. G.; CALDAS, R. P.; FALCÃO, D. M. A Methodology for Joint Data Cleaning of Solar Photovoltaic Generation and Solar Irradiation, **16th IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS 2020)**, Belgium, August 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/>.

ROCKETTE, H., ANTLE, C., KLIMKO, L.A. Maximum likelihood estimation with the Weibull model, **J. American Statistical Association**, Vol. 69, pp. 246-249, 1974.

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

SMITH, R.L. Maximum likelihood estimation in a classe of nonregular cases, **Biometrika**, Vol. 72, pp. 67-90, 1985.

TERRY, L.A., PEREIRA, M.V.F., ARARIPE NETO, T.A., SILVA, L.F.C.A., SALES, P.R.H. Coordinating the Energy Generation of the Brazilian National Hydrothermal Electrical Generating System. **Interfaces**, 16: 1 January-February, pp.16-38, 1986.

WEIBULL, W. A Statistical Theory of Strength of Materials. In: **Ingeniors Vetenskaps Akademiens Handlingar**, vol. 151, 1939.

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. Wind speed is modeled as a regression of inflows to hydropower plants (from the current or previous period) plus a residual. Hence, the tri-parametric Weibull distribution is fitted to the residuals of the monthly wind speeds, correlated and normally distributed, which also consider the cross-correlations between inflows into the hydropower plants and wind speeds from wind farms. The proposed approach seeks to preserve the mean, the standard deviation and, especially, the skewness of the historical monthly wind speeds, being particularly suitable in situations of high asymmetries. With this purpose, the proposed algorithm starts from an initial value to estimate the position parameter, which can be obtained through linear regressions, calculates estimates of the other parameters through the method of moments and, iteratively, updates the initial estimate targeting to reduce the difference between the skewness of the synthetic (generated) wind speeds and the skewness of the sample (historical) wind speeds. The proposed methodology was applied to a set of wind farms located in the Northeast and South regions of Brazil, considering different months and distinct skewness - positive and negative - and proved to be robust, always performing better than the other 17 evaluated methods. Finally, the proposed approach is general, and can be used in other applications, considering or not normal residues correlated with inflows to hydroelectric plants.*

**Key-words:** Weibull distribution; Parameter Estimation; Methods of moments; Maximum likelihood; Wind generation Stochastic Optimization.