

Modelos para Representação de Geração Eólica em Fluxo de Potência Ótimo

Ana F. P. da Costa* Marcos T. B. de Oliveira**
 Jhonatan S. Ferreira*** Arthur N. de Paula****
 Edimar J. de Oliveira† Julio C. S. de Souza‡
 Leonardo W. de Oliveira§

* Faculdade de Engenharia Elétrica, Universidade Federal de Juiz de Fora, MG, (e-mail: ana.procopio@engenharia.ufjf.br).

** Faculdade de Engenharia Elétrica, Universidade Federal de Juiz de Fora, MG, (e-mail: oliveira.marcos@engenharia.ufjf.br).

*** Faculdade de Engenharia Elétrica, Universidade Federal de Juiz de Fora, MG, (e-mail: jhonatan.sanches@engenharia.ufjf.br).

**** Faculdade de Engenharia Elétrica, Universidade Federal de Juiz de Fora, MG, (e-mail: arthur.paula@engenharia.ufjf.br).

† Faculdade de Engenharia Elétrica, Universidade Federal de Juiz de Fora, MG, (e-mail: edimar.oliveira@ufjf.edu.br).

‡ Departamento de Engenharia Elétrica, Universidade Federal Fluminense, RJ, (e-mail: julio@ic.uff.br).

§ Faculdade de Engenharia Elétrica, Universidade Federal de Juiz de Fora, MG, (e-mail: leonardo.willer@ufjf.edu.br).

Abstract: This article aims to evaluate the performance of models representative of the short-term uncertainties associated with wind generation in studies of optimal power flow (FPO). Wind generation is represented by time series of active power grouped from three different clustering techniques: k-means, ISODATA and hierarchical methods. These techniques aim to reduce the number of states necessary to model the behavior of the wind, without losing the representativeness of the original series, in order to reduce the computational effort in studies of FPO with the inclusion of wind generation. In order to verify the performance of each technique, simulations were performed in the IEEE 118 bus system. Monte Carlo Simulation, with high computational cost, is used as a benchmark.

Resumo: Este artigo tem como objetivo avaliar o desempenho de modelos representativos das incertezas de curto prazo associadas à geração eólica nos estudos do fluxo de potência ótimo (FPO). A geração eólica é representada por séries temporais de potência ativa agrupadas a partir de três diferentes técnicas de clusterização: k-means, ISODATA e métodos hierárquicos. Essas técnicas têm como objetivo reduzir a quantidade de estados necessários para modelar o comportamento do vento, sem que haja perda de representatividade da série original, a fim de reduzir o esforço computacional nos estudos do FPO com inclusão da geração eólica. A fim de verificar o desempenho de cada técnica, foram realizadas simulações no sistema IEEE 118 barras. A Simulação Monte Carlo, com alto custo computacional, é utilizada como *benchmark*.

Keywords: Wind Generation; Clustering; Optimal Power Flow; Monte Carlo Simulation; Minimal Losses

Palavras-chaves: Geração Eólica; Agrupamento de Dados; Fluxo de Potência Ótimo; Simulação Monte Carlo; Mínimas Perdas.

1. INTRODUÇÃO

Os investimentos em fontes renováveis de geração de energia elétrica têm aumentado no mundo todo. No Brasil, o ritmo de crescimento da energia eólica nos últimos anos tem sido expressivo, atingindo em 2019 uma capacidade de geração instalada de 15 GW, distribuída em cerca de 600 parques eólicos (Operador Nacional do Sistema Elétrico, 2019). Como consequência, a geração eólica é a segunda principal fonte energética do país, representando 9,2%

da matriz energética, estando atrás apenas da energia hidráulica. A partir desse cenário, estudos para analisar os possíveis impactos do comportamento dos geradores eólicos no sistema elétrico de potência se tornam cada vez mais necessários.

Em (Souza, 2008) são apresentados estudos para a identificação de diferentes tipos de agrupamento de dados, a fim de determinar um modelo reduzido que represente o comportamento do vento a partir de série históricas. Além da técnica de k-means, são consideradas as técnicas de c-

means e o Mapa Auto Organizável de Kohonen para a análise.

Em (Malta, 2009) são apresentados estudos de séries temporais de ventos com a finalidade de avaliar a viabilidade de instalação e a correlação entre usinas eólicas. O trabalho apresenta análises estatísticas para a geração de novas séries, baseando-se na análise de correlação. Essas novas séries são obtidas através do agrupamento de dados realizados com a técnica de k-means.

Embora exista uma grande quantidade de diferentes métodos de agrupamentos na área de reconhecimento de padrões (Xu and Wunsch, 2005), a maioria dos *softwares* voltados para o processamento digital de imagens de sensoriamento remoto realiza a classificação não-supervisionada baseada em métodos de agrupamentos particionais, como o k-means e o ISODATA (Iterative Self-Organizing Data Analysis Technique).

Adedeji et al. (2020) apresenta um estudo de caso para avaliar a capacidade eólica de uma região da África utilizando métodos de clusterização baseados em fuzzy-c-means que se mostraram eficazes para representar a geração eólica, sendo que este é uma versão do método k-means que será avaliada no presente artigo.

O trabalho realizado por Lin et al. (2018) apresenta uma comparação entre métodos estocásticos de fluxo de potência (Método Cumulante de Ajuste do Modelo, Método Cumulante de Ajuste de Fórmula e Simulação de Monte Carlo) considerando a correlação e observando a eficiência, precisão e complexidade de tempo de cada um.

Deng et al. (2019) propôs um novo método de fluxo de potência probabilístico com inserção de geração eólica por meio da combinação de uma técnica aprimorada de agrupamento k-means e o método cumulante tradicional. Esse novo método foi comparado como Simulação de Monte Carlo, Método de Estimativa de Pontos e Hiper cubo.

No trabalho apresentado em (Baringo and Conejo, 2013), são propostas duas metodologias alternativas para gerar com eficiência a demanda elétrica e os cenários de geração eólica, usados como dados de entrada para problemas de investimento. As metodologias são baseadas nas curvas de carga, duração do vento e no agrupamento k-means. Ambas as técnicas se mostraram eficientes mas, a técnica k-means permite representar a correlação entre carga e produção de energia eólica em locais diferentes, o que implica comparativamente em uma maior precisão.

Em (Wogrin et al., 2019) são comparadas duas metodologias de agregação de período de tempo de ponta para modelos de sistemas de potência que consideram tecnologias renováveis e de armazenamento: o agrupamento cronológico de período de tempo e a abordagem aprimorada do período representativo. Tais metodologias são usadas para reduzir o tempo computacional de modelos de otimização altamente complexos sem comprometer a qualidade dos resultados. O agrupamento cronológico de período de tempo torna o modelo matemático mais realístico. Já a abordagem aprimorada do período representativo é computacionalmente mais eficiente. O estudo conclui que a combinação das duas metodologias pode ser promissora.

Em (Baringo and Baringo, 2018) foi proposta uma abordagem de otimização robusta adaptativa estocástica para o problema de planejamento de expansão de geração e transmissão. Com o objetivo de reduzir os custos totais. Observa-se que o custo de construção de novas linhas de transmissão é relativamente menor do que o custo de construção de novas unidades geradoras. Os resultados apontam que o investimento em capacidade estocástica geralmente aumenta o valor do orçamento de incerteza de geração, enquanto o investimento em capacidade convencional aumenta normalmente o valor do orçamento de incerteza de demanda.

Dehghan et al. (2016) apresenta uma estrutura de planejamento de expansão de sistemas elétricos de potência robusto juntando a programação estocástica e a otimização robusta, modelando a incerteza da demanda, geração de energia eólica e disponibilidade de unidades e linhas. A técnica de agrupamento k-means é usada neste artigo para construir um conjunto de incertezas referentes às demandas correlacionadas e às gerações de energia eólica. A ferramenta de planejamento robusta com restrição de confiabilidade proposta é capaz de encontrar uma solução ideal, imunizada contra a pior realização possível de demandas incertas e geração de energia eólica, satisfazendo uma restrição probabilística de confiabilidade.

Seguindo essa linha de pesquisa para definir métodos de agrupamento de dados que apresentem resultados de boa qualidade em tempo computacional reduzido, a contribuição do presente trabalho consiste na realização de uma análise de desempenho de três técnicas de agrupamentos de dados: k-means, ISODATA e algoritmos Hierárquicos. Os métodos de agrupamento são utilizados para gerar cenários de despacho de geração eólica a curto prazo baseados em séries históricas. O objetivo do trabalho consiste em identificar o modelo de agrupamento mais representativo da geração, capaz de reproduzir seu comportamento intermitente em análises de fluxo de potência ótimo (FPO) com a finalidade de ter mínimas perdas elétricas no sistema com reduzido esforço computacional.

2. METODOLOGIA

A metodologia proposta consiste na inserção de cenários de despacho eólico no problema de FPO a fim de verificar o comportamento do sistema elétrico de potência. São considerados três algoritmos de agrupamento de dados: k-means, ISODATA e Algoritmos Hierárquicos (AHI). A simulação de Monte Carlo (SMC) é utilizada como *benchmark* para comparação com os métodos aglomerativos uma vez que avalia uma grande quantidade de cenários ao custo de elevado tempo computacional.

2.1 Fluxo de Potência Ótimo

Neste trabalho, a rede AC é modelada. Assim, o FPO consiste em um problema de programação não-linear que busca a minimização de perdas elétricas no sistema elétrico de potência (SEP). A função objetivo deste problema e as restrições consideradas estão representadas em (1) - (7) .

$$\min \sum_{w \in W} \text{perdas}_w \quad (1)$$

sujeito a:

$$PG_{gk} + PGW_{jk,w} - PL_k - \sum_{m \in \Omega_k} p_{km} = 0, \forall k \in B \quad (2)$$

$$QG_{gk} + QGW_{jk,w} - QL_k - \sum_{m \in \Omega_k} q_{km} = 0, \forall k \in B \quad (3)$$

$$V_{k_{MIN}} \leq V_k \leq V_{k_{MAX}}, \forall k \in B \quad (4)$$

$$\theta_{k_{MIN}} \leq \theta_k \leq \theta_{k_{MAX}}, \forall k \in B \quad (5)$$

$$0 \leq PG_g \leq PG_{g_{MAX}}, \forall g \in G \quad (6)$$

$$QG_{g_{MIN}} \leq QG_g \leq QG_{g_{MAX}}, \forall g \in G \quad (7)$$

Onde:

- B Conjunto de barras da rede;
- W Conjunto de cenários considerados;
- PG_{gk} Potência gerada pelo gerador 'g' na barra 'k';
- $PGW_{jk,w}$ Potência ativa eólica gerada na planta 'j' na barra 'k' no cenário 'w';
- PL_k Potência ativa demandada na barra 'k';
- p_{km} Fluxo de potência ativa entre as barras 'k' e 'm';
- Ω_k Conjunto de barras conectados a barra 'k';
- QG_{gk} Potência reativa do gerador 'g' na barra 'k';
- $QGW_{j,w}$ Potência reativa eólica da planta 'j' na barra 'k' no cenário 'w';
- QL_k Potência reativa demandada na barra 'k';
- $V_{k_{MIN}}$ Tensão mínima na barra 'k';
- $V_{k_{MAX}}$ Tensão máxima na barra 'k';
- $\theta_{k_{MIN}}$ Ângulo mínimo na barra 'k';
- $\theta_{k_{MAX}}$ Ângulo máximo na barra 'k';
- G Conjunto de geradores da rede;
- $PG_{g_{MAX}}$ Potência ativa máxima do gerador 'g';
- $QG_{g_{MIN}}$ Potência reativa mínima do gerador 'g';
- $QG_{g_{MAX}}$ Potência reativa máxima do gerador 'g'.

Os fluxos de potência ativa, p_{km} , e reativa, q_{km} , utilizados nas restrições (2) e (3) são calculados por (8) e (9), respectivamente. As perdas são avaliadas pela equação (10).

$$p_{km} = V_k^2 g_{km} - V_k V_m (g_{km} \cos \theta_{km} + b_{km} \sin \theta_{km}), \quad \forall k \in B \quad (8)$$

$$q_{km} = -V_k^2 (b_{km} + b_{km}^{sh}) - V_k V_m (g_{km} \sin \theta_{km} - b_{km} \cos \theta_{km}), \quad \forall k \in B \quad (9)$$

$$perdas_w = p_{km,w} + p_{mk,w} \quad (10)$$

Onde:

- V_k Magnitude da tensão na barra 'k';
- θ_{km} Diferença angular entre as barras 'k' e 'm';
- g_{km} Condutância da linha entre as barras 'k' e 'm';
- b_{km} Susceptância da linha entre as barras 'k' e 'm';
- b_{km}^{sh} Susceptância *Shunt* da linha entre as barras 'k' e 'm'.

O *software* utilizado para realizar as simulações foi o Matlab e a *fmincom* foi a função utilizada para encontrar a solução de cada um dos FPOs. A função objetivo implementada na *fmincom* está representada em (1). As expressões matemáticas (2) - (10) auxiliam a função objetivo à encontrar a solução ótima para cada cenário de geração eólica, e para cada cenário é realizado um FPO.

2.2 K-means

O método k-means (MacQueen, 1967), caracterizado por sua simplicidade e eficiência computacional, é altamente recomendável para o trabalho com grandes bases de dados numéricos. A forma mais usual de inicialização desta técnica é a escolha aleatória de k pontos para serem usados como centroides iniciais, sendo k igual ao número de grupos desejado (Souza, 2008). Definido o número de pontos de cada grupo, define-se a probabilidade de ocorrência de cada grupo. O centroide e a probabilidade são utilizados para representar o histórico.

O algoritmo k-means tem como objetivo minimizar a distância entre um dado pertencente a um grupo e o centroide deste mesmo grupo, conforme (11).

$$\min \sum_{j=1}^k \sum_{i=1}^n \|x_i^j - c_j\|^2 \quad (11)$$

Onde:

- k Número de grupos;
- n Número de realizações da série;
- j Grupo de dados;
- x_i^j Dados da série;
- c Centroide.

2.3 Algoritmo Hierárquico - AHI

Os algoritmos baseados nos métodos hierárquicos agrupam um determinado conjunto de dados em uma estrutura hierárquica de acordo com o grau de proximidade entre os indivíduo, representando os dados a partir de um dendrograma, que é uma árvore que divide os agrupamentos em diversos níveis de proximidade. A raiz do dendrograma representa o conjunto de dados. Os grupos são obtidos cortando-se o dendrograma em diferentes níveis de acordo com o número k de centroides desejado. O dendrograma pode ser criado de forma aglomerativa ou divisiva.

Os métodos com abordagem aglomerativa são mais populares do que aqueles com abordagem divisiva. Os autores em (Zhang et al., 1996) afirmam que os métodos hierárquicos não buscam encontrar os melhores centroides, mas mantém junto o par mais próximo (ou separar o par mais distante) de objetos para formar centroides. O dendrograma hierarquiza o grau de semelhança entre os agrupamentos tornando possível obter uma visão bidimensional da similaridade ou dissimilaridade de todos os dados. Esta propriedade facilita a comparação de diferentes configurações de agrupamentos para um mesmo conjunto de dados. Entretanto, para se determinar o nível apropriado do dendrograma ou o número ideal de agrupamentos algum critério de decisão deve ser empregado.

2.4 ISODATA

É um método originado da modificação do algoritmo de k-means, em que é implementada a função de fusão ou divisão dos centroides dinamicamente durante o processo iterativo, de acordo com alguns parâmetros predefinidos pelo usuário. Assim, é possível eliminar os centroides que não possuem um número mínimo de elementos especificado, evitando a distorção do agrupamento.

O algoritmo define, aleatoriamente, os primeiros centroides. Com os centroides definidos, os dados são classificados em subconjuntos com base na distância de cada ponto a cada centroide. Cada dado é agrupado ao centroide mais próximo, de acordo com a medida de distância adotada. Em seguida, calcula-se: a média de cada subconjunto; o desvio padrão dos dados pertencentes ao subconjunto; e a distância entre os centroides. Os centroides são recalculados pelas médias dos dados de cada subconjunto. Em sequência, os centroides com número de elementos (θ_N) menor do que o estabelecido são eliminados. Assim, os centroides remanescentes entram em processo divisivo ou aglomerativo.

A divisão dos centroides em agrupamentos menores ocorre quando os seguintes critérios são atendidos:

- (1) O desvio padrão em qualquer dimensão é maior que o limite especificado (θ_C);
- (2) O centroide possui um número de elementos suficiente para dividir.

Se as condições de divisão forem atendidas, então são criados dois novos pontos dentro do subconjunto. O primeiro novo ponto é criado adicionando uma unidade ao componente que possuir o maior desvio padrão com relação ao ponto médio original. Já o segundo é criado subtraindo uma unidade dessa componente. Após cada divisão, o conjunto modificado de pontos médios é usado como o conjunto de pontos para a próxima iteração.

Por outro lado, o método da junção dos centroides em agrupamentos maiores ocorre se as distâncias entre os centroides forem menor que o limite (θ_E). Caso seja obedecido o critério para a aglomeração o centroide é obtido através da média dos centroides dos dois subconjuntos que serão agrupados. O critério de parada para o ISODATA é definido por um número máximo de iterações.

2.5 Simulação Monte Carlo

A Simulação Monte Carlo (SMC) é um método estatístico baseado em amostragens aleatórias massivas para a obtenção de resultados numéricos, ou seja, repetindo sucessivas simulações um elevado número de vezes, a fim de obter resultados para diferentes cenários. Neste trabalho, a SMC é usada como *benchmark* para os métodos aglomerativos.

Para a obtenção do número de iterações da SMC, utiliza-se como critério de parada o coeficiente de variação β , que apresenta o nível de incerteza da estimativa de (Manso, 2012). Com tal critério, garante-se um número adequado de iterações e a qualidade dos resultados. O passo a passo para o cálculo do coeficiente de variação é definido em (12)-(15).

$$V\{\bar{E}[F_T(U)]\} = \frac{V[F_T(U)]}{N} \quad (12)$$

$$V[F_T(U)] = \frac{\sum_{t=1}^N \{F_T(U_t) - \bar{E}[F_T(U)]\}^2}{N-1} \quad (13)$$

$$\bar{E}[F_T(U)] = \frac{\sum_{t=1}^N \{F_T(U_t)\}}{N} \quad (14)$$

$$\beta = \frac{\sqrt{V\{\bar{E}[F_T(U)]\}}}{\bar{E}[F_T(U)]} \quad (15)$$

Onde:

F_T	Função Teste;
N	Número de iterações;
$\bar{E}[\cdot]$	Valor esperado da função teste;
$V[\cdot]$	Variância da função teste;
$V\{\cdot\}$	Variância do valor esperado da função teste;
β	Coefficiente de variação.

Para a aplicação do critério de parada da Simulação Monte Carlo não sequencial, a função teste foi representada pela potência ativa gerada pelos parques eólicos.

2.6 Comparação entre os Métodos de Clusterização

Para a comparação dos resultados, são analisadas as perdas elétricas que ocorre no SEP. Como uma simulação é composta de várias iterações, a comparação foi feita com base nas médias ponderadas, considerando as n rodadas de cada simulação.

Para a SMC foi realizada a média aritmética de todos os valores das variáveis correspondentes a cada simulação. Já no caso em que a geração eólica é representada por técnicas de agrupamento de dados, em todos os casos foi realizada a média ponderada dos dados, de acordo com a probabilidade de ocorrência de cada centroide representativo.

Para a realização deste estudo, o software utilizado para as simulações foi o MATLAB R2015a em um processador Intel® Core™ i3^{7th} e 4GB de memória RAM.

3. ESTUDO DE CASO

A Figura 1 mostra o sistema teste IEEE 118 barras, (Chopade and Bikdash, 2012), utilizado nas simulações. Este sistema é composto por 118 barras, 186 linhas e 54 geradores. Neste trabalho, o sistema foi adaptado de modo a receber, nas barras 53, 67 e 84, três usinas eólicas, cada uma contendo 20 aerogeradores de 5 MW. Considera-se que as usinas eólicas possuem um fator de potencia de 0,92 atrasado para representar a potência reativa consumida.

Nesse estudo, os valores mínimos e máximos das tensões nas barras foram mantidos entre 0,95 pu e 1,05 pu, respectivamente. As potências ativas e reativas variaram entre as capacidades mínimas e máximas de cada planta geradora.

A seguir serão apresentados os 4 estudos de caso considerando as séries eólicas representadas pelos algoritmos aglomerativos k-means, AHI e ISODATA e um quarto estudo de caso com a SMC. Todos os resultados são relativos as perdas que ocorrem no sistema. Para as técnicas k-means e AHI foram consideradas simulações com 200 e

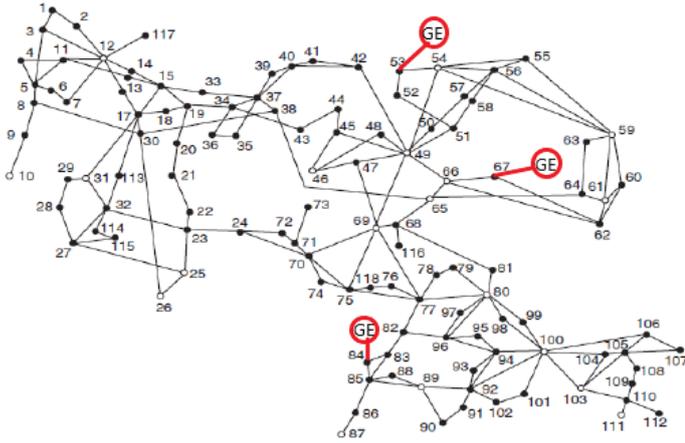


Figura 1. Sistema IEEE 118 Barras.

10 centroides, que correspondem ao número de cenários do FPO.

A simulação da técnica de agrupamento ISODATA ocorre de maneira diferente da técnica de k-means e AHI. Nesse tipo de simulação, são pré estabelecidos pelo usuário alguns parâmetros limites e a partir destes limites, o algoritmo retorna o número de agrupamentos proporcionais. Os parâmetros limites a serem definidos pelo usuário se encontram na lista a seguir.

- θ_N Limite do número de elementos para a eliminação de um agrupamento;
- θ_C Limite da distância para a união de um agrupamento;
- θ_E Limite de desvio padrão para a divisão de um agrupamento;
- k_{max} Número máximo de agrupamentos;
- L Número máximo de centroides que podem ser combinados em uma única iteração;
- I Número máximo de iterações permitidas.

A escolha desses parâmetros limites são definidas pelo usuário e são dependentes da dimensão do problema, assim como da correlação existente entre as séries originais utilizadas. Para o caso em estudo, foi realizada uma análise de sensibilidade do resultado do algoritmo com relação a esses parâmetros resultando nos valores apresentados na Tabela 1. A análise foi baseada em Ball and Hall (1965) e usou como critério a potência ativa do parque eólico.

Tabela 1. Limites dos parâmetros.

θ_N	θ_C	θ_E	k_{max}	L	I
1640	0,03	0,01	200	100	200

Para a SMC foi considerado o critério de parada apresentada em Manso (2012), em que o coeficiente de variação β possui tolerância de 2%. Considerando este critério, o número total de iterações resultantes da SMC foi de 2024.

Conforme apresentado anteriormente, serão demonstrados os resultados para as perdas na rede elétrica. Na Tabela 2 encontram-se as médias ponderadas das perdas para cada um dos agrupamentos e da SMC. Observando a Tabela 2 verifica-se que a média ponderada que mais se aproxima da SMC, 8,98MW, foi encontrada através do agrupamento AHI de 10 centroides, com 9,00MW. E a técnica com média ponderada com valor que menos se aproximou da

SMC foi o k-means de 10 cenários, com 8,89MW, mas ainda ficou com valor próximo ao da SMC.

Tabela 2. Perdas Médias.

Técnica	Perdas (MW)
SMC	8,98
K-Means = 200	9,02
K-Means = 10	8,86
AHI = 200	9,06
AHI = 10	9,00
ISODATA	8,92

Outro resultado apresentado são as porcentagens de cenários da SMC que ultrapassaram os máximos encontrados em cada uma das outras técnicas. Na Tabela 3 estão as porcentagens dessas ultrapassagens para as perdas. As técnicas do AHI, de 10 e 200 cenários, e o ISODATA não tiveram seus máximos ultrapassados pela SMC.

Tabela 3. Porcentagens de Máximos de Perdas Ultrapassados.

Técnica	Cenários Ultrapassados (%)
K-Means = 200	0,30
K-Means = 10	0,69
AHI = 200	0
AHI = 10	0
ISODATA	0

Logo, analisando de forma geral e comparando cada um dos resultados apresentados nas Tabelas 2 e 3, verifica-se que a técnica que melhor se adequou foi a AHI com 10 cenários e a ISODATA.

3.1 Comparativo entre as simulações

A Tabela 4 mostra os tempos computacionais para cada simulação realizada. Para cada simulação do FPO o tempo de processamento foi em média 80 segundos para cada cenário, logo, o número reduzido de agrupamentos resulta em menor esforço computacional.

Tabela 4. Tempos Computacionais.

Técnica	Número de Centroides	Tempo Computacional (h)
Monte Carlo	-	45,65
K-Means	200	4,51
K-Means	10	0,27
Hierárquicos	200	4,48
Hierárquicos	10	0,22
ISODATA	12	0,27

A Tabela 4 mostra que a utilização de métodos de agrupamento possibilitou uma significativa redução no tempo computacional de cada simulação. Observando mais uma vez a Tabela 2, é possível ver que tal ganho de eficiência não é acompanhado de distorções significativas na qualidade dos resultados obtidos.

4. CONCLUSÕES

Neste artigo foi avaliada a aplicação das técnicas k-means, Algoritmos Hierárquicos (AHI) e ISODATA para agrupamento de séries originais de estados de geração eólica no fluxo de potência ótimo (FPO). De modo a avaliar o

efeito da utilização de tais técnicas, no que diz respeito à eficiência e à eficácia na solução do problema, comparações com a Simulação de Monte Carlo (SMC) foram realizadas utilizando-se o sistema IEEE 118 barras. Dos resultados obtidos, os seguintes pontos podem ser enfatizados:

- Os resultados com a técnica AHI, de 10 centroides, e o ISODATA foram as que mais se aproximaram da SMC;
- A quantidade de centroides tem visível influência na relação custo-benefício das soluções, devendo sua escolha levar a um bom compromisso entre a precisão dos resultados e o tempo para obtê-los;
- As técnicas de k-means com número de centroide $k = 10$, AHI com $k = 10$ centroides e ISODATA foram as que apresentaram o menor tempo computacional comparado com todos os outros tipos de simulação.

As diferentes metodologias de clusterização avaliadas foram eficientes para representar as séries originais de potência ativa dos três parques eólicos correlacionados. Os resultados obtidos observados para o FPO mostram ainda que a qualidade das soluções obtidas não é significativamente afetada pelo uso das técnicas de clusterização. Isto revela o potencial do uso de tais técnicas como uma boa alternativa para a redução do tempo computacional nas análises de planejamento e operação na presença de geração eólica.

AGRADECIMENTOS

Os autores agradecem a CAPES, CNPq, FAPEMIG e INERGE pelo suporte durante o desenvolvimento do trabalho.

REFERÊNCIAS

- Adedeji, P.A., Akinlabi, S., Madushele, N., and Olatunji, O.O. (2020). Wind turbine power output very short-term forecast: A comparative study of data clustering techniques in a pso-anfis model. *Journal of Cleaner Production*, 254, 120135. doi:https://doi.org/10.1016/j.jclepro.2020.120135.
- Ball, G. and Hall, D. (1965). *Isodata, a Novel Method of Data Analysis and Pattern Classification*. Architectural design. Stanford Research Institute.
- Baringo, L. and Baringo, A. (2018). A stochastic adaptive robust optimization approach for the generation and transmission expansion planning. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 33(1), 792–802.
- Baringo, L. and Conejo, A.J. (2013). Correlated wind-power production and electric load scenarios for investment decisions. *Applied Energy*, 101, 475–482.
- Chopade, P. and Bikdash, M. (2012). Modeling for survivability of smart power grid when subject to severe emergencies and vulnerability. *Southeastcon, 2012 Proceedings of IEEE*. doi:10.1109/SECon.2012.6196924.
- Dehghan, S., Amjady, N., and Conejo, A.J. (2016). Reliability-constrained robust power system expansion planning. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 31(3), 2383–2392.
- Deng, X., Zhang, P., Jin, K., He, J., X., W., and Wang, Y. (2019). Probabilistic load flow method considering large-scale wind power integration. *J. Mod. Power Syst. Clean Energy*, 7(4), 813–825.
- Lin, C., Bie, Z., Zhou, B., Wang, T., and Wang, T. (2018). Comparison of different methods in stochastic power flow with correlated wind power generation. *Science Direct*, 67–72.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, 281–297. University of California Press, Berkeley, Calif.
- Malta, C.S. (2009). *Estudos de Séries Temporais de Vento Utilizando Análises Estatísticas e Agrupamentos de Dados*. Monografia, UFRJ.
- Manso, L.A.F. (2012). Avaliação do desperdício de energia eólica utilizando fluxo de potência cronológico. In *Congresso Brasileiro de Automática (CBA)*, Online. Campina Grande, Brasil.
- Operador Nacional do Sistema Elétrico, O.N.S. (2019). Boletim mensal de geração eólica março/2019.
- Souza, T.A. (2008). *Agrupamento de Séries Temporais de Vento para Avaliação da Disponibilidade de Geração de Usinas Eólicas*. Dissertação, UFRJ.
- Wogrin, S., Tejada-Arango, D.A., Pineda, S., and Morales, J.M. (2019). What time-period aggregation method works best for power system operation models with renewables and storage? *2019 International Conference on Smart Energy Systems and Technologies (SEST)*, 1–6.
- Xu, R. and Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3), 645–678.
- Zhang, T., Ramakrishnan, R., and Livny, M. (1996). Birch: an efficient data clustering method for very large databases. In *SIGMOD '96 Proceedings of the 1996 ACM SIGMOD international conference on Management of data*, 103–114. Montreal, Canada.