

CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ
ISSN on-line 2317-4535 / ISSN impresso 1413-9022 - v. 48, p.1 - 36, 2020
DOI: 10.12957/cadest.2020.55395

UMA ABORDAGEM PARA A REPRESENTAÇÃO DAS INCERTEZAS DA FONTE DE GERAÇÃO EÓLICA NO MODELO NEWAVE

Maria Elvira P. Maceira

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

Albert C.G. Melo

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

José F. M. Pessanha

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

Cristiane B. Cruz

Centro de Pesquisas de Energia Elétrica
criscruz@cepel.br

Victor A. Almeida

Centro de Pesquisas de Energia Elétrica
andrade@cepel.br

Thatiana C. Justino

Centro de Pesquisas de Energia Elétrica
thatiana@cepel.br

Resumo

Propõe-se uma abordagem para modelar a incerteza da produção eólica mensal no problema de despacho de operação, aderente ao algoritmo de solução utilizado (PDDE) no modelo NEWAVE. A abordagem compreende: (i) agrupamento estatístico dos regimes de ventos; (ii) avaliação de funções de transferência mensais entre ventos e produção eólica; (iii) geração de séries sintéticas de ventos correlacionadas às afluições aos reservatórios das hidroelétricas; e (iv) representação das produções eólicas mensais no algoritmo PDDE do NEWAVE. Resultados iniciais com configurações reais do sistema interligado brasileiro apontam para a sua efetividade.

Palavras-chave: *Planejamento da expansão e da operação; Geração eólica; Estatística multivariada; Séries temporais; Geração de séries sintéticas; Otimização estocástica.*

1. Introdução

O planejamento da expansão e da operação em sistemas hidrotérmicos tem sido realizado, em muitos países, desagregando os problemas de planejamento nos horizontes de longo, médio e curto prazos, de acordo com características específicas de seus sistemas de energia (HELSETH & MELO, 2020). No 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; MACEIRA *et al.*, 2016).

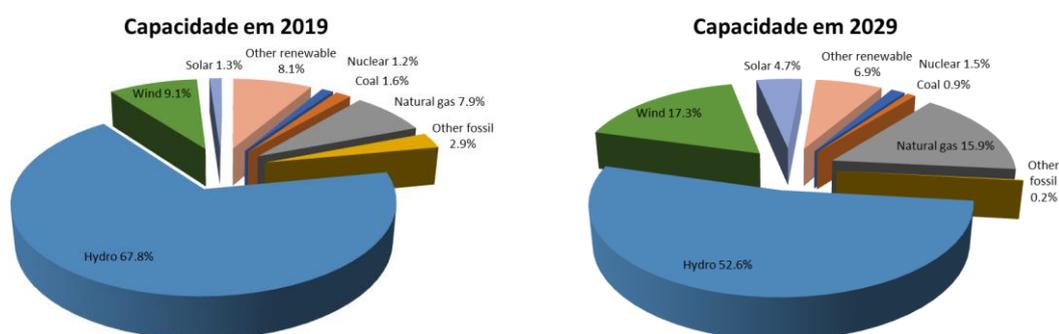
Um dos modelos-chave desta cadeia para as atividades de planejamento da expansão e da operação energética do sistema interligado é o modelo NEWAVE (MACEIRA *et al.*, 2008, 2018) que, desde 1998, vem sendo utilizado em estudos oficiais pelas entidades setoriais (Operador Nacional do Sistema – ONS, Câmara de Comercialização de Energia Elétrica – MAE / 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. O modelo NEWAVE representa o problema de despacho utilizado no planejamento de longo e médio prazo como um problema de programação linear estocástica multi-estágio, utilizando como estratégia de solução o algoritmo de programação dinâmica dual estocástica – PDDE (PEREIRA & PINTO, 1991, MACEIRA, 1993); a sua função-objetivo é minimizar o valor esperado do custo total de operação (custo do acionamento de usinas térmicas e de eventuais déficits de energia no futuro), considerando ainda mecanismos de aversão ao risco hidrológico.

No caso do Brasil, o sistema de geração apresenta uma alta participação de fontes renováveis de energia, incluindo as tecnologias hidroelétricas, biomassa, eólica e, mais recentemente, solar, respondendo por mais de 80% do consumo de eletricidade do país (MELO *et al.*, 2020). Apresenta ainda predominância da hidroeletricidade, caracterizada por grandes reservatórios com capacidade de regularização plurianual, dispostos em cascatas complexas.

Com relação à fonte eólica, ela tem experimentado um crescimento acelerado mundialmente e também no Brasil. Em maio de 2019, esta fonte já contava com mais de 14.968 MW instalados em mais de 600 parques, localizados principalmente nas regiões Nordeste (que concentra 85% da capacidade instalada) e Sul, correspondendo a 9,1% de

toda a capacidade instalada no País. De acordo com o Plano Decenal de Expansão de Energia 2020-2029 (MME e EPE, 2020), 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, conforme apresentado na Figura 1.1. A relevância da fonte eólica também pode ser constatada em termos regionais: a energia eólica representa 35,6% da capacidade instalada da região Nordeste e em 2018 foi responsável por 51,6% do suprimento de energia desta região.

Figura 1.1 – Capacidade Instalada no Brasil - Anos 2019 e 2029



Fonte: Autores, a partir de (MME e EPE, 2020)

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, no planejamento da operação de longo, médio e curto prazos, como é o caso da geração eólica.

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.

Com o intuito de aprimorar a representação atual, encontra-se em desenvolvimento uma metodologia para a representação das incertezas associadas às fontes renováveis intermitentes (eólica e solar) inicialmente no modelo NEWAVE, aderentes às características intrínsecas das modelagens adotadas nesse modelo, com o objetivo de manter o problema estocástico e de grande porte resolvido pelo NEWAVE tratável computacionalmente. Assim, o desafio é definir uma estratégia de consideração de cenários de produção eólica e solar que seja aderente ao algoritmo de solução utilizado, PDDE, e que permita ainda a obtenção de resultados com a precisão adequada e dentro de um tempo computacional razoável. Essa mesma metodologia poderá ser aplicada ao modelo DECOMP (DINIZ *et al.*, 2018) a partir do segundo mês do seu horizonte. Dessa forma, daqui em diante se fará referência apenas ao modelo NEWAVE.

A metodologia está sendo desenvolvida e implementada em estágios, iniciando com a incorporação da incerteza da fonte eólica. Neste caso, devido a não disponibilização no Brasil de valores verificados (medidos) de produção eólica e velocidade dos ventos, em bases horárias, torna-se necessário a utilização de dados de reanálises.

O objetivo deste trabalho é descrever uma proposta metodológica para a representação das incertezas da fonte de geração eólica no modelo NEWAVE. A metodologia proposta é multidisciplinar, empregando análise exploratória de dados, métodos estatísticos multivariados, processos estocásticos e otimização estocástica. Resultados iniciais da aplicação das etapas da abordagem proposta a configurações reais do sistema interligado brasileiro também são apresentados e discutidos.

2. O Modelo de Planejamento da Operação de Longo e Médio Prazos

No modelo NEWAVE (MACEIRA *et al.*, 2008, 2018), o problema de planejamento da operação é representado como um problema de programação linear estocástico multi-estágio. Sua 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). Custos de combustível das usinas termoelétricas 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 o modelo NEWAVE 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", por exemplo, as últimas p energias naturais afluentes de cada reservatório, individual ou equivalente. As incertezas nas afluições - vazões quando a representação do sistema se dá a UHEs ou energias afluentes quando a representação do sistema se dá a REEs - são modeladas por um modelo auto-regressivo periódico da ordem p , PAR(p) (SALAS *et al.*, 1980; FRANCES & PAAP, 2006) empregado no modelo GEVAZP (MACEIRA & BEZERRA, 1997; PENNA *et al.*, 2011), a fim de considerar as correlações temporais das afluições, e um esquema de simulação de Monte Carlo é usado para construir iterativamente as funções de custo futuro em cada estágio. O conjunto de funções de custo futuro representa a política ótima de operação.

Para resolver o problema de programação linear estocástico multi-estágio, o NEWAVE utiliza o algoritmo de Programação Dinâmica Dual Estocástica (PDDE) (PEREIRA & PINTO, 1991, MACEIRA, 1993). 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).

Além de fornecer a política de operação por meio da função de custo futuro para o modelo de programação de curto prazo – DECOMP (DINIZ *et al.*, 2018), o modelo NEWAVE é empregado nas seguintes atividades setoriais: Plano Decenal de Expansão de Energia (PDE) – MME/EPE; Programa de Operação Energética – ONS; Cálculo do Preço de Liquidação das Diferenças – CCEE; Cálculo da Garantia Física das UHEs –

MME/EPE; definição de parâmetros dos leilões de compra de energia – MME/EPE; e estratégias corporativas dos diversos agentes do setor elétrico brasileiro.

A formulação compacta do problema de planejamento da operação de médio/longo prazo representada no modelo NEWAVE, em sua forma recursiva, está apresentada em (1).

As variáveis de decisão x_t são associadas aos níveis de reservatórios x_t^v e à alocação dos recursos hídricos x_t^{gh} e térmicos x_t^{gt} . A incerteza das afluições aos reservatórios é representada pelo vetor ξ_t .

$$\min_{x_1} c_1 x_1 + E_{\xi_2} [\vartheta_2(x_1, \xi_2)]$$

$$s. a. g_1(x_1) = b_1$$

$$x_1 \in X$$

(1a)

$$\vartheta_t(x_{t-1}, \xi_t) = \min_{x_t} c_t x_t + E_{\xi_{t+1} | \xi_t, \dots, \xi_{t+1-p}} [\vartheta_{t+1}(x_t, \xi_{t+1})]$$

$$s. a. g_t(x_t) = b_t(x_{t-1}, \xi_{t-j}, j=1, \dots, p)$$

$$x_t \in X$$

$$t = 2, \dots, T$$

(1b)

2.1 Estratégia de Solução

Dado que não é possível percorrer a árvore completa de cenários de afluições, o algoritmo de PDDE – que é uma versão amostral da decomposição de Benders aninhada (BIRGE, 1985) – é aplicado para resolver o problema. A estratégia de solução consiste em percorrer uma sub-árvore de cenários de afluições, que é escolhida da distribuição original da variável aleatória, iterativamente através de duas simulações:

Simulação forward, do estágio $t=1$ até $t=T$ percorrendo toda a sub-árvore (totalizando $nsim$ cenários), com a finalidade de gerar novos estados $\hat{x}^t := \{\hat{x}_t^v, \hat{\xi}_{t-p}, p = 1, \dots, P\}$ para os quais a função de custo futuro será avaliada e novos cortes de Benders construídos na próxima recursão *backward*.

Recursão backward, do estágio $t=T$ até $t=1$, 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*. As variáveis duais associadas a esses subproblemas de programação linear são utilizadas para construir um corte de Benders (conjunto de restrições lineares multivariadas que representam a função de custo futuro), produzindo uma limite inferior do custo total esperado de operação em todo o horizonte de estudo (ZINF) no valor do estado associado ao nó da sub-árvore em questão $\hat{x}_{t-1,s}$.

As modelagens do subproblema de cada nó da sub-árvore (t,s) do período t (LSP), do cenário *forward* s e do cenário *backward* ω é apresentado na equação (2). Por simplificação, apresenta-se a formulação relativa a REEs sem a consideração de mecanismos de aversão a risco. Maiores detalhes são apresentados em (MACEIRA *et al.*, 2018).

A função objetivo (2a) do LSP compreende os custos de geração térmica, déficit de energia e eventuais violações de restrições operativas (vazão mínima, desvio de água, geração hidráulica mínima, e etc), onde β é a taxa de desconto e α representa a função de custo futuro. As principais restrições em cada período são o atendimento a demanda de energia em cada patamar de carga e submercado (2b) e as equações do balanço hídrico nos reservatórios (2c) para todos os k subsistemas que compõem o sistema. Compõem também o conjunto de restrições, aquelas que representam a função de custo futuro (conjunto de cortes de Benders) (2d), construídas a partir das variáveis duais associadas aos sub-problemas de programação linear, onde p_ω é a probabilidade condicional do cenário *backward* ω , $z^{t,s,\omega*}$ é o valor ótimo e $\pi^{t,s,\omega*}(\cdot) = \partial z^{t,s,\omega*} / \partial x_{t-1}(\cdot)$ é o vetor de derivadas parciais da função objetivo em relação ao estado do sistema no cenário *forward* s .

$$\min \vartheta_{t,s,\omega} = \vartheta_t(x_{t,s,\omega}, \xi_{t,\omega}) = c_t x_{t,s,\omega}^{gt} + \left(\frac{1}{1 + \beta} \right) \alpha^t (x_{t,s,\omega}^v, \xi_{t-p,\omega}, p = 0, \dots, p_m - 1) \quad (2a)$$

$$s. a. x_{t,s,\omega}^{gt} + x_{t,s,\omega}^{gh} = d_t \quad (2b)$$

$$x_{t,s,\omega}^v - x_{t-1,s,\omega}^v + x_{t,s,\omega}^q = \xi_{t,\omega} \quad (2c)$$

$$\vartheta_{t-1}(x_{t-1}) \geq \sum_{\omega=1, \dots, nleq} p_\omega [z^{t,s,\omega*} + \langle \pi^{t,s,\omega*}(\hat{x}_{t-1,s}), x_{t-1} - \hat{x}_{t-1,s} \rangle] \quad (2d)$$

$$\underline{x}_t \leq x_{t,s,\omega} \leq \bar{x}_t \quad (2e)$$

Outras restrições que são representadas na modelagem do problema LSP e que não aparecem em (2), para efeito de simplificação, são: geração hidráulica mínima e máxima para cada UHE/REE e por patamar de carga; perdas a fio d'água para cada reservatório; limites de armazenamento em cada reservatório; armazenamento mínimo operativo em cada reservatório; retirada/acréscimo de água para outros usos em cada reservatório; deplecionamento mínimo para cada reservatório; limite máximo de geração de cada usina térmica; limite mínimo de intercâmbio entre os subsistemas por patamar de carga; fluxo nas interligações de um subsistema que não tem carga e nem geração associada; agrupamento de intercâmbio em cada patamar de carga; geração térmica

antecipada em cada patamar de carga; função de produção hidroelétrica por usina hidroelétrica e patamar de carga; turbinamento máximo para cada usina hidroelétrica e patamar de carga. Em particular, na representação a reservatório equivalente de energia duas equações são utilizadas para representar o balanço hídrico: controlável e a fio d'água.

3. Descrição Geral da Abordagem Proposta

De forma ampla, a pesquisa em curso tem como objetivo a definição de uma estratégia de consideração de cenários de produção de fontes renováveis intermitentes, como a eólica e a solar, que seja aderente ao algoritmo de solução utilizado (PDDE) no modelo NEWAVE. Como corolário, deve permitir a obtenção de resultados com a precisão adequada e dentro de um tempo computacional razoável, quando aplicada a sistemas interligados de grande porte, com predominância hidroelétrica, como é o caso do sistema brasileiro.

A metodologia está sendo desenvolvida e implementada em estágios, sendo que no estágio inicial será incorporada a incerteza da fonte eólica. Desta forma, o objetivo básico deste trabalho é descrever a abordagem proposta para a modelagem da incerteza da produção eólica *mensal* no problema de despacho de operação, resolvido pelo algoritmo de PDDE do modelo NEWAVE, através da modelagem da incerteza do processo estocástico dos ventos ou diretamente pela modelagem da incerteza da produção eólica.

Dadas a quantidade de parques eólicos atualmente instalados no Brasil (cerca de 640) e as perspectivas de elevado aumento neste número, torna-se fundamental investigar a forma de representação dos parques eólicos no modelo NEWAVE, à semelhança do que já ocorre com a representação das usinas hidroelétricas por UHEs ou REEs. Neste sentido, os parques eólicos poderão ser representados de forma individual (Centrais Geradoras Eólicas - EOLs) ou por meio de parques eólicos equivalentes (PEEs).

Para a obtenção dos PEEs, faz-se necessário efetuar a identificação e o agrupamento dos regimes de ventos por meio de técnicas estatísticas, considerando diversas granularidades espaciais, e.g., parques eólicos individualizados (EOLs), subestações agregadoras dos parques eólicos conectados ao SIN (46 subestações no sistema Sintegre do ONS, sendo 34 na região nordeste e 12 na região sul) e regiões/sub-

regiões geográficas do Brasil. Adicionalmente, a granularidade temporal, i.e., dados em base horária ou mensal, também deve ser estudada.

Adicionalmente, serão modeladas as correlações espaciais entre os parques eólicos (EOLs/PEEs) e também as correlações espaciais entre EOLs/PEEs e UHEs/REEs.

Em um primeiro momento, a estrutura de correlação temporal porventura verificada no processo estocástico dos ventos *médios mensais* (m/s) em algum EOL ou PEE não será considerada de forma explícita no modelo de geração de séries sintéticas de ventos, podendo ser representada de forma indireta, por meio da correlação espacial verificada entre os processos estocásticos dos ventos nos EOLs/PEEs e afluições às UHEs/REEs. Como consequência, nenhuma variável de estado será adicionada ao algoritmo de PDDE do modelo NEWAVE, não havendo, portanto, acréscimo da cardinalidade da função de custo futuro (multivariada), que representa a política ótima de operação.

Tanto para a simulação *forward*, quanto para a recursão *backward* e também para a simulação da operação do sistema interligado com a política de operação obtida, há, no módulo GEVAZP do modelo NEWAVE, uma estratégia de geração de séries sintéticas multivariadas de afluições aos reservatórios das UHEs/REEs, utilizando modelos autorregressivos periódicos - PAR(p). A proposta deste trabalho é estender este módulo para torná-lo um *modelo integrado de geração de séries sintéticas multivariadas mensais de afluições e ventos*, observando os aspectos listados anteriormente.

Uma vez obtidos os cenários sintéticos de ventos mensais para os EOLs ou PEEs, é necessário estimar a produção eólica associada para ser considerada no problema de despacho mensal do modelo NEWAVE. A abordagem proposta consiste em construir *funções de transferência mensais* (FTMs) entre vento e produção eólica mensais, a partir do levantamento das *curvas de potência probabilísticas mensais* (produção versus velocidade média mensal) dos EOLs/PEEs. Entre as técnicas que poderiam ser empregadas para este fim, incluem-se: regressões lineares e modelos aditivos generalizados. Ainda, em havendo dados, seria relevante que as FTMs pudessem ser avaliadas sazonalmente e atualizadas anualmente.

Uma questão importante no desenvolvimento e testes da metodologia proposta, 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 (EPE e CCEE) têm declarado que, por questões de confidencialidade com os proprietários dos dados (os parques eólicos), a sua disponibilização não é possível.

Com relação aos dados públicos no Brasil, a EPE disponibiliza algumas informações sobre a velocidade de vento no sistema de Acompanhamento de Medições Eólicas (AMA) agregados por unidades da Federação; embora interessantes, os dados disponibilizados não apresentam resoluções espaciais e temporais necessárias para a realização de análises mais sofisticadas e abrangentes (PESSANHA, ALMEIDA & MELO, 2020a; PESSANHA *et al.*, 2007). Com relação aos dados de produção eólica, a partir de 2018, a CCEE passou a disponibilizar dados verificados em base horária; porém o histórico disponibilizado, no momento, ainda é insuficiente. Por sua vez, o ONS disponibiliza desde 2018 valores previstos de velocidades de ventos e produção eólicas, em base semi-horária, para o conjunto das 46 subestações que fazem parte do seu sistema Sintege; novamente o histórico disponibilizado ainda é insuficiente, além de se tratar de valores não observados.

Deste modo, tornou-se necessário a realização de avaliações adicionais, incluindo dados oriundos de reanálises, por exemplo, os dados oriundos do MERRA-2 (global), disponíveis para qualquer local do planeta e com resolução horária (STAFFELL & PFENNINGER, 2016).

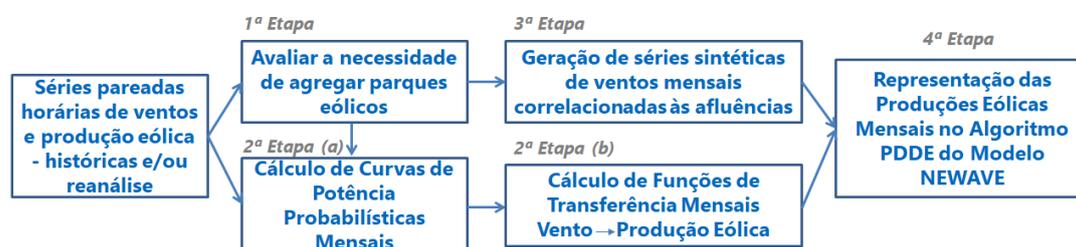
Em resumo, a abordagem proposta e descrita neste trabalho é multidisciplinar, conjugando o emprego de diversas técnicas, dentre as quais: análise exploratória de dados - análise estatística de agrupamentos e métodos estatísticos multivariados (JOHNSON e WICHERN, 1998); métodos estatísticos para análise de regressão, desde os tradicionais modelos de regresso linear até abordagens mais flexíveis, como modelos aditivos generalizados (generalized additive models - GAM) (WOOD, 2006); processos estocásticos e séries temporais (BOX & JENKINS, 1976); SALAS *et al.*, 1980; FRANCES & PAAP, 2006); otimização estocástica - programação dinâmica dual estocástica (PEREIRA & PINTO, 1991, MACEIRA, 1993).

Por fim, a abordagem proposta, conforme ilustrado na Figura 3.1, é composta por quatro etapas interligadas, com desenvolvimentos simultâneos: (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 afluências aos reservatórios das hidroelétricas; e (iv) representação das produções eólicas mensais no algoritmo PDDE do modelo NEWAVE.

Uma descrição sucinta de cada uma dessas etapas, assim como resultados já obtidos com a aplicação da abordagem proposta a configurações reais do sistema interligado serão apresentados e discutidos nas próximas seções.

Figura 3.1. Representação das incertezas da fonte eólica no modelo NEWAVE – diagrama esquemático das etapas da metodologia proposta



Fonte: Autores

4. Agrupamento Estatístico dos Regimes de Ventos – Etapa 1

Conforme apresentado na Seção 3, a Etapa 1 da abordagem metodológica proposta consiste na identificação e no agrupamento dos regimes de ventos nas diversas regiões do Brasil em PEEs.

A agregação de regimes de vento em PEEs proposta neste trabalho é baseada em métodos estatísticos multivariados e compreende duas etapas sequenciais: (i) a técnica de Análise Fatorial Exploratória (JOHNSON e WICHERN, 1998) é aplicada à matriz de covariâncias entre os regimes de ventos nos PEEs reduzindo a dimensionalidade dos dados; e (ii) a identificação de grupos de parques eólicos com regimes de vento correlacionados é realizada por meio de análise estatística de agrupamentos, e.g., utilizando o método K-Means ou algoritmo Ward aplicados às coordenadas dos PEEs nos eixos fatoriais (JOHNSON e WICHERN, 1998).

Para isso, deve-se formar uma matriz de dados na qual cada coluna guarda a série temporal de velocidade de vento em um aproveitamento eólico. Na sequência, a identificação de grupos de regimes eólicos semelhantes envolve a avaliação da similaridade das séries temporais de velocidade de ventos nos aproveitamentos. Neste

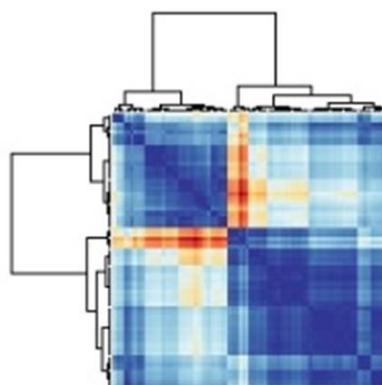
caso, deve-se analisar as correlações entre as séries de velocidade de vento, i.e., entre as colunas da matriz de dados.

Por exemplo, dado que na Seção 3 foi justificada a necessidade de se utilizar dados oriundos de reanálises, a Figura 4.1 apresenta um mapa de calor (*heatmap*) da matriz de correlações entre as séries das médias mensais de velocidade de vento a 100 m de altura, oriundas do MERRA-2 (global) para o período de 2001-2017 em 498 parques eólicos que totalizam uma capacidade instalada de cerca de 12.676 MW, agregados em 79 municípios de 8 estados da região Nordeste. Uma análise pictórica do mapa revela a existência de 2 até 5 regimes eólicos.

Embora a matriz de correlações na Figura 4.1 identifique a existência de grupos de parques eólicos espacialmente correlacionados, eles não estão claramente delimitados. A discriminação dos grupos pode ser alcançada, por exemplo, por meio da aplicação da análise fatorial exploratória - AFE, seguida da análise de agrupamentos.

Seja X a matriz de dados de reanálise, em que cada coluna guarda a série temporal da velocidade de vento em uma localidade com aproveitamentos eólicos. Assim, para o caso de n localidades, cada uma com uma série temporal com q registros horários de velocidade vento, a matriz X tem dimensões $n \times q$. A partir da matriz X pode-se obter a matriz de covariâncias S entre as velocidades de vento dos n parques eólicos. A matriz S tem dimensões $n \times n$ e cada elemento S_{ij} contém a covariância entre as velocidades de vento nas localidades i e j .

Figura 4.1. Mapa de calor da matriz de correlações entre as médias mensais de 79 localidades com parques eólicos na região Nordeste



Fonte: Autores

A AFE é uma técnica para redução da dimensionalidade dos dados na qual admite-se que a velocidade do vento x_i em cada parque eólico i seja expressa em função da soma

dos efeitos de m ($m < n$) regimes eólicos (fatores latentes $F_j \forall j=1,m$) mais uma componente específica ε_i da i -ésima localidade com aproveitamento eólico:

$$x_i = l_{i1}F_1 + l_{i2}F_2 + \dots + l_{im}F_m + \varepsilon_i \quad \forall i=1,n \quad (3)$$

em que l_{ij} é o peso do i -ésimo parque eólico no j -ésimo fator latente.

Os pesos de cada parque nos fatores são determinados com base nos autovalores e autovetores da matriz de covariâncias entre as velocidades de vento tal que $S \cong L^T L$, onde L é a matriz dos pesos dos parques nos fatores (JOHNSON e WICHERN, 1998).

A aplicação da AFE na matriz de covariâncias entre as 79 localidades indica que cerca de 95% da variância total presente nos dados está concentrada nos 3 primeiros fatores latentes. Portanto, nesta etapa foram adotados 3 fatores latentes, permitindo ainda obter uma representação gráfica em 3 dimensões da estrutura de correlações entre os regimes de vento que atingem as localidades, conforme apresentada na Figura 4.2(a). Cada dimensão corresponde a um eixo fatorial e pode ser interpretada como sendo um regime de ventos. Cada ponto no diagrama ilustrado na Figura 4.2(a) representa um município com aproveitamento eólico. As distâncias entre os pontos na Figura 4.2(a) refletem as correlações entre os regimes de vento, sendo que pontos próximos indicam regimes correlacionados.

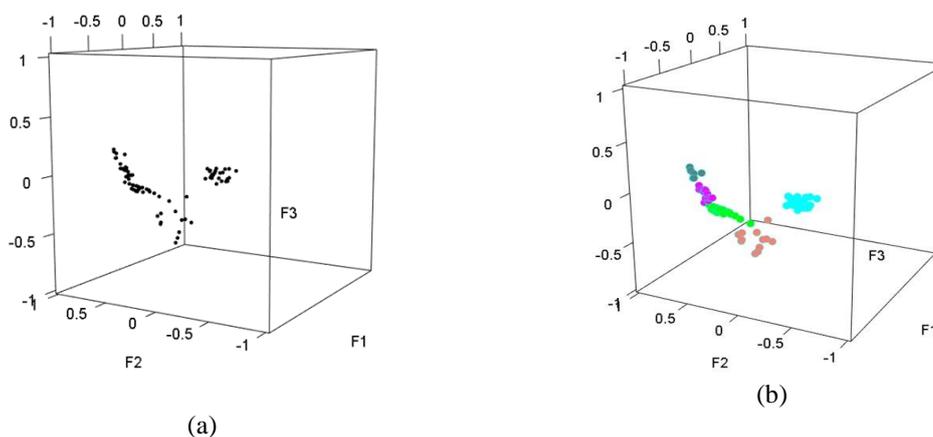
Por sua vez, a identificação de clusters de parques eólicos com regimes de vento correlacionados pode ser realizada por meio de análise de agrupamentos - *cluster analysis* (JOHNSON e WICHERN, 1998; MELO *et al.*, 2018a; MELO *et al.*, 2018b; PESSANHA *et al.*, 2018) (JOHNSON e WICHERN, 1998).

A análise de agrupamentos é uma técnica estatística útil na segmentação de um conjunto de n objetos (localidades com parques eólicos), caracterizados por m atributos, em subconjuntos mutuamente exclusivos, denominados clusters, de tal forma que os objetos em um mesmo subconjunto sejam semelhantes entre si, mas ao mesmo tempo diferentes dos objetos pertencentes aos outros subconjuntos. Entre os principais métodos de análise de agrupamentos destaca-se o *K-Means*, aqui empregado.

Para a aplicação do *K-Means*, os vetores com as coordenadas dos aproveitamentos eólicos nos eixos fatoriais foram empilhados em uma nova matriz de dados, neste caso uma matriz com três colunas, já que 3 fatores latentes foram selecionados na etapa da AFE. Neste caso, a partição em 5 agrupamentos capturou cerca de 90% da variabilidade dos dados, medida pela razão entre a inércia entre os *clusters* e a inércia total dos dados.

A Figura 4.2(b) ilustra o diagrama das localidades avaliadas nos três eixos fatoriais apresentado na Figura 4.2(a), incluindo agora a classificação dos aproveitamentos eólicos nos 5 agrupamentos identificados.

Figura 4.2 – Diagrama das localidades avaliadas nos três eixos fatoriais (a) e agrupamentos dos aproveitamentos eólicos (b)

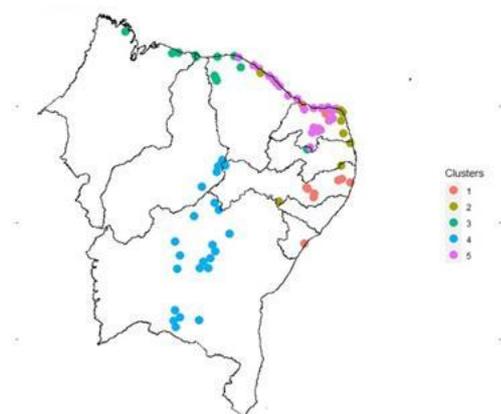


Fonte: Autores

Já na Figura 4.3 tem-se a representação espacial dos agrupamentos identificados na Figura 4.2. Vale destacar a reduzida superposição entre os agrupamentos, conseguida por meio da utilização da AFE.

Uma vez identificados os agrupamentos dos regimes de ventos, a abordagem descrita na Seção 5 será aplicada aos dados agrupados, obtendo-se, por fim, as funções de transferência entre médias mensais (FTMs) da velocidade de vento e produção eólica para os grupos eólicos identificados.

Figura 4.3 – Representação espacial dos agrupamentos dos aproveitamentos eólicos



Fonte: Autores

Como desenvolvimento futuro, será investigada a granularidade espacial mais adequada para utilização no processo de geração de séries sintéticas de velocidades de

ventos mensais, dentro do algoritmo PDDE do modelo NEWAVE. Também serão avaliadas as vantagens/limitações de se utilizar, na identificação de grupos de regimes eólicos semelhantes, matrizes de correlações entre as séries dos valores horários e das médias mensais de velocidade de vento.

5. Avaliação das Funções de Transferência Mensais (FTMs) entre Ventos e Produção Eólica – Etapa 2

Na metodologia proposta, após a geração de séries sintéticas de ventos mensais (Etapa 3), é preciso obter a correspondente produção eólica de cada parque eólico para ser considerada no problema de despacho mensal da operação do modelo NEWAVE (Etapa 4). Para isso há a necessidade de ajustar funções de transferência (FTMs) capazes de descrever o relacionamento entre as médias mensais de velocidade de vento e de produção eólica mensal.

Por sua vez, as FTMs entre ventos e produção eólica são obtidas a partir de curvas de potência probabilísticas *mensais* (CPPMs). As abordagens para a construção das CPPMs e FTMs são descritas a seguir.

5.1 - Cálculo de Curvas de Potência Probabilísticas Mensais – Etapa 2(a)

O cômputo de curvas de potência probabilísticas *mensais* (CPPMs) requer a análise de dados pareados de velocidade de vento e de produção eólica, um requisito difícil de ser alcançado, sobretudo quando há muitos aproveitamentos eólicos, e também devido a não disponibilização de uma base pública de dados verificados (medidos).

Uma alternativa para contornar a dificuldade na obtenção de dados verificados e pareados, é lançar mão, de previsões de velocidade de vento, acompanhados das respectivas previsões de produção eólica, que recentemente (meados de 2018) passaram a ser disponibilizados diariamente pelo sistema Sintegre do ONS, até 168 horas à frente e em base semi-horária. Adicionalmente, no Sintegre são disponibilizadas as funções de transferência usadas para converter as previsões de velocidade em previsões de produção eólica, uma função ajustada para cada semi-hora do dia.

Apesar do período de disponibilização ser relativamente curto, foi elaborado no âmbito deste trabalho um procedimento que permite utilizar este conjunto de funções de transferência de cada intervalo semi-horário para transformar as reanálises das velocidades horárias de vento em estimativas horárias de produção eólica.

O procedimento desenvolvido, ilustrado na Figura 5.1, consiste na aplicação das funções disponibilizadas no Sintegre às séries temporais de reanálise de velocidades de vento horárias, obtidas do MERRA 2, cobrindo o período de 1980 a 2019 (40 anos). Com isto são obtidas estimativas de produção de eólica horárias para o mesmo período, assim como curvas de potência probabilísticas *horárias*. Em seguida, as estimativas de produção de eólica horárias são agrupadas por mês para a obtenção de séries temporais das médias mensais da velocidade de vento e produção eólica. A partir destes dados, são obtidas as curvas de potência probabilísticas mensais (diagramas de dispersão).

Chama-se a atenção que os dados no sistema Sintegre são fornecidos para grupos de aproveitamentos eólicos, cada um associado à uma subestação do sistema elétrico, a qual pode ser visualizada como um PEE. A desagregação por parque eólico pode ser obtida de forma aproximada mediante fatores de participação, também fornecidos no Sintegre, e que permitem quantificar a contribuição de cada parque eólico na subestação em que está conectado.

Figura 5.1 – Procedimento para a obtenção de Curvas de Potência Probabilísticas Mensais utilizando dados do SINTEGRE e MERRA 2 – Diagrama Esquemático



Fonte: Autores

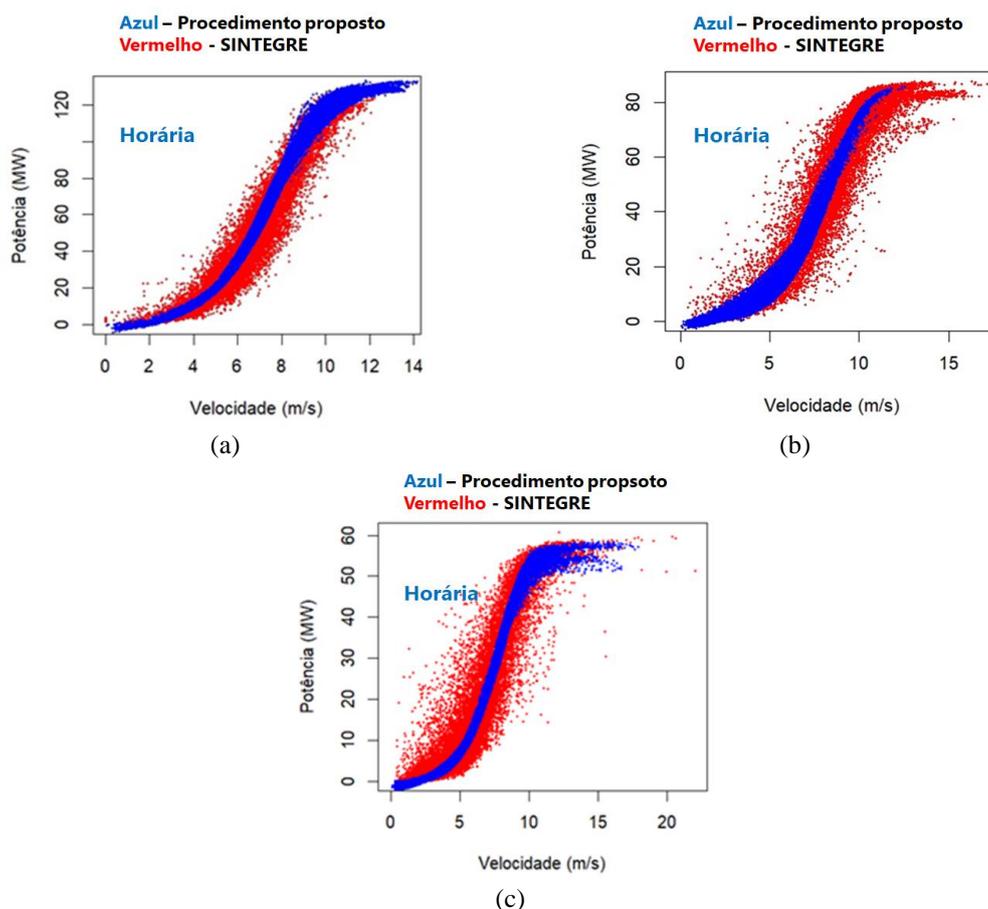
Para ilustrar a factibilidade e o desempenho do procedimento proposto, foram analisados dados dos PEEs de três subestações disponibilizados no Sintegre: duas para a região Nordeste – PEE NE-L, no litoral, e PEE NE-I, no interior; e uma para a região Sul - PEE Sul.

Para cada uma das subestações consideradas, a Figura 5.2 apresenta (na cor vermelha) as curvas de potência probabilísticas horárias (diagramas de dispersão) obtidas com os dados do Sintegre. Nesta mesma Figura também são apresentadas (na cor azul) as curvas de potência probabilísticas horárias obtidas com o procedimento desenvolvido, i.e., pela aplicação das funções disponibilizadas no Sintegre às séries temporais de

reanálise de velocidade de vento, obtidas do MERRA 2. Verifica-se que as curvas de potência probabilísticas horárias obtidas com o procedimento desenvolvido estão contidas nos diagramas de dispersão oriundos do Sintegre, evidenciando a razoabilidade tanto do procedimento elaborado quanto da utilização de dados de reanálise.

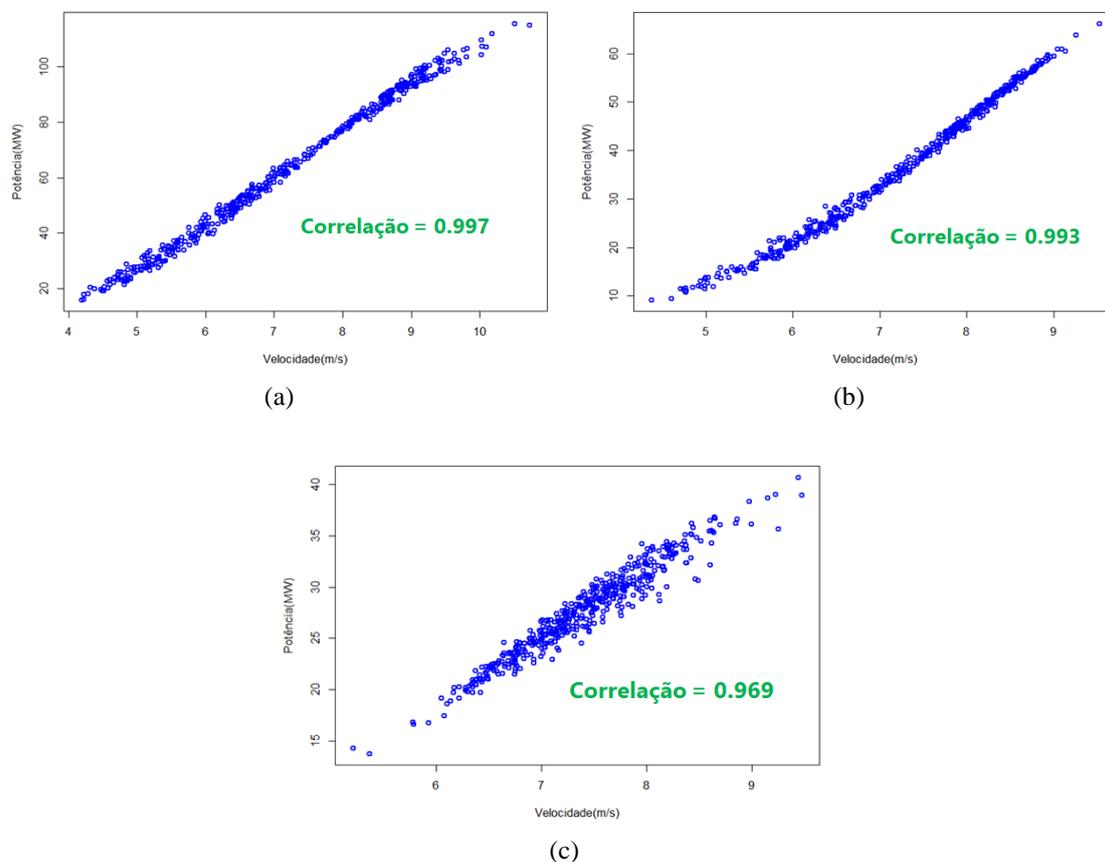
Na sequência as estimativas horárias foram agrupadas por mês para a obtenção de séries temporais das médias mensais da velocidade de vento e produção eólica. A partir destes dados, foram obtidas as curvas de potência probabilísticas mensais (diagramas de dispersão) para os PEEs de cada uma das subestações anteriores, conforme apresentados na Figura 5.3. A partir desta figura se constata que, para os PEEs analisados, as correlações entre as médias mensais da velocidade de vento e produção eólica são superiores a 99% para o Nordeste, e superior a 97% para o Sul.

Figura 5.2 – Curvas de potência probabilísticas *horárias* - dados do Sintegre (vermelho) e procedimento proposto usando dados do MERRA 2 (azul) – PEEs NE-L (a), NE-I (b) e Sul (c)



Fonte: Autores

Figura 5.3 – Curva de potência probabilística *mensal* obtida pela aplicação do procedimento proposto aos dados de reanálise oriundos do MERRA 2 – PEEs NE-L (a), NE-I (b) e Sul (c)



Fonte: Autores

5.2 - Avaliação das Funções de Transferências Mensais (FTMs) entre Ventos e Produção Eólica– Etapa 2(b)

Quando se comparam as Figuras 5.2 e 5.3, constata-se que, como esperado, a dispersão das curvas de potência probabilísticas mensais são suavizadas em relação àquelas em escala horária. Esta característica possibilita a construção de Funções de Transferências Mensais (FTMs) entre vento e produção eólica.

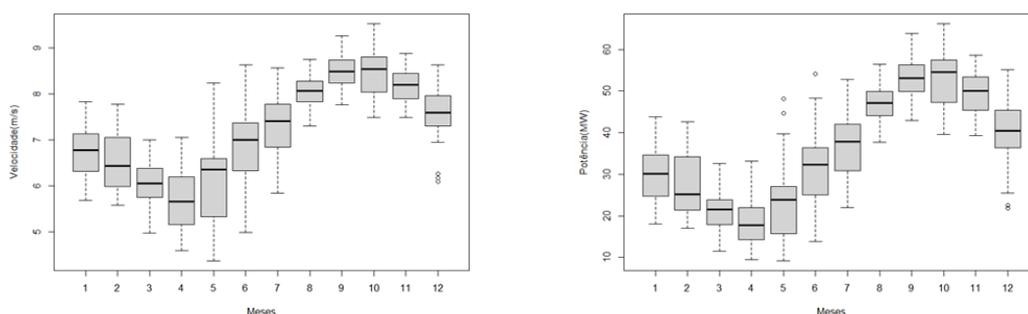
As abordagens para construção das FTMs entre ventos e produções eólicas mensais incluem os métodos estatísticos para análise de regressão, desde os tradicionais modelos de regressão linear até abordagens mais flexíveis, como os modelos aditivos (por exemplo, regressões lineares por parte e *splines*) (WOOD, 2006).

Em havendo dados suficientes, um aspecto a ponderar é se, para um dado PEE, dever-se-ia construir uma única FTM – válida para todos os meses do ano, ou várias FTMs, com intuito de capturar eventuais sazonalidades presentes. Isto irá depender do

comportamento das CPPMs associadas a cada PEE, as quais devem ser obtidas e analisadas sob esta perspectiva. Devido a limitação de espaço, este aspecto é ilustrado apenas para o PEE NE-L, abordado na Seção 5.1. Conclusões semelhantes são obtidas para os demais PEEs.

A partir das séries temporais das médias mensais da velocidade de vento e produção eólica calculadas na Seção 5.1 para o PEE NE-L, a Figura 5.4 exhibe, para cada mês do ano, os *boxplots* dos correspondentes valores médios de velocidade de vento e produção eólica, onde se observa um comportamento sazonal, com as medianas apresentando valor mínimo em abril e valores máximos em setembro e outubro; também se verifica que os meses de valores médios mais elevados apresentam menor dispersão que aqueles com valores mais baixos.

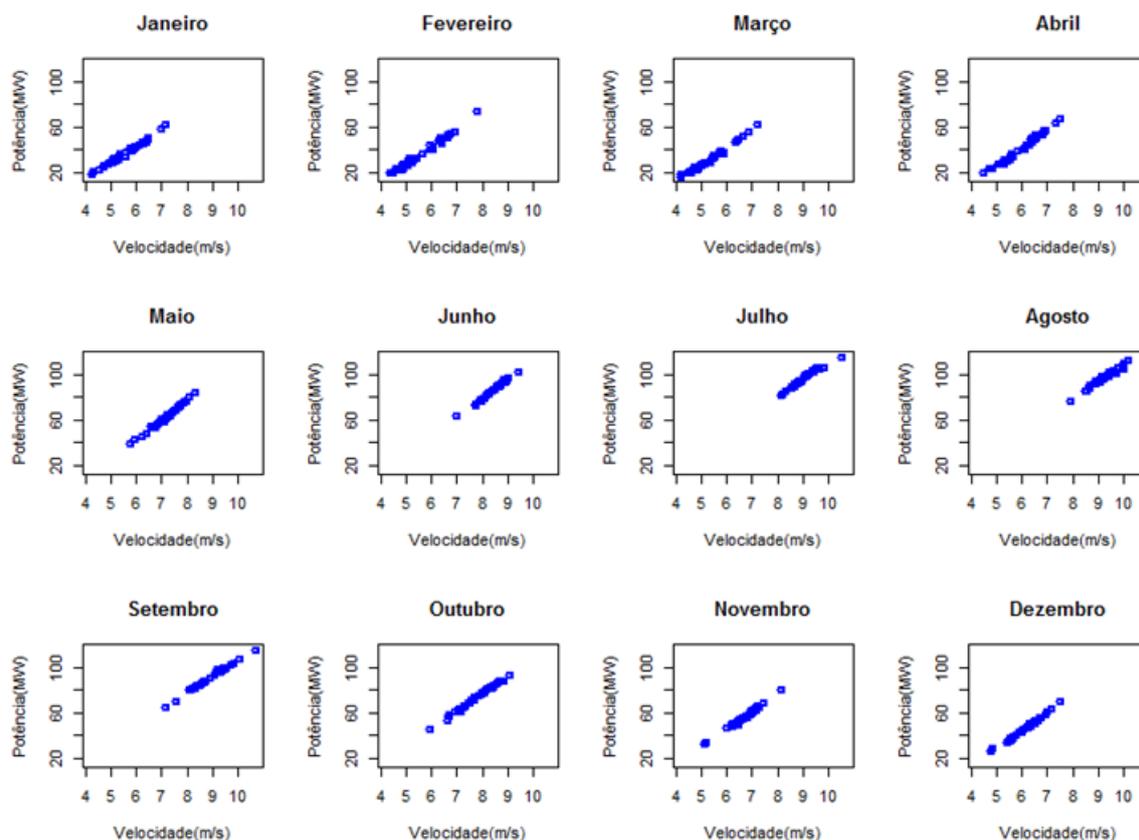
Figura 5.4 – *Boxplots* das médias mensais de velocidades de vento e produção eólica – PEE NE-L



Fonte: Autores

Este comportamento sazonal sugere a construção e análise de CPPMs para cada mês do ano, conforme apresentado na Figura 5.5. O agrupamento das CPPMs de cada mês pode ser realizada por várias técnicas, e.g., utilizando o Mapa de Kohonen (KOHONEN, 1989; PESSANHA *et al.*, 2020b). Uma inspeção visual dessas CPPMs sugere o seu agrupamento, por exemplo, em três classes: uma associada aos meses mais quentes (Dez-Jan-Fev-Mar-Abr) e denominada DJFMA, outra aos meses mais frios (Jun-Jul-Ago-Set-Out) – JJASO, e a última, relacionada a meses de transição (Mai e Nov) - MN.

Figura 5.5 – Curvas de Potência Probabilísticas Mensais para cada mês do ano – PEE NE-L



Fonte: Autores

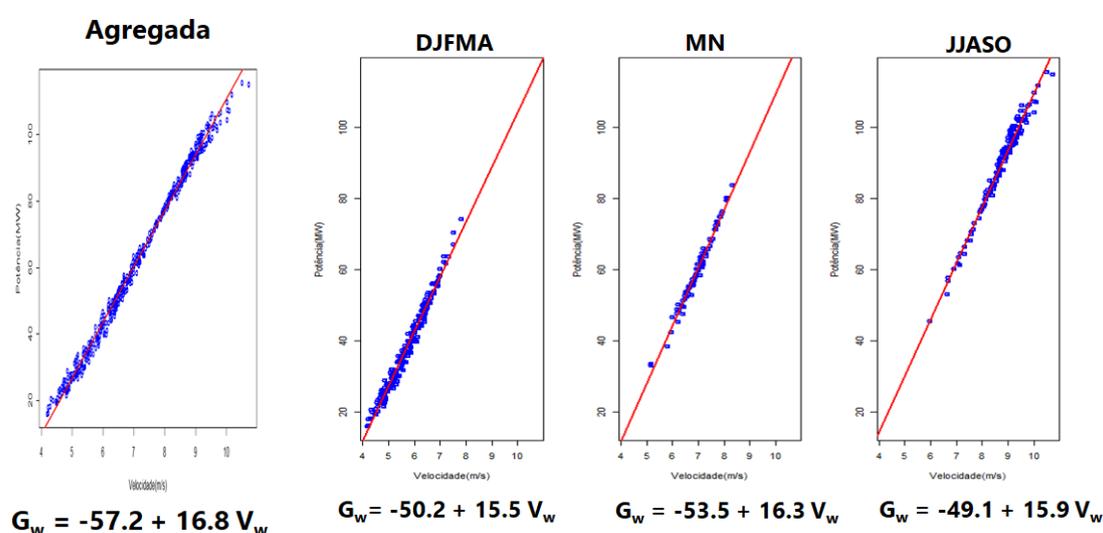
Tanto para a CPPM agregada (obtida considerando todos os meses), quanto para cada uma das três CPPMs sazonais obtidas anteriormente, foram construídas Funções de Transferência (FTMs) entre ventos e produções eólicas mensais por meio da aplicação do método de regressão linear simples. As FTMs ajustadas estão indicadas na Figura 5.6, tanto em forma gráfica quanto por meio das equações de regressão. Observa-se que o ajuste da FTM por meio de uma equação afim se constitui em uma boa representação; este resultado não surpreende dado o elevado valor da correlação entre as médias mensais da velocidade de vento e produção eólica obtida para este PEE (99.7%). Constata-se ainda que não há ganho, neste caso, em se adotar várias FTMs, sendo bastante razoável utilizar a FTM agregada.

Dependendo da forma da CPPM, o ajuste da FTM associada por uma única regressão linear pode não ser adequada. Neste caso, recomenda-se a utilização de modelos aditivos, e.g., regressões lineares por parte e *splines* (WOOD, 2006). Embora não pareça ser o caso do PEE NE-L, mas com o intuito de ilustrar a sua modelagem, a Figura 5.7

apresenta duas possíveis FTMs agregadas ajustadas utilizando o modelo de regressão linear por partes, considerando duas possibilidades para os pontos de quebra (*break points*), i.e., 5,3 e 9,2. Observe que, neste caso, na equação de regressão linear por parte aparece uma variável fictícia (*dummy*) V^* , onde $V^* = 0$ se a velocidade média mensal do vento for inferior ao ponto de quebra; caso contrário, $V^* = 1$.

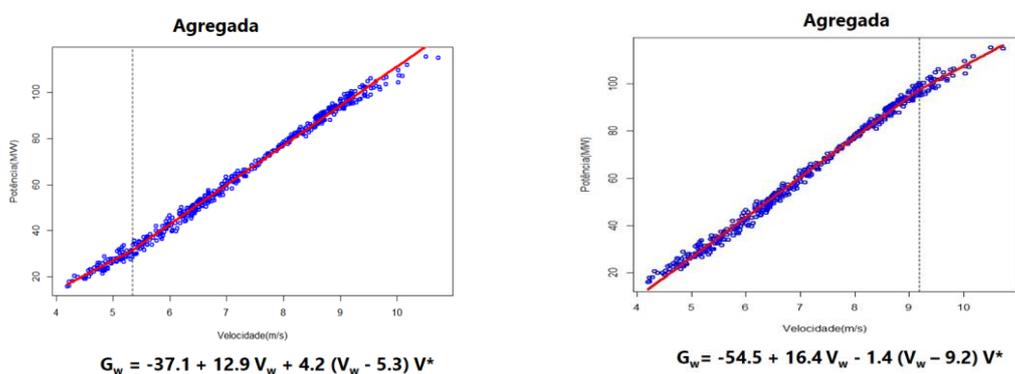
Finalmente, recomenda-se que as FTMs sejam atualizadas a cada ano, à medida que mais dados sejam disponibilizados e incorporados às análises.

Figura 5.6 – Funções de Transferência Mensais - regressão linear simples – PEE NE-L



Fonte: Autores

Figura 5.7 – Funções de Transferência Mensais - regressões lineares por partes – PEE NE-L



Fonte: Autores

6. Geração de Séries Sintéticas de Ventos – Etapa 3

Conforme descrito na Seção 2, faz parte do modelo NEWAVE um módulo responsável pela geração de séries sintéticas de afluências (energias afluentes aos REEs

ou vazões afluentes aos reservatórios de usinas hidroelétricas), denominado GEVAZP (MACEIRA & BEZERRA, 1997; PENNA *et al.*, 2011). São geradas séries sintéticas para a simulação *forward*, para a recursão *backward* e para a simulação da operação do sistema interligado com a política de operação obtida.

O modelo GEVAZP emprega a classe de modelos de séries temporais em que parte ou todos os parâmetros variam dentro do ano, os quais são conhecidos como modelos auto-regressivos com parâmetros periódicos e denotados por modelos PAR(p).

O modelo auto-regressivo periódico, caracterizado por PAR(p_1, \dots, p_s), pode ser escrito como:

$$\phi^m(B) \left(\frac{Z_t - \mu_m}{\sigma_m} \right) = a_t \quad (4)$$

onde:

Z_t é a variável aleatória do processo estocástico com s períodos sazonais e, nesta aplicação, representa a afliência ao REE ou à UHE no estágio t ;

O índice de tempo t pode ser considerado como uma função do ano T , e do período sazonal m : $t = (T - 1) s + m$;

$$\phi^m(B) = (1 - \phi_1^m B - \dots - \phi_{p_m}^m B^{p_m});$$

B é o operador defasagem no estágio t . $B Z_t = Z_{t-1}$;

μ_m é a média do processo estocástico do período sazonal m correspondente ao estágio t ;

σ_m é o desvio-padrão do processo estocástico do período sazonal m correspondente ao estágio t .

A série a_t não correlacionada temporalmente é independente de Z_t , possui média zero e variância $\sigma_a^{2(m)}$ e pode ser escrita como uma função das auto-correlações $\rho^m(k)$ de Z_t e dos coeficientes auto-regressivos periódicos (SALAS *et al.*, 1980; FRANSES & PAAP, 2006).

A função de auto-correlação $\rho^m(k)$ de Z_t é obtida multiplicando-se ambos os lados da equação (1) por $\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right)$ e aplicando o valor esperado a todos os termos desta equação. Ela satisfaz a equação de diferenças (sistema de equações de Yule-Walker). Para cada período m , pode-se obter um conjunto de equações lineares, o sistema de equações de Yule Walker estendido para modelos PAR(p). Os coeficientes auto-regressivos periódicos $\phi_1^m, \phi_2^m, \dots, \phi_{p_m}^m$ são estimados resolvendo o sistema de equações lineares de

Yule Walker de cada período m , onde os parâmetros μ_m e σ_m^2 e a função de auto-correlação $\rho^m(k)$ de Z_t são obtidos pelas suas estimativas de momento.

A função de auto-correlação parcial (FACP) é outra forma de se representar a estrutura de dependência temporal de um modelo. É útil na identificação da ordem do modelo auto-regressivo que deve ser ajustado a uma série temporal histórica. Em (HIPEL & MCLEOD, 1994) foi definido a FACP periódica para um modelo PAR(p) de forma semelhante à feita para um modelo de auto-regressivo não periódico - AR(p). Para cada período m , a FACP periódica é definida pelo conjunto dos últimos parâmetros auto-regressivos dos modelos auto-regressivos de ordem k , $k = 1, \dots$, deste período m , usualmente denominados de ϕ_{kk}^m . Como decorrência dessa definição, a FACP do período m deve ser igual a zero após o lag p_m , quando a ordem do modelo AR(p) nesse período é p_m (HIPEL & MCLEOD, 1994). Também como consequência dessa definição, a resolução de cada um dos sistemas de equações de Yule-Walker fornece ϕ_{kk}^m até lag $k = p_m$. Maiores detalhes do modelo podem ser encontrados em (MACEIRA & BEZERRA, 1997; PENNA *et al.*, 2011).

Para gerar cenários sintéticos de afluições com o modelo PAR(p) foi empregada uma metodologia na qual, inicialmente, é necessário obter uma amostra de ruídos, a_t , não correlacionados temporalmente, uma para cada período de tempo e para cada cenário de afluições. Atualmente, essas amostras são produzidas por amostragem seletiva a partir de uma distribuição Lognormal multivariada (PENNA *et al.*, 2011). O método de amostragem seletiva consiste na aplicação de técnicas de agrupamento ao processo de geração de cenários de afluições. Nesse caso, as técnicas de agrupamento são usadas para obter amostras de séries de ruídos, a_t , multivariados e não correlacionados no tempo.

O esquema adotado para geração de cenários sintéticos de afluições pode ser resumido da seguinte forma:

- ✓ Obter as afluições incrementais mensais históricas para cada UHE ou REE;
- ✓ Escolher a ordem do modelo AR de cada período sazonal para cada UHE ou REE, examinando as parcelas da FACP;
- ✓ Estimar os coeficientes dos modelos PAR(p) através dos sistemas de equações de Yule-Walker;
- ✓ Gerar uma amostra muito grande (por exemplo, de tamanho 100.000) de resíduos a_t normais, não correlacionados temporalmente e espacialmente, usando amostragem aleatória simples, onde são tratados como equiprováveis;

- ✓ Aplicar o método K-means para reduzir a cardinalidade da amostra original; os resíduos resultantes passam, então, a ser tratados como não equiprováveis;
- ✓ Com o intuito de gerar aflúências mensais multivariadas, assume-se que ruídos normais padrão não correlacionados espacialmente, a_t , podem ser transformados em ruídos espacialmente correlacionados, e_t , através da seguinte relação:

$$e_t = D a_t \quad (5)$$

A matriz D pode ser estimada por (KELMAN *et al.*, 1987):

$$DD^T = \hat{U} \quad (6)$$

onde \hat{U} pode ser, em princípio, a estimativa da matriz de covariâncias dos resíduos a_t . Na prática, o comportamento dos resíduos não segue o comportamento das aflúências: os resíduos não são espacialmente correlacionados. No entanto, a fim de se preservar as dependências espaciais entre os REEs/UHEs utiliza-se a correlação espacial entre as aflúências em substituição à correlação espacial entre os resíduos;

- ✓ Uma distribuição Lognormal três parâmetros é ajustada aos resíduos espacialmente correlacionados a fim de melhor reproduzir a assimetria observada nesse tipo de processo estocástico;
- ✓ Os cenários sintéticos são obtidos usando-se a equação (4);
- ✓ A soma das aflúências incrementais ao longo da cascata de UHEs, em cada período e cenário, é realizada para obter as aflúências totais.

Maiores detalhes da metodologia para geração de cenários sintéticos são obtidos em (MACEIRA & BEZERRA, 1997, PENNA *et al.*, 2011).

Em 2019, foi proposto a introdução de um novo termo na equação auto-regressiva periódica (TREISTMAN *et al.*, 2020), que representa a média das 12 últimas aflúências, a fim de incorporar mais memória nos cenários de aflúências gerados, conseguindo reproduzir com mais acurácia a auto-correlação anual lag 1 e também a função de auto-correlação mensal das aflúências. Foi denominado de modelo auto-regressivo periódico com componente anual, PAR(p)-A e pode ser escrito como:

$$\phi^m(B) \left(\frac{ENA_t - \mu_m}{\sigma_m} \right) - \psi^m \left(\frac{A_{t-1} - \mu_{m-1}^A}{\sigma_{m-1}^A} \right) = a_t \quad (7)$$

onde A_{t-1} representa a média das últimas 12 observações do processo estocástico ENA_t com referência a $(t-1)$:

$$A_{t-1} = \sum_{\tau=1}^{12} \frac{ENA_{t-\tau}}{12} \quad (8)$$

e μ_{m-1}^A , σ_{m-1}^A são, respectivamente, a média e o desvio-padrão de A_{t-1} .

O esquema para geração de cenários sintéticos de aflúências não sofre alterações.

Um dos objetivos deste trabalho é ajustar um modelo estocástico para gerar séries sintéticas de ventos aos parques eólicos individuais (EOLs) ou equivalentes (PEEs) do SIN que atenda às seguintes condições, elencada na Seção 3:

- A estrutura de correlação temporal, porventura verificada no processo estocástico dos ventos médios mensais (m/s) em algum EOL ou PEE, não será representada de forma explícita no modelo de geração de séries sintéticas. Este condicionante se impõe para evitar um acréscimo no número de variáveis de estado consideradas no algoritmo de PDDE, hoje já bem elevado (144, no caso da representação por REEs ou 1.771, quando se modela por UHEs, considerando $p_m = 11$);
- Será representada a correlação espacial verificada entre os processos estocásticos dos ventos e afluições aos REEs ou UHEs. Esta estatística apresenta-se elevada em muitos dos meses do ano, conforme verificado na Tabela 6.1. Esta Tabela apresenta as correlações entre os PEEs analisados na Seção 5 (NE-L, NE-I e Sul) e alguns dos REEs considerados no Programa Mensal de Operação (PMO) do ONS. Estes resultados são seguidos por outros parques eólicos do País, daí a importância de ser levado em consideração;

Tabela 6.1 - Correlação espacial entre os PEEs NE-L, NE-I e Sul e alguns REEs do SIN.

NE-L	Sudeste	Paraná	Nordeste	Norte	NE-I	Sudeste	Paraná	Nordeste	Norte	Sul	Paraná	Belo Monte	Sul	Iguaçu
Fevereiro	-0.38	-0.45	-0.36	-0.36	Maio	0.47	0.31	0.49	0.23	Fevereiro	0.51	0.29	0.33	0.37
Agosto	-0.54	-0.60	-0.36	-0.30	Junho	-0.45	-0.62	-0.27	-0.39	Abril	-0.32	-0.57	0.26	0.40
Setembro	-0.70	-0.89	-0.56	-0.69	Setembro	-0.33	-0.44	-0.22	-0.31	Maio	0.16	-0.32	0.35	0.32
Outubro	-0.59	-0.69	-0.49	-0.75	Dezembro	-0.52	0.20	-0.28	-0.52	Outubro	-0.38	-0.30	0.36	0.31

Fonte: Autores

A proposta deste trabalho é adotar um **modelo integrado de geração de séries sintéticas de afluições e ventos** que considere os condicionantes listados. Em uma primeira proposta, a variável aleatória do processo estocástico com s períodos sazonais que representa o vento em um parque eólico no estágio t é dada por:

$$\left(\frac{V_t - \mu_m^v}{\sigma_m^v}\right) = \text{componente explicativa} + a_t \quad (9)$$

Reescrevendo:

$$V_t = \text{componente explicativa} + \sigma_m^v a_t \quad (10)$$

A *componente explicativa* pode ser a própria média do processo estocástico vento do período sazonal m correspondente ao estágio t , μ_m^v , ou conter uma parcela relacionada às afluições do próprio estágio t , ou do estágio $t-1$.

O esquema adotado para geração de cenários sintéticos de afluições deve ser então modificado a fim integrar o processo estocástico dos ventos, dando origem ao **modelo integrado de afluições e ventos mensais**, descrito por (as alterações estão marcadas em negrito):

- ✓ Obter as afluições incrementais mensais históricas para cada usina hidrelétrica ou reservatório equivalente de energia;
- ✓ Escolher a ordem do modelo AR de cada período sazonal para cada UHE ou REE, examinando as parcelas da FACP;
- ✓ Estimar os coeficientes dos modelos PAR(p) através dos sistemas de equações de Yule-Walker;
- ✓ Gerar uma amostra muito grande (por exemplo, 100.000) de resíduos a_t Normais, não correlacionados temporalmente e espacialmente, usando amostragem aleatória simples onde são tratados como equiprováveis. **Será também gerada amostra de mesma cardinalidade para os ruídos dos ventos;**
- ✓ Aplicar o método K-means para reduzir a cardinalidade da amostra original; os resíduos resultantes passam, então, a ser tratados como não equiprováveis. **A amostra de resíduos correspondentes aos parques eólicos também terá a cardinalidade reduzida empregando-se a mesma metodologia;**
- ✓ Com o intuito de gerar **afluições e ventos mensais multivariados**, assume-se que ruídos normais padrão não correlacionados espacialmente, a_t , podem ser transformados em ruídos espacialmente correlacionados, e_t , através das relações (5) e (6), repetidas abaixo:

$$e_t = D a_t$$

A matriz D pode ser estimada por:

$$DD^T = \hat{U}$$

onde \hat{U} passa a ser, em princípio, a estimativa da **matriz de covariâncias dos resíduos a_t , correspondentes às afluições e ventos**. Na prática, o comportamento dos resíduos não segue o comportamento das afluições, nem dos ventos, nem entre afluições e ventos: os resíduos não são espacialmente correlacionados. No entanto, a fim de se preservar as dependências espaciais entre os REEs/UHEs, **entre os EOLs/PEEs, e entre REEs/UHEs e EOLs/PEEs, utiliza-se a correlação espacial entre as afluições, entre os ventos e entre as afluições e ventos** em substituição à correlação espacial entre os resíduos;

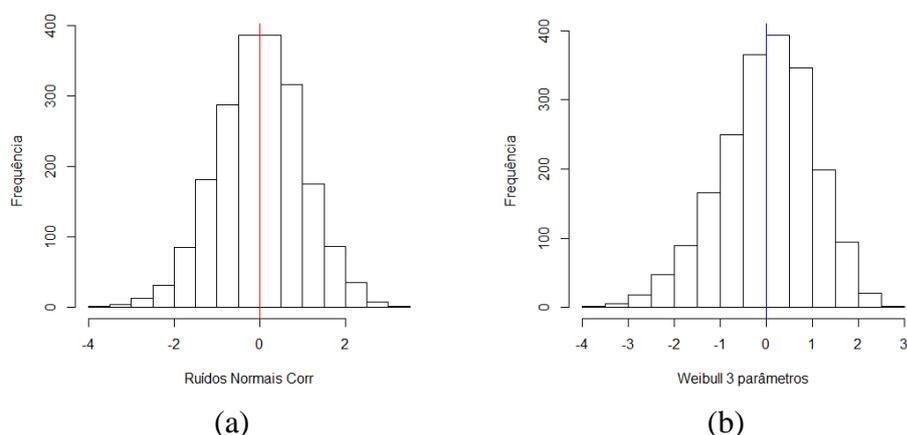
- ✓ Uma distribuição Lognormal três parâmetros é ajustada aos resíduos espacialmente correlacionados a fim de melhor reproduzir a assimetria observada nesse tipo de processo estocástico. **Este passo também será aplicado aos resíduos correspondentes aos EOLs/PEEs, já que as séries históricas de vento, medidas ou obtidas por reanálise também podem apresentar uma assimetria elevada conforme ilustrado na Tabela 6.2.** Esta tabela apresenta também as principais estatísticas para as séries temporais das velocidades mensais de vento dos PEEs NE-L, NE-I e Sul;

No entanto, caso a assimetria observada seja negativa, a distribuição Weibull três parâmetros substitui a Lognormal.

Na Figura 6.1 estão ilustradas as distribuições de frequência dos ruídos Normais e dos ruídos Weibull, correlacionados, referentes ao mês de julho. Elas resultam de uma geração de 100.000 resíduos Normais, inicialmente não correlacionados temporalmente e espacialmente, que foram agregados através do método K-means

em 2.000 elementos e a seguir correlacionados, Figura 6.1(a). A seguir esses ruídos foram transformados em ruídos Weibull também correlacionados, Figura 6.1(b).

Figura 6.1 – Distribuição de frequências dos ruídos Normais e dos ruídos Weibull



Fonte: Autores

Tabela 6.2 – Principais estatísticas para as séries temporais das velocidades mensais de vento - PEEs NE-L, NE-I e Sul

		Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
NE - L	média	7,76	7,53	6,92	7,04	7,85	8,60	9,17	9,59	9,66	9,37	8,89	8,38
	d.padrão	0,73	0,82	1,05	1,13	0,71	0,49	0,49	0,45	0,57	0,52	0,31	0,41
	c. variação	0,09	0,11	0,15	0,16	0,09	0,06	0,05	0,05	0,06	0,06	0,03	0,05
	assimetria	-0,30	-0,64	-0,69	-0,68	-1,27	-0,78	-0,43	0,23	0,09	-0,20	0,26	-0,22
NE - I	média	5,48	5,68	5,47	5,97	6,31	6,76	7,05	7,15	7,09	6,63	5,84	5,46
	d.padrão	0,53	0,64	0,60	0,46	0,39	0,40	0,44	0,39	0,45	0,52	0,52	0,50
	c. variação	0,10	0,11	0,11	0,08	0,06	0,06	0,06	0,05	0,06	0,08	0,09	0,09
	assimetria	-0,13	0,46	0,08	-0,83	-0,54	-1,25	1,17	0,28	0,08	-0,60	-0,64	0,00
Sul	média	6,68	6,35	6,39	6,40	6,25	6,22	6,59	6,68	7,28	7,22	7,09	6,84
	d.padrão	0,60	0,50	0,57	0,61	0,60	0,52	0,38	0,56	0,50	0,60	0,55	0,59
	c. variação	0,09	0,08	0,09	0,10	0,10	0,08	0,06	0,08	0,07	0,08	0,08	0,09
	assimetria	-0,32	-0,29	0,59	0,39	-0,22	0,38	-0,42	1,57	0,18	1,07	0,30	0,40

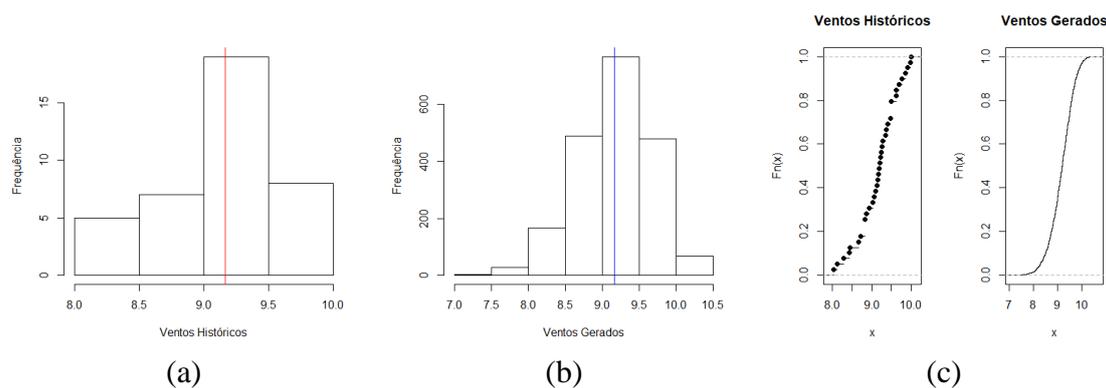
Fonte: Autores

Na Figura 6.2 estão ilustradas, para o PEE NE-L e no mês de julho, as distribuições de frequência dos ventos históricos de uma amostra de 40 elementos (a) e sintéticos, de uma amostra de 2.000 elementos (b). Já a Figura 6.2(c) apresenta as respectivas distribuições de frequência acumuladas deste mês. Os resultados indicam boa aderência entre as séries sintéticas gerada pelo modelo Weibull três parâmetros e as séries históricas.

- ✓ Os cenários sintéticos são obtidos usando-se as expressões (4) ou (7) e (10);
- ✓ A soma das afluências incrementais ao longo da cascata de UHEs, em cada período e cenário, é realizada para obter as afluências totais.

Conforme evidenciado na Tabela 6.1, existem correlações cruzadas relevantes entre os PEEs e REEs analisados, na sua grande maioria é negativa, mas valores positivos também podem ser encontrados. Também se verificou que os PEEs analisados apresentam distribuições de frequências com assimetrias elevadas em algumas meses (Tabela 6.2), em geral negativas, diferentemente das assimetrias dos REEs, que são sempre positivas.

Figura 6.2 – Histogramas e distribuição de frequências acumuladas dos ventos históricos e sintéticos



Fonte: Autores

Durante o decorrer da pesquisa, estas estatísticas serão calculadas para outros EOLs e PEEs, especialmente para aqueles resultantes da agregação estatística dos regimes de ventos obtidos na Etapa 1 da abordagem proposta, e orientarão a geração das séries sintéticas de produção de velocidades de ventos.

7. Representação das Produções Eólicas Mensais no Algoritmo PDDE do Modelo NEWAVE – Etapa 4

Após a geração de séries sintéticas de ventos mensais (Etapa 3), é preciso obter a correspondente produção de cada parque eólico individual (EOL) ou equivalente (PEE) para ser considerada no problema de despacho mensal da operação do modelo NEWAVE. As produções eólicas são computadas por meio de funções de transferência entre as médias mensais da velocidade de vento e da produção eólica (FTMs), obtidas na Etapa 2 da metodologia proposta.

Uma vez obtidas as produções eólicas dos parques individualizados ou agregados, as mesmas poderão ser abatidas diretamente dos mercados dos subsistemas associados, da mesma maneira como atualmente se procede, onde as estimativas médias mensais não consideram a incerteza dessa fonte. Outra alternativa, mais atraente e aqui adotada,

consiste em representá-las no problema de despacho como uma fonte de geração disponível, porém com custo de operação zero, conforme descrito a seguir.

Para este fim, a formulação compacta do problema LSP, em cada estágio t e cenário *forward* s , e cenário *backward* ω , descrito em (2), será apresentada em sua formulação mais detalhada (MACEIRA *et al.*, 2020), modificando-se ou acrescentando-se as equações em que as variáveis de produção eólica aparecem de forma explícita, conforme em (11), além da função objetivo. Também, por simplificação, serão omitidos os subscritos s e ω .

Por fim, no modelo integrado proposto e descrito por (10), a *componente explicativa* pode ser a própria média do processo estocástico dos ventos do período sazonal m , ou conter uma parcela relacionada às afluências do próprio estágio t , ENA_t , ou do estágio $t-1$, ENA_{t-1} . Em consequência, chama-se a atenção que cada uma dessas opções de modelagem impacta de forma distinta a construção dos cortes de Benders relativa às variáveis de estado afluências aos REEs nos períodos $t-1$, ..., $t-p$, e a sua derivação foge ao escopo deste trabalho.

<p><i>Função Objetivo</i></p> $ \begin{aligned} z_t = \min & \sum_{k=1}^{NSBM} \left(\sum_{c=1}^{NPMC} \left(\sum_{\substack{ictl=1 \\ LAG_{ictl}=0}}^{NCLT_k} cterm_{t,ictl} \cdot GT_{t,ictl,c} \right) \right) + \\ & \sum_{m=1}^{NSBM} \left(\sum_{c=1}^{NPMC} \left(\sum_{ilag=1}^{LAGMAX} \sum_{\substack{ictl=1 \\ LAG_{ictl}=ilag}}^{NCLT_m} \left(\frac{1}{1+\beta} \right)^{ilag} \cdot CTERM_{t+ilag,ictl} \cdot GT_{t+ilag,ictl,c} \right. \right. \\ & \left. \left. + \sum_{idef=1}^{NPDF} CDEF_{t,m,idef} \cdot DEF_{t,m,idef,c} \right) \right) \\ & + \sum_{i=1}^{NREE} \left(\sum_{c=1}^{NPMC} p^{GHMIN} \cdot \delta_{t,i,c}^{GHMIN} + PEN^{EARM} \cdot \delta_{t,i}^{EARM} + PEN^{DESV} \cdot (\delta_{t,i}^{DSVC} + \delta_{t,i}^{DSVF}) + PEN_1^{EVMIN} \cdot \delta_{i,1}^{EVMIN} \right. \\ & \left. + PEN_2^{EVMIN} \cdot \delta_{i,2}^{EVMIN} \right) \\ & + \sum_{j=1, j \neq m}^{NNSBM} p^{INT} \cdot INT_{t,j,m,c} + \sum_{i=1}^{NREE} p^{EVERT} \cdot EVERT_{t,i} + \sum_{m=1}^{NSBM} \left(\sum_{c=1}^{NPMC} p^{EXC} \cdot EXC_{t,m,c} \right) \\ & + \sum_{c=1}^{NPMC} \sum_{ko=1}^{NSBM} \sum_{kd=1}^{NSBM} pen^{INTMIN} \cdot \delta_{t,ko,kd,c}^{INTMIN} \\ & + \frac{1}{1+\beta} \cdot \alpha_{t+1} \end{aligned} $	<p>(11a)</p>
<p><i>Equação de atendimento à demanda em cada subsistema k, para cada patamar de carga c, por estágio t</i></p>	<p>(11.b)</p>

$\sum_{i=1}^{NREE_m} (perdc_{t,i,c} \cdot a_{t,i}^{ACPL} \cdot GHC_{t,i,c} + perdf_{t,i,c} \cdot fpeng_{t,c} \cdot GFIOl_{t,i}) + \sum_{iclt=1}^{NCLT_m} perdt_{t,iclt,c} \cdot GT_{t,iclt,c}$ $+ \sum_{idef=1}^{NPDF} DEF_{t,m,idef,c} + \sum_{j=1, j \neq m}^{NSBM} [perdi_{t,j,m,c} \cdot INT_{t,j,m,c} - perdi_{t,m,j,c} \cdot INT_{t,m,j,c}] - EXC_{t,m,c}$ $+ \sum_{u=1}^{NPEE_m} GW_{t,u,c} = merc_{t,m,c} + cadic_{t,m,c}$ $- \left(submot_{t,m} + pqusi_{t,m} + \sum_{iclt=1}^{NCLT_m} gmin_{t,m,iclt} \right) \cdot fpeng_{t,c} - \sum_{l=1}^{LAGMAX} sgt_{t-l,m,c}$	
<p><i>Equação de balanço hídrico controlável por estágio t para cada REE i</i></p>	(11.c)
<p><i>Equação de balanço hídrico a fio d'água por estágio t para cada REE i</i></p>	(11.d)
<p><i>Perdas a fio d'água por estágio t para cada REE i</i></p>	(11.e)
<p><i>Geração hidráulica mínima por estágio t para cada REE i e patamar de carga c</i></p>	(11.f)
<p><i>Geração hidráulica máxima por estágio t para cada REE i e patamar de carga c</i></p>	(11.g)
<p><i>Limites de armazenamento de cada reservatório equivalente i, por estágio t</i></p>	(11.h)
<p><i>Armazenamento mínimo operativo de cada reservatório equivalente i, por estágio t</i></p>	(11.i)
<p><i>Máxima violação do armazenamento mínimo operativo até o estágio t</i></p>	(11.j)
<p><i>Retirada/Acréscimo de água para outros usos por estágio t para cada REE i</i></p>	(11.k)
<p><i>Deplecionamento mínimo por estágio t para cada REE i</i></p>	(11.l)
<p><i>Limite máximo de geração de cada usina térmica j, por estágio t</i></p>	(11.m)
<p><i>Fluxo nas interligações do subsistema i que não tem carga e nem usinas, por estágio t</i></p>	(11.n)
<p><i>Limite mínimo de intercâmbio entre os subsistemas i e j por estágio t e patamar de carga c</i></p>	(11.o)
<p><i>Agrupamento de intercâmbio por estágio t em cada patamar de carga c</i></p>	(11.p)
<p><i>Somatório da geração térmica antecipada por estágio t em cada patamar de carga c</i></p>	(11.q)
<p><i>Produção eólica através da função de transferência mensal (uma única equação de regressão) para cada PEE u por estágio t</i></p> $\sum_{c=1}^{NPMC} GW_{t,u,c} \leq b_{t,u}^W + a_{t,u}^W v_{t,u}$	(11.r)
<p><i>Cortes de Benders icor = 1,... (função de custo futuro)</i></p>	(11.s)

onde:

- α_t Valor esperado do custo total de operação do estágio t até o final do período de planejamento T, também chamado de função de custo futuro do estágio t;
- $GHC_{t,i,c}$ Geração hidráulica controlável do reservatório equivalente i, para o patamar de carga c, no estágio t;
- $GFIOl_{t,i}$ Geração hidráulica não controlável do reservatório equivalente i, no estágio t;
- $GT_{t,iclt,c}$ Geração da classe térmica iclt, para o patamar de carga c, no estágio t;

$GW_{t,u,c}$	Geração eólica do parque eólico equivalente u , para o patamar de carga c , no estágio t ;
$DEF_{t,k,idef,c}$	Energia não suprida no subsistema k , para o patamar de déficit $idef$, para o patamar de carga c , no estágio t ;
$EVERT_{t,i}$	Energia vertida do reservatório equivalente i , no estágio t ;
$V_{t,u}$	Vento do parque eólico equivalente u , no estágio t ;
$submot_{t,m}$	Energia de submotorização do subsistema m , no estágio t ;
$pqusi_{t,m}$	Geração das usinas não despachadas centralizadamente do subsistema m , no estágio t ;
$SGT_{t-1,m,c}$	Geração térmica antecipada do subsistema k , para o patamar de carga c , definida no estágio $t-1$;
$gtmin_{t,iclt,c}$	Geração térmica mínima da classe térmica $iclt$, do subsistema m , para o patamar de carga c , no estágio t ;
$INT_{t,j,m,c}$	Intercâmbio de energia do subsistema j para o subsistema m , para o patamar de carga c , no estágio t ;
$EXC_{t,m,c}$	Excesso de energia associado à equação de atendimento à demanda do subsistema m , para o patamar de carga c , no estágio t ;
$merc_{t,m,c}$	Mercado de energia do subsistema m , para o patamar de carga c , no estágio t ;
$cadic_{t,m,c}$	Carga adicional do subsistema m , para o patamar de carga c , no estágio t ;
$perdc_{t,i,c}$	Perda hidráulica controlável no reservatório equivalente i , no patamar de carga c e no estágio t ;
$fpeng_{t,c}$	Fator de ponderação da energia associada ao patamar de carga c ;
$perdf_{t,i,c}$	Perda hidráulica fio d'água no reservatório equivalente i , no patamar de carga c e no estágio t ;
$perdt_{t,iclt,c}$	Perda térmica na classe térmica $iclt$, no patamar de carga c e no estágio t ;
$perdi_{t,j,m,c}$	Perda por intercâmbio do submercado j para o m no patamar de carga c e no estágio t ;
$a_{t,u}^W$	Coeficiente angular da reta que representa a função de transferência mensal no parque eólico equivalente u , no estágio t ;
$b_{t,u}^W$	Termo independente da reta que representa a função de transferência mensal no parque eólico equivalente u , no estágio t ;
$\delta_{t,i,c}^{GHMIN}$	Violação da restrição de geração hidráulica mínima no reservatório equivalente i , para o patamar de carga c , no estágio t ;
$\delta_{t,i}^{EARM}$	Violação da restrição de armazenamento mínimo operativo no reservatório equivalente i , no estágio t ;
$\delta_{t,i}^{DSVC}$	Violação da restrição de desvio de energia controlável no reservatório equivalente i , no estágio t ;
$\delta_{t,i}^{DSVF}$	Violação da restrição de desvio de energia nas usinas à fio d'água no reservatório equivalente i , no estágio t ;
$\delta_{t,i,1}^{EVMIN}$	Violação do patamar 1 da restrição de vazão mínima no reservatório equivalente i , no estágio t ;
$\delta_{t,i,2}^{EVMIN}$	Violação do patamar 2 da restrição de vazão mínima no reservatório equivalente i , no estágio t ;

$\delta_{t,i,j,c}^{INTMIN}$	Violação da restrição de intercâmbio mínimo do subsistema i para o subsistema j , para o patamar de carga c , no estágio t ;
β	Taxa de desconto;
$cterm_{t,iclt}$	Custo de geração da classe térmica $iclt$ no estágio t ;
$cdef_{t,k,idef}$	Custo de não atendimento à demanda de energia do subsistema k e do patamar de déficit $idef$ no estágio t ;
pen^{GHMIN}	Penalidade por violação da restrição de geração hidráulica mínima;
pen^{EARM}	Penalidade por violação da restrição de armazenamento mínimo operativo;
pen^{DESV}	Penalidade por violação da restrição de retirada de água para outros usos;
pen_1^{EVMIN}	Penalidade por violação do patamar 1 da meta de vazão mínima;
pen_2^{EVMIN}	Penalidade por violação do patamar 2 da meta de vazão mínima;
pen^{INTMIN}	Penalidade por violação da restrição de intercâmbio mínimo;
p^{INT}	Penalidade de intercâmbio;
p^{EVERT}	Penalidade de vertimento;
p^{EXC}	Penalidade de excesso de energia;
$NSBM$	Número de subsistemas;
$NREE$	Número de reservatórios equivalentes;
$NCLT_k$	Número de classes térmicas do subsistema k ;
$NPMC$	Número de patamares de carga;
$NPDF$	Número de patamares de déficit;
$LAGMAX$	Número máximo de períodos para antecipação térmica das usinas a GNL;
$NPEE$	Número de parques eólicos equivalentes;
$a_{t,i}^{ACPL}$	Fator de acoplamento hidráulico referente à parcela controlável gerada no próprio reservatório equivalente i no estágio t ;

8. Conclusões e Trabalhos Futuros

Este artigo apresentou uma abordagem metodológica geral para modelar a incerteza da produção eólica mensal no problema de despacho de operação, aderente ao algoritmo de solução utilizado (programação dinâmica dual estocástica - PDDE) no modelo NEWAVE, permitindo a obtenção de resultados com a precisão adequada e dentro de um tempo computacional razoável, quando aplicada a sistemas interligados de grande porte, com predominância hidroelétrica, como é o caso do sistema brasileiro.

A abordagem proposta é composta por quatro etapas: (i) agrupamento estatístico dos regimes de ventos; (ii) avaliação de funções de transferência mensais entre ventos e produção eólica, a partir de curvas de potência probabilísticas mensais; (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.

Como não está disponível uma base de dados com séries horárias consistidas e pareadas de velocidades de ventos e produções eólicas, para os parques existentes, ou pelo menos para aqueles vencedores nos leilões públicos de compra de energia elétrica, tornou-se necessária a utilização de dados oriundos de reanálises.

Resultados iniciais da aplicação das etapas da metodologia proposta a configurações reais do sistema interligado brasileiro apontam para a sua efetividade, ao mesmo tempo lançam luz sobre possíveis aprimoramentos e novos desenvolvimentos.

Por fim, a experiência aqui obtida norteará o desenvolvimento de abordagem similar para a modelagem de outra fonte renovável intermitente – a geração solar fotovoltaica, para a qual se vislumbra um crescimento superior a 3,5 vezes nos próximos dez anos.

Referências

- 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.
- BIRGE, J.R. Decomposition and partitioning methods for multistage stochastic linear programs, **Operations Research**. Vol. 33, nº. 5, 989-1007, 1985.
- BOX, G.E.P.; JENKINS, G.M. **Time series analysis forecasting and control**. Holden-Day, 1976.
- DINIZ, A.L., COSTA, F.S., MACEIRA, M.E.P., SANTOS, T.N., BRANDÃO, L.C., CABRAL, R.N. Short/Mid-Term Hydrothermal Dispatch and Spot Pricing for Large-Scale Systems - the Case of Brazil, **20th Power Systems Computation Conference**, Dublin, Ireland, June 2018.
- FRANSES, P.H.; PAAP, R. **Periodic Time Series Models**, Oxford University Press, 2006.
- 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>.
- HIPPEL, W.H., MCLEOD, A.I. **Time Series Modeling of Water Resources and Environmental System**. Elsevier. 1994.
- JOHNSON, R.A. and WICHERN, D.W. **Applied Multivariate Analysis**, Prentice Hall, New Jersey, 1998.
- KELMAN, J., BARTH, F. T., POMPEU, C. T., FILL, H. D., TUCCI, C. E. M., BRAGA JR., B. P. F. **Modelos para Gerenciamento de Recursos Hídricos**, ABRH, 1987.
- KOHONEN, T. **Self-Organization and Associative Memory**; Springer Berlin Heidelberg, vol. 8., 1989.
- 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., CRUZ, C. B., DAMÁZIO, J. M., MELO, A. C. G., DINIZ, A. L., TREISTMAN, F. Consideração do Modelo Auto-Regressivo Periódico Anual (PAR(p)-A) no Algoritmo de Programação Dinâmica Dual Estocástica. **Relatório Técnico CEPEL 2002/ 020**, Julho de 2020.

MELO, A.C.G.; PESSANHA, J.F.M.; JUSTINO, T.C.; MACEIRA, M.E.P. Metodologia e Determinação de Perfis Típicos Diários de Carga através da Combinação de Técnicas Estatísticas de Análise de Agrupamentos e na Análise Exploratória de Dados, **Relatório Técnico CEPEL 10349/18**, 2018.

MELO, A.C.G.; PESSANHA, J.F.M.; JUSTINO, T.C.; MACEIRA, M.E.P. Uma Metodologia baseada em Técnicas Estatísticas de Agrupamento e Análise Exploratória de Dados para a Definição de Patamares de Carga em Estudos de Planejamento da Operação e da Expansão de Sistemas Elétricos de Potência, **XIV Simpósio de Especialistas em Planejamento da Expansão e da Operação de Sistemas Elétricos – SEPOPE**, Recife-PE, 30 de setembro a 3 de outubro, 2018.

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.

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.; BESSA, R.J.G.S.B.; PEÇAS LOPES, J.A. Propostas para a previsão probabilística da geração eólica, **XXIV Seminário Nacional de Produção e Transmissão de Energia Elétrica**, 2007.

PESSANHA, J.F.M.; MELO, A.C.G.; JUSTINO, T.C.; MACEIRA, M.E.P. “Combining Statistical Clustering Techniques and Exploratory Data Analysis to Compute Typical Daily Load Profiles – Application to the Expansion and Operational Planning in Brazil”, **IEEE Probabilistic Methods Applied do Power Systems – PMAPS 2018**, Boise, Idaho, USA, 24-28 June, 2018.

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 Simpósio Brasileiro de Sistemas Elétricos**, Santo André, 2020.

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.

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

STAFFELL, I., PFENNINGER, S. (2016). Using Bias-Corrected Reanalysis to Simulate Current and Future Wind Power Output. **Energy** 114, pp. 1224-1239.

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.

TREISTMAN, F., MACEIRA, M.E.P., DAMÁZIO, J., OLIVEIRA, C.C.B., Periodic Time Series Model with Annual Component Applied to Operation Planning of Hydrothermal Systems, **16th International Conference on Probabilistic Methods Applied to Power Systems – PMAPS**, on-line, July 2020. WOOD, S.N. Generalized Additive Models: An Introduction with R, Chapman & Hall/CRC, 2006.

AN APPROACH TO REPRESENTING THE UNCERTAINTIES OF THE WIND GENERATION SOURCE IN THE NEWAVE MODEL

Abstract

This work proposes a methodological approach to model the uncertainty of monthly wind power production in the operation dispatch problem, adhering to the solution algorithm used (stochastic dual dynamic programming - SDDP) in the NEWAVE model, allowing to obtain results with the appropriate accuracy and within of a reasonable computational time, when applied to large interconnected and hydro-dominated systems, as is the case of the Brazilian system. The proposed approach is composed of four stages: (i) statistical clustering of wind regimes; (ii) evaluation of monthly transfer functions between winds and wind power production, based on monthly probabilistic power curves; (iii) generation of synthetic wind speed scenarios / wind power production correlated to inflows to hydropower reservoirs; and (iv) representation of monthly wind power production in the SDDP algorithm implemented in the NEWAVE model. A brief description of each of these steps is presented, and preliminary results are discussed. Initial results of applying the steps of the proposed methodology to real configurations of the Brazilian interconnected system point to its effectiveness, at the same time shedding light on possible improvements and new developments.

Keywords: *Expansion and operation planning; Wind power; Multivariate statistical techniques; Time series; Generation of synthetic sequences; Stochastic optimization.*