

Título: Uma Abordagem para a Representação das Incertezas da Fonte de Geração Eólica no Planejamento da Operação de Longo, Médio e Curto Prazos.**Nº de Páginas: 34****Nº de Anexos: 0****Dados do Cliente:**

Comissão Permanente Análise de Metodologias e Programas Computacionais do Setor Elétrico - GT-Metodologia/CPAMP

E-mail: gtmet.cpamp@ccee.org.br

Departamento: Departamento de Otimização Energética e Meio Ambiente - DEA**Centro de Lucro:** C230000101**Centro de Custo:** C203000006/C203000022**Resumo:** Este relatório descreve uma metodologia para a representação das incertezas da produção eólica, 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 tratável computacionalmente, assim como obter resultados com a precisão adequada, 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; (iii) geração de séries sintéticas de ventos/produção eólica; e (iv) obtenção da produção eólica mensal no algoritmo de PDDE do modelo NEWAVE.**Autores:**Maria Elvira Piñeiro Maceira – Cepel
Albert Cordeiro Geber de Melo – Cepel
José Francisco Moreira Pessanha – Cepel
Cristiane Barbosa da Cruz Oliveira – Cepel
Victor Andrade de Almeida – Cepel
Thatiana Conceição Justino - Cepel**Palavras-Chave:**

Planejamento da operação de longo/médio prazo; modelos de séries temporais; incerteza na geração eólica; otimização estocástica; função de transferência; geração de séries sintéticas.

Classificação de acesso: Público**Aprovação e data de emissão****André Luiz Diniz Souto Lima**
Chefe do Departamento de Otimização Energética e Meio Ambiente**Tel.:** 21-2598-6046**E-mail:** diniz@cepel.br**Aprovação e data de emissão****Mauricio Barreto Lisboa****Diretor de Pesquisa, Desenvolvimento e Inovação**

Uma Abordagem para a Representação das Incertezas da Fonte de Geração Eólica no Planejamento da Operação de Longo, Médio e Curto Prazos

Relatório de Projeto CEPEL 2667/2020

PROJETO NEWAVE

Modelo de Planejamento da Operação de Sistemas Hidrotérmicos Interligados de Longo e Médio Prazos

PROJETO VENTOS

Sistema de Previsão de Ventos e Geração Eólica

PROJETO GEVAZP

Geração de Cenários Sintéticos de Energias e Vazões

Outubro/2020

newave@cepel.br

ventos@cepel.br

gevazp@cepel.br

Sumário

SUMÁRIO EXECUTIVO	4
1. INTRODUÇÃO	5
2. O MODELO DE PLANEJAMENTO DA OPERAÇÃO DE LONGO E MÉDIO PRAZOS	8
2.1 Estratégia de Solução	9
3. DESCRIÇÃO GERAL DA ABORDAGEM PROPOSTA	11
4. AGRUPAMENTO ESTATÍSTICO DOS REGIMES DE VENTOS	14
5. AVALIAÇÃO DAS FUNÇÕES DE TRANSFERÊNCIA MENSIS (FTMS) ENTRE VENTOS E PRODUÇÃO EÓLICA	19
6. GERAÇÃO DE SÉRIES SINTÉTICAS DE VENTOS/PRODUÇÃO EÓLICA	23
7. OBTENÇÃO DA PRODUÇÃO EÓLICA MENSAL NO ALGORITMO DE PDDE DO MODELO NEWAVE	29
8. CONCLUSÕES E TRABALHOS FUTUROS	30
9. AGRADECIMENTOS	31
10. REFERÊNCIAS	32

SUMÁRIO EXECUTIVO

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, nos modelos de planejamento da operação de longo, médio e curto prazos do CEPEL, que seja aderente aos algoritmos de solução utilizados em cada modelo.

Motivação: A matriz elétrica brasileira apresenta uma alta participação de fontes renováveis, respondendo por mais de 80% do consumo de eletricidade do país, e com preponderância da tecnologia hidroelétrica. As fontes intermitentes, especialmente a eólica, tem experimentado um crescimento acelerado, mundialmente e também no Brasil, e em 2019 já contava com mais de 14.968 MW instalados em mais de 600 parques, localizados principalmente nas regiões Nordeste e Sul. De acordo com o Plano Decenal de Expansão de Energia 2020-2029, estima-se que em 2029 a capacidade instalada da fonte eólica aumente mais de 2,5 vezes, atingindo 39,500 MW, enquanto a solar crescerá cerca de 5 vezes. Apesar das vantagens, 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. Assim, torna-se fundamental o desenvolvimento e o aprimoramento de metodologias para representar as incertezas das fontes renováveis intermitentes – eólica e solar - nos modelos de planejamento da operação de longo, médio e curto prazos.

Objetivos: Descrever uma metodologia para a representação das incertezas da produção eólica, 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 tratável computacionalmente. Essa mesma metodologia também poderá ser aplicada ao modelo DECOMP a partir do segundo mês do seu horizonte.

Metodologia: A metodologia está sendo desenvolvida e implementada em estágios, iniciando com a incorporação da incerteza da fonte eólica e pelo modelo NEWAVE. A abordagem proposta é multidisciplinar, conjugando o emprego de diversas técnicas, dentre as quais: análise exploratória de dados - análise estatística de agrupamentos e métodos estatísticos multivariados; 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); processos estocásticos e séries temporais; e otimização estocástica - programação dinâmica dual estocástica (PDDE). A metodologia proposta 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; (iii) geração de séries sintéticas de ventos/produção eólica; e (iv) obtenção da produção eólica mensal no algoritmo de PDDE do modelo NEWAVE.

Resultados: Aprimoramento da modelagem das fontes intermitentes nos modelos de planejamento da operação de longo, médio e curto prazos. Desenvolvimento e implementação de uma estratégia de consideração de cenários de produção eólica que seja aderente ao algoritmo de solução utilizado no modelo NEWAVE (a PDDE). E ainda que permita manter o problema estocástico e de grande porte, resolvido pelos modelos, tratável computacionalmente, assim como a obtenção de resultados com a precisão adequada, quando aplicada a sistemas interligados de grande porte, com predominância hidroelétrica, como é o caso do sistema brasileiro.

1. INTRODUÇÃO

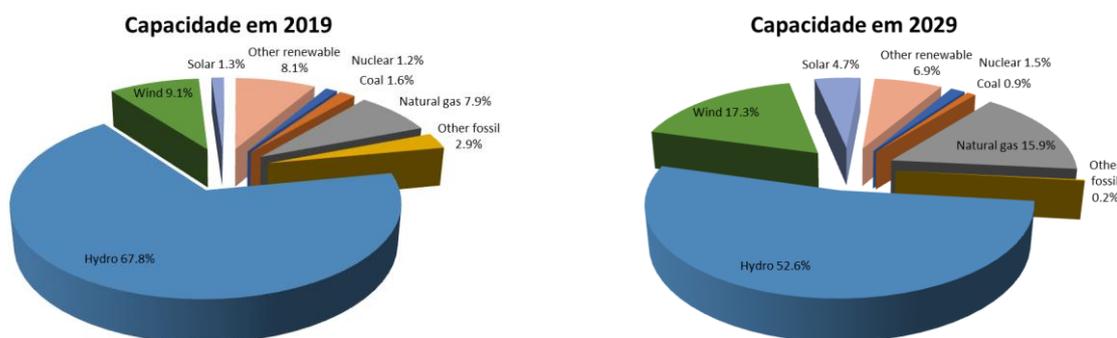
O planejamento da expansão e da operação em sistemas hidrotérmicos tem sido realizado, em muitos países, desagregando os problemas de planejamento nos horizontes de longo, médio e curto prazos, de acordo com características específicas de seus sistemas de energia [1]. No Brasil, o problema de planejamento 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 [2][3].

Um dos modelos desta cadeia para as atividades de planejamento da expansão e da operação energética de longo e médio prazos do sistema interligado é o modelo NEWAVE [4][5] 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 [6][7]; 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. Já no planejamento da operação de curto prazo é empregado o modelo DECOMP [8], que, desde 2001, vem sendo utilizado em estudos oficiais pelo ONS e pelo MAE / CCEE. O modelo DECOMP emprega a técnica de programação dinâmica dual [9], e vem sendo empregado, com horizonte de dois meses. O primeiro mês é determinístico, discretizado em semanas e o segundo mês é estocástico e representado em base mensal. A função de custo futuro calculada no terceiro mês do horizonte de planejamento pelo modelo NEWAVE é acoplada ao final do segundo mês do modelo DECOMP, obtendo-se o despacho de operação do SIN (geração hidroelétrica, geração térmica, intercâmbios entre os subsistemas e custos marginais de operação - CMO) para a semana operativa.

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. 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 [10], 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.

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 horária 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.



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

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

Atualmente, em consonância com a Resolução ANEEL Nº 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 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 relatório é descrever o estágio atual de 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 de algumas etapas da abordagem proposta também são apresentados e discutidos.

2. O MODELO DE PLANEJAMENTO DA OPERAÇÃO DE LONGO E MÉDIO PRAZOS

No modelo NEWAVE, 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 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) [11][12] 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 [13][14]. 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 afluências - vazões quando a representação do sistema se dá a UHEs ou energia quando a representação do sistema se dá a REEs - são modeladas por um modelo auto-regressivo periódico da ordem p , PAR(p) [15][16] empregado no modelo GEVAZP [17][18][19], a fim de considerar as correlações temporais das afluências, 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). 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. [3][5].

Além de fornecer a política de operação por meio da função de custo futuro para o modelo de programação de curto prazo – DECOMP, 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).

$$\begin{aligned}
 \min_{x_1} c_1 x_1 + E[\phi_2(x_1, \xi_2)] \quad (1a) \quad & \phi_t(x_{t-1}, \xi_t) = \min_{x_t} c_t x_t + E_{\xi_{t+1}|\xi_t, \dots, \xi_{t+1-p}} [\phi_{t+1}(x_t, \xi_{t+1})] \quad (1b) \\
 \text{s. a. } g_1(x_1) = b_1 \quad & \text{s. a. } g_t(x_t) = b_t(x_{t-1}, \xi_{t-j}, j=1, \dots, p) \\
 x_1 \in X \quad & x_t \in X
 \end{aligned}$$

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

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

2.1 Estratégia de Solução

Dado que computacionalmente 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 [9] – 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 n_{sim} 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 um 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 (t,s) do período t , 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 [5].

A função objetivo (2a) é composta pelos custos de geração térmica, custo de déficit de energia e eventuais violações de restrições operativas (vazão mínima, desvio de água, geração hidráulica mínima, etc.) onde β é a taxa de desconto mensal e CF representa o custo futuro. As principais restrições em cada período são a demanda energética (2b) e o balanço hídrico (2c) em cada submercado e REE. Os limites das variáveis de decisão x_t são apresentados em (2d).

$$\begin{aligned} \phi^{t,s,\omega} &= \phi_t(x_t^{s,\omega}, \xi_t^\omega) \\ &= \min c_{t,gt} x_{gt}^{t,s,\omega} + c_{t,def} x_{def}^{t,s,\omega} + c_{t,viol} x_{viol}^{t,s,\omega} + \left(\frac{1}{1+\beta} \right) CF^t(x_v^{t,s,\omega}, \xi^{t-p,\omega,p=0,\dots,P-1}) \end{aligned} \quad (2a)$$

s. a.

$$x_{gt}^{t,s,\omega} + x_{gh}^{t,s,\omega} + x_{geol}^t + x_{gsol}^t + x_{def}^{t,s,\omega} = d^t \quad (2b)$$

$$x_v^{t,s,\omega} - x_v^{t-1,s,\omega} + x_q^{t,s,\omega} = \xi^{t,\omega}, \quad (2c)$$

$$\underline{x}^t \leq x^{t,s,\omega} \leq \overline{x}^t \quad (2d)$$

São apresentadas a seguir as equações de atendimento à demanda, balanço hídrico, função de produção e função de custo futuro, para cada estágio e REE.

$$\sum_{k \in NREE} GH_{t,c,k} + GFIOl_{t,c,k} + \sum_{j \in NUT_m} GT_{t,c,j} + \sum_{i \in \theta_k} (F_{t,c}(i,k) - F_{t,c}(k,i)) + DEF_{t,c,m} - EXC_{t,c,m} = DLIQ_{t,c,m} \quad (3a) \text{ [Equação de atendimento à demanda para cada patamar de carga } c \text{ em cada submercado } m]$$

$$EA_{t+1,k} = FDIN_{t,k} EA_{t,k} + FC_{t,k} EC_{t,k} - GH_{t,c,k} - EVT_{t,k} - EVP_{t,k} - EDVC_{t,k} \quad (3b) \text{ [Equação de balanço hídrico]}$$

$$\alpha_{t+1} - \sum_{k \in NREE} \bar{\pi}_{EA_{1,t+1,k}} EA_{t+1,k} + \sum_{j=1}^p \bar{\pi}_{EAF_{1,j,t+1,k}} EAF_{t-j+1,k} \geq \bar{\delta}_{1,t+1} \quad (3c) \text{ [Inequações representando a Função de Custo Futuro]}$$

⋮

$$\alpha_{t+1} - \sum_{k \in NREE} \bar{\pi}_{EA_{q,t+1,k}} EA_{t+1,k} + \sum_{j=1}^p \bar{\pi}_{EAF_{q,j,t+1,k}} EAF_{t-j+1,k} \geq \bar{\delta}_{q,t+1}$$

$t = 1, \dots, T; k = 1, \dots, NREE; m = 1, \dots, NSUBM$

onde, para cada estágio t , tem-se: $GH_{t,c,k}$ geração hidráulica do REE k para o patamar de carga c ; $GFIOl_{t,c,k}$ geração fio d'água líquida do REE k para o patamar de carga c ; $GT_{t,c,j}$ geração termoelétrica da usina j para o patamar de carga c ; $F_{t,c}(i,k)$ intercâmbio do subsistema i para k no patamar de carga c ; $DEF_{t,c,m}$ corte de carga no subsistema m no patamar de carga c ; $EXC_{t,c,m}$ excesso de energia no subsistema m no patamar de carga c ; $DLIQ_{t,c,m}$ demanda líquida no subsistema m no patamar de carga c ; $EA_{t,k}$ energia armazenada no REE k ; $FDIN_{t,k}$ fator de correção devido à mudança de configuração no REE k ; $FC_{t,k}$ fator de correção da energia controlável; $EC_{t,k}$ energia controlável afluyente ao REE k ; $EVT_{t,k}$ energia vertida pelo REE k ; $EVP_{t,k}$ energia evaporada pelo REE k ; $EDVC_{t,k}$ energia desviado do REE k ; α_t valor esperado do custo total de operação do estágio t até T , também chamada função de custo futuro do estágio t ; $\bar{\pi}_{EA_{1,t+1,k}}$ multiplicador de Lagrange associado à energia armazenada no REE k ; $\bar{\pi}_{EAF_{1,t+1,k}}$ multiplicador de Lagrange associado à energia afluyente ao REE k ; $\bar{\delta}_{1,t+1}$ constante.

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 relatório é descrever uma 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 (UEEs) ou por meio de parques eólicos equivalentes (PEEs).

Para a obtenção dos PEEs, faz-se necessário efetuar a identificação e o agrupamento dos regimentos de ventos por meio de técnicas estatísticas, considerando diversas granularidades espaciais, e.g., parques eólicos individualizados (UEEs), 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 (UEEs/PEEs) e também as correlações espaciais entre UEEs/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 UEE 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 UEEs/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 auto-regressivos 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 UEEs 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 UEEs/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 [20]-[25]. 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 Sintegra; 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 [26].

Em resumo, a abordagem proposta é multidisciplinar, conjugando o emprego de diversas técnicas, dentre as quais: análise exploratória de dados - análise estatística de agrupamentos e métodos estatísticos multivariados [27]; 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) [28]; processos estocásticos e séries temporais [29][15][16]; otimização estocástica - programação dinâmica dual estocástica [6][7].

Por fim, a abordagem proposta 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;
- (iii) geração de séries sintéticas de ventos/produção eólica; e

(iv) obtenção da produção eólica mensal no algoritmo de PDDE do modelo NEWAVE.

Uma descrição sucinta de cada uma dessas etapas, assim como de resultados preliminares, será apresentada nas próximas seções.

4. AGRUPAMENTO ESTATÍSTICO DOS REGIMES DE VENTOS

Conforme apresentado na Seção 3, uma das etapas da abordagem metodológica proposta consiste na identificação e no agrupamento dos regimes de ventos nas diversas regiões do Brasil. Também foi justificada a necessidade de se utilizar dados oriundos de reanálises. 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. Para isso deve-se analisar as correlações entre as séries de velocidade de vento, i.e., entre as colunas da matriz de dados. Por exemplo, 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.

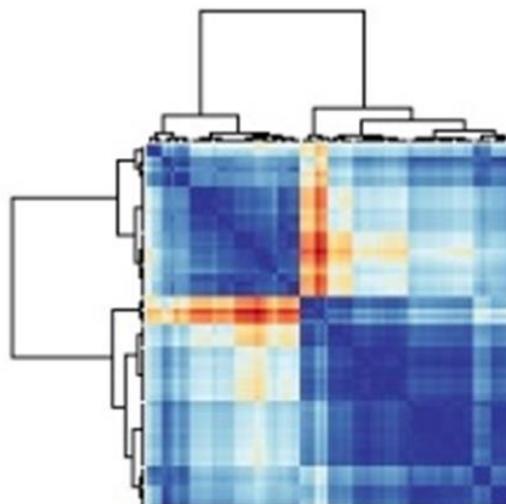


Figura 4.1 - Mapa de calor da matriz de correlações entre as médias mensais de 79 localidades com parques eólicos na região Nordeste [20]

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 com a aplicação da análise fatorial exploratória - AFE [27], seguida da análise de agrupamentos [27].

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 .

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:

$$x_i = l_{i1}F_1 + l_{i2}F_2 + \dots + l_{im}F_m + \varepsilon_i \quad \forall i=1,n \quad (4)$$

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

A partir da combinação linear em (4) e da premissa de independência entre F_1, F_2, \dots, F_m e $\varepsilon_i \forall i=1,n$ obtém-se a seguinte decomposição da matriz de covariâncias:

$$S = L^T L + \Phi \quad (5)$$

em que L é uma matriz de dimensões $n \times m$ em que cada linha guarda os pesos de cada parque eólico i nos m fatores latentes, i.e., cada linha é formada pelo elementos $l_{i1}, \dots, l_{im} \forall i=1,n$.

A parcela $L^T L$ é denominada por comunalidade e captura a contribuição dos m regimes de vento para a variabilidade da velocidade nos n parques eólicos. Já a parcela Φ é uma matriz diagonal, cujos elementos capturam a variabilidade da velocidade de vento em cada parque eólico que não é explicada pelos regimes de vento.

O ajuste do modelo em (4) pode ser realizada por meio do cálculo dos n autovetores e_1, e_2, \dots, e_n e respectivos autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ da matriz S . Adicionalmente, a soma dos autovalores é igual ao traço da matriz S , i.e., a variância total dos ventos nas n localidades com parques eólicos. A determinação do número de fatores latentes m e a formação da L consiste em encontrar um valor para m tal que $S \cong L^T L$.

Pelo Teorema da Decomposição Espectral [27] tem-se que:

$$S = \lambda_1 e_1 e_1^T + \lambda_2 e_2 e_2^T + \dots + \lambda_m e_m e_m^T \quad (6)$$

Dado que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, os primeiros autovalores concentram a maior parcela da variância total, então os primeiros termos da soma em (6) são os que mais contribuem para a formação da matriz S . Assim, para que $L^T L$ forneça uma boa aproximação da matriz S , o número de fatores m deve satisfazer a seguinte condição [27]:

$$100\%(\lambda_1 + \lambda_2 + \dots + \lambda_m)/(\lambda_1 + \lambda_2 + \dots + \lambda_n) \geq 80\% \quad (7)$$

Definido o número de fatores m , a matriz L pode ser gerada com base nos autovetores de S associados aos primeiros m autovalores:

$$L = [\sqrt{\lambda_1} e_1 \quad \sqrt{\lambda_2} e_2 \quad \dots \quad \sqrt{\lambda_m} e_m] \quad (8)$$

Se a condição em (7) é satisfeita com menos de três fatores ($m \leq 3$) pode-se gerar uma visualização das n localidades com parques eólicos em um sistema de m eixos fatoriais, um mapa que permite a rápida identificação de grupos de parques eólicos espacialmente correlacionados, conforme ilustrado na Figura 4.2.

A aplicação da AFE na matriz de correlações entre as 79 localidades permite obter uma representação gráfica, em 3 dimensões (Figura 4.2), da estrutura de correlações entre os regimes de vento que atingem as localidades. 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 representa um município com aproveitamento eólico. As distâncias entre os pontos na Figura 4.2 refletem as correlações entre os regimes de vento, sendo que pontos próximos indicam regimes correlacionados.

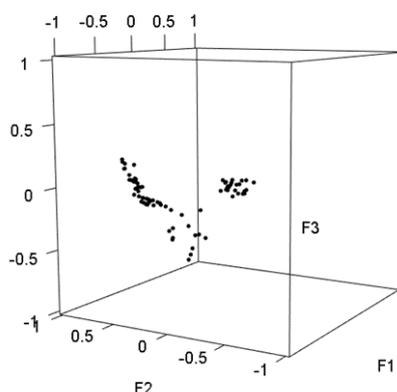


Figura 4.2 - Diagrama das localidades avaliadas nos três eixos fatoriais [20]

Por sua vez, a identificação de *clusters* de parques eólicos com regimes de vento correlacionados pode ser realizada por meio do algoritmo K-Means [27]. Para isso, os vetores com as coordenadas dos aproveitamentos eólicos nos eixos fatoriais devem ser empilhados em uma nova matriz de dados, neste caso uma matriz com três colunas.

A análise de agrupamentos - *cluster analysis* [27], [30]-[32] 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.

O K-Means classifica o conjunto de objetos em um número k de *clusters* especificado a priori. O critério utilizado pelo método consiste em minimizar a variabilidade dentro dos agrupamentos, expressa pela soma dos quadrados dos desvios entre as observações e o centroide do *cluster* no qual as observações foram alocadas (Within Sum of Squares - WSS):

$$WSS = \sum_{j=1}^n \left(\sum_{i, x_i \in C_j} \|x_i - c_j\|^2 \right) \quad (9)$$

em que x_i , $i=1, n$, é o i -ésimo objeto, C_j é o j -ésimo cluster, c_j é o centroide do j -ésimo cluster e n é o número de objetos a serem classificados.

Cada objeto pertence a apenas um dos k clusters, portanto, a classificação dos n objetos pode ser representada por uma matriz binária U de dimensões $n \times k$, onde $u_{ij} = 1$ se o i -ésimo objeto pertence ao j -ésimo cluster e $u_{ij} = 0$, caso contrário.

O método K-Means pode ser implementado segundo o algoritmo a seguir, onde a matriz U e os centroides dos clusters são obtidos de forma iterativa:

- 1) Inicialize os centroides dos clusters pelo sorteio de k objetos entre os n a serem classificados.
- 2) Determine a matriz U de acordo com o critério do centroide mais próximo, i.e., aloque cada objeto ao cluster com o centroide mais próximo.
- 3) Calcule o valor da função objetivo WSS em (9). Pare se o valor da função estiver abaixo de uma tolerância pré-especificada, se a melhoria em relação à iteração anterior for desprezível ou se o máximo de iterações for alcançado.
- 4) Atualize os centroides dos clusters e volte para o passo 2.

A classificação resultante dos aproveitamentos eólicos pelo método K-Means está apresentada na Figura 4.3.

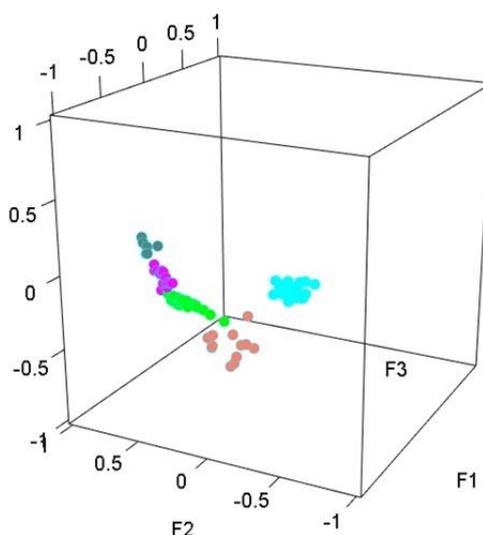


Figura 4.3 - Clusters de aproveitamentos eólicos [20]

Já na Figura 4.4 tem-se a representação espacial dos agrupamentos identificados na Figura 4.3. Vale destacar a reduzida superposição entre os agrupamentos.

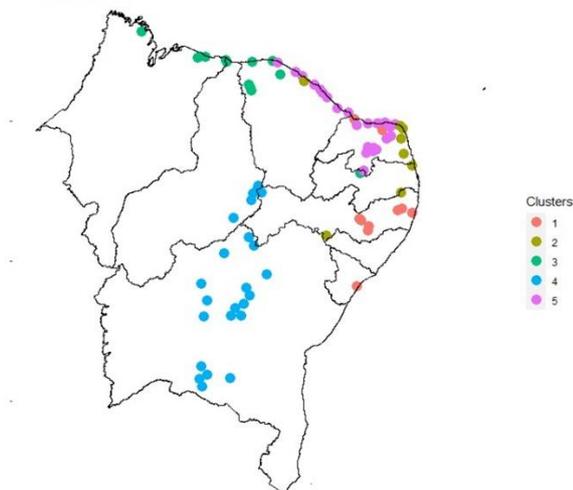


Figura 4.4 – Clusters de parques eólicos

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.

No presente trabalho de pesquisa, 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 MENSAS (FTMS) ENTRE VENTOS E PRODUÇÃO EÓLICA

As considerações anteriores apontam para a necessidade de ajustar funções de transferência capazes de descrever o relacionamento entre as médias mensais de velocidade de vento e de produção eólica mensal. O ajuste de tais funções 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 esta 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 (2018) passaram a ser disponibilizados diariamente pelo sistema Sintegre do ONS, até 168 horas à frente e em base semi-horária. Com isto, pode-se obter estimativas das *curvas de potência probabilísticas* em escala 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. Contudo, tais funções de transferência são fornecidas para grupos de aproveitamentos eólicos, cada um associado à uma subestação. A desagregação por parque eólico pode ser aproximada mediante fatores de participação, também fornecidos no Sintegre ONS, e que permitem quantificar a contribuição de cada parque eólico na subestação em que está conectado.

Apesar do período de disponibilização ser relativamente curto – pouco mais de um ano – 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.

Para ilustrar a factibilidade deste procedimento foram analisados dados dos PEEs de duas subestações disponibilizados no Sintegre para a região Nordeste: Extremos II 230kV – próxima ao litoral do Rio Grande do Norte, e Morro do Chapéu 69kV - no interior da Bahia. Adicionalmente, foram analisados dados disponibilizados pelo Sintegre para a subestação de Viamão 3 230 kV, localizada no Rio Grande do Sul.

Para cada uma das subestações consideradas, as Figuras 5.1 a 5.3 apresentam (na cor vermelha) as curvas de potência probabilísticas *horárias* (diagramas de dispersão) obtidas com os dados do Sintegre. Nestas mesmas Figuras 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.

Morro do Chapéu

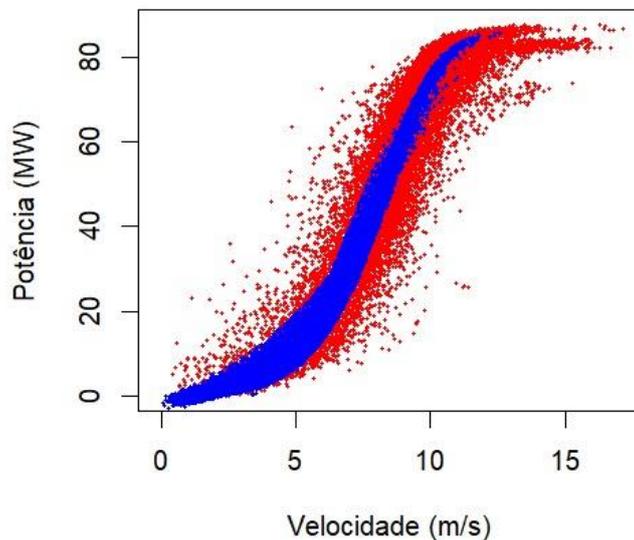


Figura 5.1 - Curvas de potência probabilísticas horárias - dados do Sintegre (vermelho) e procedimento elaborado usando o MERRA 2 (azul) - PEE da SE Morro do Chapéu 69 kV

Extremoz II

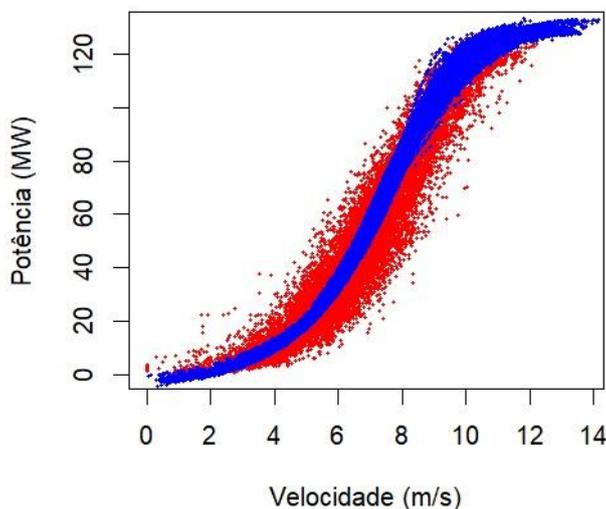


Figura 5.2 - Curvas de potência probabilísticas horárias - dados do Sintegre (vermelho) e procedimento elaborado usando o MERRA 2 (azul) - PEE da SE Extremoz II 230 kV

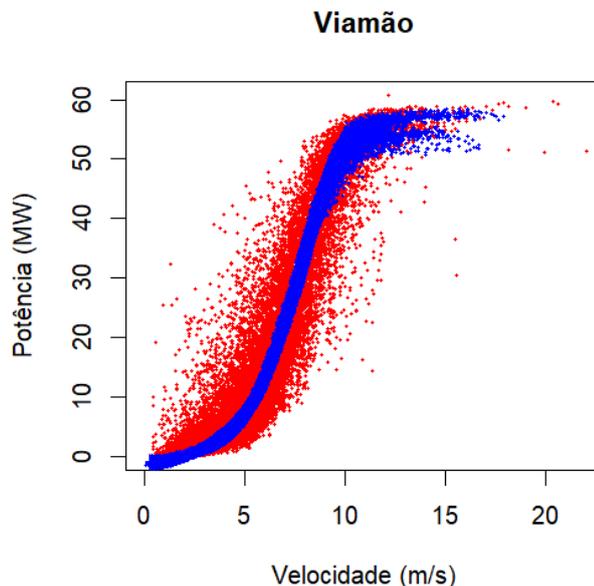


Figura 5.3 - Curvas de potência probabilísticas horárias - dados do Sintegre (vermelho) e procedimento elaborado usando o MERRA 2 (azul) – PEE da SE Viamão 230 kV

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 apresentadas nas Figuras 5.4 a 5.6.

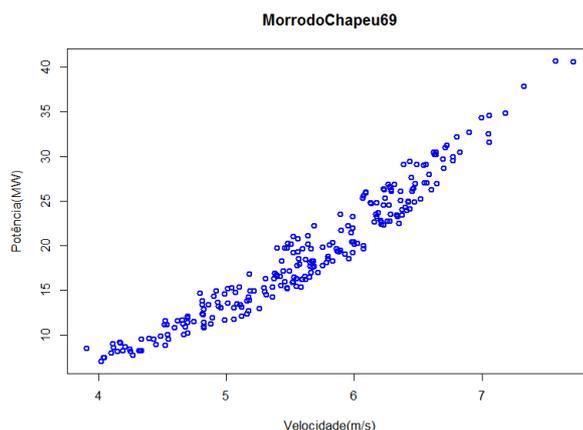


Figura 5.4 - Curva de potência probabilística mensal obtida pela aplicação do procedimento proposto aos dados de reanálise oriundos do MERRA 2 – PEE da SE Morro do Chapéu 69 kV

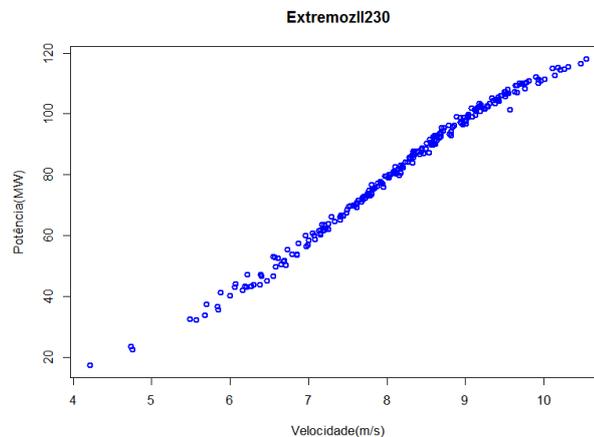


Figura 5.5 - Curvas de potência probabilísticas mensais obtida pela aplicação do procedimento proposto aos dados de reanálise oriundos do MERRA 2

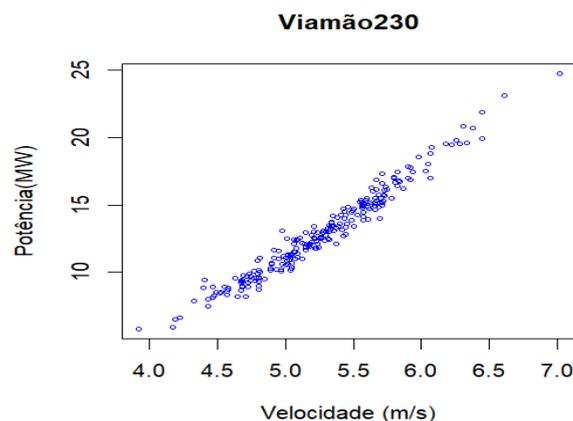


Figura 5.6 - Curvas de potência probabilísticas mensais obtida pela aplicação do procedimento proposto aos dados de reanálise oriundos do MERRA 2

Quando se compara as Figuras 5.1, 5.2 e 5.3 com as Figuras 5.4, 5.5 e 5.6, 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 produção eólica mensais incluirão os métodos estatísticos para análise de regressão, desde os tradicionais modelos de regressão linear até abordagens mais flexíveis, como os modelos aditivos generalizados (generalized additive models - GAM) [28].

6. GERAÇÃO DE SÉRIES SINTÉTICAS DE VENTOS/PRODUÇÃO EÓLICA

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. São gerados séries sintéticas para a simulação *forward*, 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 dos parâmetros, ou todos, variam dentro do ano, são conhecidos como modelos auto-regressivos com parâmetros periódicos, e são denotados por modelos PAR(p).

O modelo auto-regressivo periódico, caracterizado por PAR(p1, ..., ps), pode ser escrito como:

$$\phi^m(B) \left(\frac{Z_t - \mu_m}{\sigma_m} \right) = a_t \quad (10)$$

onde:

Z_t é a variável aleatória do processo estocástico com s períodos sazonais e, nesta aplicação, representa a afluição ao REE ou 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 sazonal período 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 [15][16].

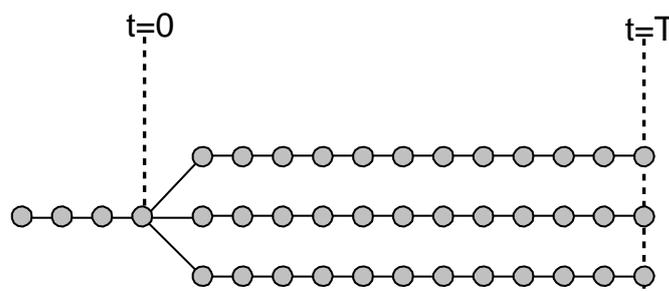
A função de auto-correlação $\rho^m(k)$ de Z_t é obtida multiplicando-se ambos os lados da equação (4) por $\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right)$ e aplicando o valor esperado a todos os termos desta equação. Ela satisfaz a equação de diferenças (sistema de equações de Yule-Walker). 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 respectivas 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 [33], os autores definem 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 [33]. 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 [17]-[19].

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 [18]. O método de amostragem seletiva consiste na aplicação de técnicas de agrupamento no 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. A Figura 6.1 ilustra o esquema de geração de cenários sintéticos em um período de tempo com T estágios.



Fonte: Autores

Figura 6.1 - Esquema de geração de cenários sintéticos

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

- ✓ 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 usina hidroelétrica ou reservatório equivalente de energia, 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 normalmente distribuídos, 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 afluições 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:

$$e_t = D a_t \quad (11)$$

A matriz D pode ser estimada por:

$$DD^T = \hat{U} \quad (12)$$

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 afluições: 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 afluições 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 de afluições são obtidos usando-se a expressão (4);
- ✓ A soma das afluições incrementais ao longo da cascata de UHEs, em cada período e cenário, é realizada para obter as afluições totais.

Maiores detalhes da metodologia para geração de cenários sintéticos são obtidos em [17][18].

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

- A estrutura de correlação temporal, porventura verificada no processo estocástico ventos médios mensais (m/s) em algum UEE 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;
- Será representada a correlação espacial verificada entre os processos estocásticos ventos e afluições aos REEs ou UHEs. Esta estatística apresenta-se elevada em muitos dos meses do ano, conforme apresentado na Tabela 6.1 que mostra as correlações entre PEEs que compõem as subestações Extremoz II, Morro do Chapéu Viamao 3 e alguns dos REEs considerados no PMO. Estes resultados são seguidos por outros parques eólicos do país, daí a importância de ser levado em consideração;

Tabela 6.1 - Correlação espacial entre PEEs das subestações Morro do Chapéu, Bom Jardim e ExtremozII e alguns REEs do SIN.

<i>Extremoz II</i>		Sudeste	Paraná	Nordeste	Norte
	Fevereiro	-0,38	-0,45	-0,36	-0,36
	Agosto	-0,54	-0,60	-0,36	-0,30
	Setembro	-0,70	-0,89	-0,56	-0,69
	Outubro	-0,59	-0,69	-0,49	-0,75
<i>Morro do Chapéu</i>		Sudeste	Paraná	Nordeste	Norte
	Maio	0,47	0,31	0,49	0,23
	Junho	-0,45	-0,62	-0,27	-0,39
	Setembro	-0,33	-0,44	-0,22	-0,31
	Dezembro	-0,52	0,20	-0,28	-0,52
<i>Viamao3</i>		Paraná	Belo Monte	Sul	Iguaçu
	Fevereiro	0,51	0,29	0,33	0,37
	Abril	-0,32	-0,57	0,26	0,40
	Maio	0,16	-0,32	0,35	0,32
	Outubro	-0,38	-0,30	0,36	0,31

A proposta deste trabalho é adotar um *modelo integrado de geração de séries sintéticas de afluências 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:

$$\left(\frac{V_t - \mu_m^v}{\sigma_m^v}\right) = \text{componente explicativa} + a_t \tag{13}$$

Reescrevendo:

$$V_t = \text{componente explicativa} + \sigma_m^v a_t \tag{14}$$

A *componente explicativa* pode ser a próprio média do processo estocástico vento do período sazonal m correspondente ao estágio t, μ_m^v , correspondendo ao modelo *naive* ou ainda, por exemplo, conter uma parcela relacionada às afluências do próprio estágio t.

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

- ✓ Obter as afluências incrementais mensais históricas para cada usina hidrelétrica ou reservatório equivalente de energia;
- ✓ Escolher a ordem do modelo AR de cada período sazonal para cada usina hidroelétrica ou reservatório equivalente de energia, 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 correspondente de mesma cardinalidade para os 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 **afluências 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 expressões (5) e (6). A matriz \hat{U} passa a ser, em princípio, a estimativa da **matriz de covariâncias dos resíduos a_t , correspondentes às afluências e ventos**. Na prática, o comportamento dos resíduos não segue o comportamento das afluências, nem dos ventos, nem entre afluências e ventos: os resíduos não são espacialmente correlacionados. No entanto, a fim de se preservar as dependências espaciais entre os REEs/UHEs, **entre os UEEs/PEEs, e entre REEs/UHEs e UEEs/PEEs, utiliza-se a correlação espacial entre as afluências, entre os ventos e entre as afluências 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 UEEs/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, porém se ajustará a distribuição Lognormal e/ou a distribuição Weibull três parâmetros de tal forma a representar tanto assimetrias positivas quanto negativas**. Esta tabela apresenta também as principais estatísticas dos PEEs Morro do Chapéu e ExtremozII associados às subestações de mesmo nome, ambas na região Nordeste e, do PEE Bom Jardim associados à subestação Bom Jardim na região Sul;
- ✓ Os cenários sintéticos de afluências e ventos são obtidos usando-se as expressões (4) e (8) respectivamente;
- ✓ A soma das afluências incrementais ao longo da cascata de UHEs, em cada período e cenário, é realizada para obter as afluências totais.

Tabela 6.2 - Correlação espacial entre PEEs das subestações Morro do Chapéu, Bom Jardim e ExtremozII e alguns REEs do SIN.

		Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
<i>Extremoz II</i>													
	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
	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
<i>Morro do Chapéu</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
	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
<i>Viamao3</i>													
	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
	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

Conforme evidenciado nas Tabelas 6.1 e 6.2, 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 alguns meses, em geral negativas, diferentemente das assimetrias dos REEs, que são sempre positivas.

Durante o decorrer da pesquisa, estas estatísticas serão calculadas para outros UEEs e PEEs, especialmente para aqueles resultantes da agregação estatística dos regimes de ventos obtidos na etapa 2 da abordagem proposta, e orientarão a geração das séries sintéticas de produção de velocidades de ventos.

7. OBTENÇÃO DA PRODUÇÃO EÓLICA MENSAL NO ALGORITMO DE PDDE DO MODELO NEWAVE

No caso da geração de séries sintéticas de ventos mensais, é necessário a obtenção da correspondente produção eólica de cada parque eólico individual (UEE) ou equivalente (PEE) – obtido conforme descrito na Seção 5 – para ser considerada no problema de despacho mensal da operação do modelo NEWAVE. A produção eólica será obtida por meio das 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, é representá-las no problema de despacho como uma fonte de geração disponível, porém com custo de operação zero.

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; (iii) geração de séries sintéticas de ventos/produção eólica; e (iv) obtenção da produção eólica mensal no algoritmo de 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.

Dados o caráter multidisciplinar desta pesquisa e a complexidade do tema, as etapas da abordagem proposta estão sendo desenvolvidas e aprimoradas de forma simultânea e paulatina, com retro-alimentação a partir dos resultados que estão sendo obtidos.

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

9. AGRADECIMENTOS

Os autores agradecem ao apoio dos estagiários de Engenharia Elétrica da UERJ, Juliane Alves Celestino dos Santos e Felipe Botelho Barbosa.

10. REFERÊNCIAS

- [1] 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>.
- [2] 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.
- [3] 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.
- [4] 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, Glasgow, Scotland, 2008.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] BIRGE, J.R. Decomposition and partitioning methods for multistage stochastic linear programs, Operations Research. Vol. 33, no. 5, 989-1007, 1985.
- [10] MME - MINISTÉRIO DE MINAS E ENERGIA/EPE - EMPRESA DE PESQUISA ENERGÉTICA, Plano Decenal de Expansão de Energia 2029. Relatório Técnico, Maio 2020.
- [11] 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.
- [12] 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.

- [13] 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.
- [14] 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.
- [15] SALAS, J.D.; DELLEUR, J.W.; YEVJEVICH, V.; LANE, W.L Applied Modeling of Hydrologic Time Series, Water Resources Publications, 1980.
- [16] FRANCES, P.H.; PAAP, R. Periodic Time Series Models, Oxford University Press, 2006.
- [17] 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.
- [18] 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.
- [19] PENNA, D.D.J., MACEIRA, M.E.P., DAMÁZIO, J.M., TREISTMAN, F., ARAÚJO, H.S., "Manual de Referência do Modelo GEVAZP", Relatório Técnico CEPEL nº 13955/18.
- [20] PESSANHA, J.F.M.; ALMEIDA, V.A.; MELO, A.G.C. 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.
- [21] PESSANHA, J.F.M.; MELO, A.C.G.; BESSA, R.J.G.S.B.; ALMEIDA, V.A. Uma metodologia para geração de cenários de produção eólica compatíveis com as correlações espaciais entre os regimes de ventos, XXV Seminário Nacional de Produção e Transmissão de Energia Elétrica, Belo Horizonte, 2019.
- [22] PESSANHA, J.F.M.; MELO, A.C.G.; ALMEIDA, V.A. "Impact of Wind Speed Correlations on Probabilistic Power Flow by using the Nataf Transformation", IEEE Probabilistic Methods Applied to Power Systems – PMAPS 2018, Boise, Idaho, USA, 24-28 June, 2018.
- [23] PESSANHA, J.F.M.; MELO, A.C.G.; BESSA, R.J.G.S.B. Transformação de Nataf e estimação não paramétrica de densidades condicionais na modelagem da geração eólica no problema de fluxo de potência probabilístico, XXII Congresso Brasileiro de Automática, João Pessoa, 2018.
- [24] PESSANHA, J.F.M.; CASTELLANI, V.; JUSTINO, T.C.; PENNA, D.D.J.; MACEIRA, M.E.P. Uma metodologia para filtragem de medições anemométricas, Learning and Nonlinear Models, v. 10, p. 90-98, 2012.
- [25] 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, Curitiba, 2007.
- [26] STAFFELL, I., PFENNINGER, S. (2016). Using Bias-Corrected Reanalysis to Simulate Current and Future Wind Power Output. Energy 114, pp. 1224-1239.

-
- [27] JOHNSON, R. A. and WICHERN, D. W. Applied Multivariate Analysis, Fourth Edition, Prentice Hall, New Jersey, 1998.
- [28] WOOD, S.N. Generalized Additive Models: An Introduction with R, Chapman & Hall/CRC, 2006.
- [29] BOX, G.E.P.; JENKINS, G.M. Time series analysis forecasting and control. San Francisco: Holden-Day, 1976.
- [30] 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 10349/2018, Agosto de 2018.
- [31] 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.
- [32] 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.
- [33] HIPPEL, W.H., MCLEOD, A.I. Time Series Modeling of Water Resources and Environmental System. Elsevier. 1994.