

# Técnicas Não Supervisionadas de Aprendizado de Máquina Aplicadas na Avaliação da Confiabilidade Composta de Sistemas Elétricos de Potência

Fernando A. Assis\*, Alex Júnior C. Coelho\*, Lucas D. Rezende\*,  
Leonidas C. Resende\*, Armando M. Leite da Silva\*\*

\* Departamento de Engenharia Elétrica, Universidade Federal de São João del-Rei - UFSJ, São João del-Rei, MG, Brasil (e-mails: fernandoassis@ufsj.edu.br, alexjuniordacunhacoelho@gmail.com, lucas.diniz40@hotmail.com.br, leonidas@ufsj.edu.br).

\*\* Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, PUC-Rio, Rio de Janeiro, RJ, Brasil (e-mail: am.leitedasilva@gmail.com).

**Abstract:** This paper proposes a simple and new method for assessing the composite reliability of electrical power systems. The nonsequential Monte Carlo simulation method is combined with unsupervised machine learning techniques in order to reduce the computational burden involved in the process of estimating composite reliability indices. The proposed approach allows different unsupervised techniques to be employed, in order to obtain significant reductions in processing times, without losing the accuracy of the estimated indices. The results obtained with the use of three different classification techniques (*Kohonen self-organizing map*, *K-means*, and *K-medoids*) are presented and deeply analyzed.

**Resumo:** Neste artigo é proposta uma metodologia simples e nova para avaliação da confiabilidade composta de sistemas elétricos de potência. O método de simulação Monte Carlo não sequencial é combinado com técnicas não supervisionadas de aprendizado de máquina com o intuito de reduzir o esforço computacional envolvido no processo de estimativa dos índices de confiabilidade composta. A metodologia permite que diferentes técnicas não supervisionadas sejam empregadas, tendo em vista a obtenção de reduções significativas nos tempos de processamento, sem que haja perda de precisão dos índices de desempenho estimados. Os resultados obtidos com a utilização de três diferentes técnicas de classificação (*Kohonen self-organizing map*, *K-means*, and *K-medoids*) são apresentados e amplamente analisados.

**Keywords:** Clustering algorithms; Composite reliability; Monte Carlo simulation; Pattern recognition; Unsupervised machine learning.

**Palavras-chaves:** Algoritmos de classificação; Aprendizado de máquina não supervisionado; Confiabilidade composta; Reconhecimento de padrão; Simulação Monte Carlo.

## 1. INTRODUÇÃO

O crescimento contínuo da demanda por energia elétrica exige que os sistemas elétricos de potência estejam em constante expansão. Neste contexto, é de suma importância que os planejadores destes sistemas conciliem segurança da operação com os aspectos econômicos durante o processo de tomada de decisão. Na determinação de planos de expansão que garantam níveis adequados de segurança e, ao mesmo tempo, tarifas reduzidas para o consumidor final, oriundas da minimização dos custos de investimento e operação, metodologias baseadas na avaliação probabilística de confiabilidade dos sistemas se apresentam como ferramentas extremamente úteis e flexíveis (Leite da Silva et al., 2010; Miloca et al., 2015).

Em geral, a avaliação probabilística de confiabilidade permite conhecer os riscos de falha na operação do sistema, sendo levadas em consideração as incertezas associadas à disponibilidade de seus equipamentos. Estes riscos são traduzidos

por meio de índices de desempenho estimados durante o processo de avaliação da confiabilidade (Pereira & Balu, 1992; Alan & Billinton, 1996). Um dos grandes desafios enfrentados no emprego de técnicas baseadas na avaliação probabilística de confiabilidade durante as etapas de planejamento está relacionado aos custos computacionais (tempos de processamento) exigidos. Essa dificuldade é intensificada para problemas de planejamento de longo prazo, que envolvem a necessidade de se avaliar a confiabilidade composta de geração e transmissão do sistema (Alan & Billinton, 1996). Na avaliação de confiabilidade composta, devem ser levadas em consideração as incertezas associadas tanto à disponibilidade de geradores quanto dos equipamentos dedicados ao transporte de energia. Neste caso, para se atingir a convergência na estimativa dos índices de confiabilidade, pode ser necessária a realização de um grande número de avaliações de fluxo de potência e de ações corretivas/remediadoras (e.g., redespacho de geração e corte de carga), as quais são baseadas na solução de problemas de otimização.

Com o intuito de acelerar o processo da avaliação de confiabilidade composta de sistemas elétricos, diversos trabalhos da literatura relacionada vêm propondo estratégias inovadoras. Destaque pode ser dado ao emprego de técnicas de redução de variância (e.g., González-Fernández et al., 2013 e Shu et al., 2014), métodos de representação *fuzzy* (e.g., Saraiva et al., 1996), computação paralela e distribuída (e.g., Borges et al., 2001), métodos de otimização inteligentes baseados em população (e.g., Benidris et al., 2015), e técnicas de reconhecimento de padrão (e.g., Luo et al., 1999; Luo et al., 2000, e Urgan et al., 2019). De forma geral, muitas dessas estratégias configuram ferramentas híbridas, que combinam o emprego de técnicas diversas com métodos de simulação Monte Carlo (Alan & Billinton, 1996).

Buscando melhor aproveitar e investigar o potencial de técnicas não supervisionadas de aprendizado de máquina baseadas em reconhecimento de padrões, é proposta no presente trabalho uma metodologia simples para avaliação probabilística de confiabilidade composta de sistemas elétricos. Basicamente, a metodologia combina técnicas não supervisionadas de aprendizado de máquina (NSAM) com o método de simulação Monte Carlo (SMC) não sequencial (Alan & Billinton, 1996). Tal combinação permite estimar índices de confiabilidade composta de forma eficiente e com redução significativa do esforço computacional envolvido, quando comparada à avaliação realizada via simulação Monte Carlo não sequencial convencional. É importante destacar que a metodologia proposta não é atrelada a apenas uma única técnica NSAM, o que permite flexibilidade em sua utilização. Tal observação é ilustrada e analisada a partir de resultados promissores obtidos por meio da metodologia proposta e emprego de três diferentes técnicas NSAM: Kohonen *self-organizing map* (SOM) (Kohonen, 1982), *K-means* (Jain, 2010) e *K-medoids* (Park & Jun, 2009) clustering. O sistema teste IEEE-RTS (*IEEE Reliability Test System*) (IEEE APM, 1979), considerando dois diferentes cenários de geração e carga, é utilizado para a realização dos estudos apresentados e analisados neste trabalho.

## 2. CONFIABILIDADE COMPOSTA

A avaliação da confiabilidade composta de um sistema elétrico consiste em examinar, em termos de adequação estática, a capacidade deste em atender à carga, estando os seus equipamentos de geração e de transporte de energia (e.g., linhas de transmissão e transformadores) sujeitos à ocorrência de falhas durante a operação.

Na avaliação probabilística de confiabilidade, o diagnóstico de adequação do sistema pode ser traduzido por meio de estimativas de índices de confiabilidade (Pereira & Balu, 1992), ligadas, normalmente, ao evento de falha do sistema (i.e., necessidade de se promover o corte de carga, seja por insuficiência de geração, por incapacidade de transporte de energia ou pelo efeito combinado dos equipamentos de geração e transmissão disponíveis). Índices de confiabilidade, tais como LOLP (*Loss of Load Probability*), LOLE (*Loss of Load Expectation*), EPNS (*Expected Power Not Supplied*) e EENS (*Expected Energy Not Supplied*), são estimados com base em informações da natureza estocástica de disponibilidade dos equipamentos e do comportamento aleatório da carga.

A forma como os índices de confiabilidade são estimados depende da técnica de avaliação utilizada. No método de SMC não sequencial, os índices LOLP e EPNS são calculados através de estimativas do valor esperado de uma função teste (Alan & Billinton, 1996), dada por:

$$\tilde{E}[F] = \frac{1}{N} \sum_{v=1}^N F(x^v) \quad (1)$$

em que  $N$  é o número total de estados gerados,  $x^v$  representa o estado  $v$  amostrado, e  $F$  é a função teste empregada de acordo com o índice. As funções teste para estimar os índices LOLP e EPNS são dadas, respectivamente, por (2) e (3):

$$F_{\text{LOLP}}(x^v) = \begin{cases} 1, & \text{se estado } x^v \text{ é de falha} \\ 0, & \text{caso contrário} \end{cases} \quad (2)$$

$$F_{\text{EPNS}}(x^v) = \begin{cases} \Delta P, & \text{se estado } x^v \text{ é de falha} \\ 0, & \text{caso contrário} \end{cases} \quad (3)$$

onde  $\Delta P$ , em MW, representa, para cada estado de falha, o corte de carga realizado.

Por fim, os índices LOLE e EENS podem ser obtidos, respectivamente, conforme relações (4) e (5):

$$\text{LOLE} = T \times \text{LOLP} \quad (4)$$

$$\text{EENS} = T \times \text{EPNS} \quad (5)$$

onde  $T$ , em horas, é o período de estudo (e.g., 8736 horas para avaliação da confiabilidade no período de um ano).

As incertezas da estimativa relacionadas aos cálculos dos índices conforme (1) podem ser obtidas como:

$$\beta = \sqrt{\frac{V(\tilde{E}[F])}{\tilde{E}[F]}} \quad (6)$$

sendo  $V(\tilde{E}[F])$  a variância do estimador, dada por:

$$V(\tilde{E}[F]) = \frac{V(F)}{N} \quad (7)$$

onde  $V(F)$  é a variância da função teste. As incertezas associadas aos índices de confiabilidade podem indicar se o número  $N$  de estados gerados é suficiente para uma adequada estimativa de seus valores.

## 3. AVALIAÇÃO DA CONFIABILIDADE COMPOSTA VIA TÉCNICAS NÃO SUPERVISIONADAS DE APRENDIZADO DE MÁQUINA

A metodologia proposta neste trabalho visa à aceleração do processo de avaliação de confiabilidade composta de sistemas elétricos de potência. De forma geral, o método de SMC não sequencial, baseado na representação a espaço de estados, é combinado a uma técnica NSAM para redução do número de estados avaliados por análises de fluxo de potência na estimativa dos índices de confiabilidade (Pereira & Balu, 1992; Alan & Billinton, 1996). Um fluxograma básico, que representa o algoritmo da metodologia proposta é apresentado na Fig. 1. Detalhes da dinâmica inerente ao algoritmo são discutidos a seguir.

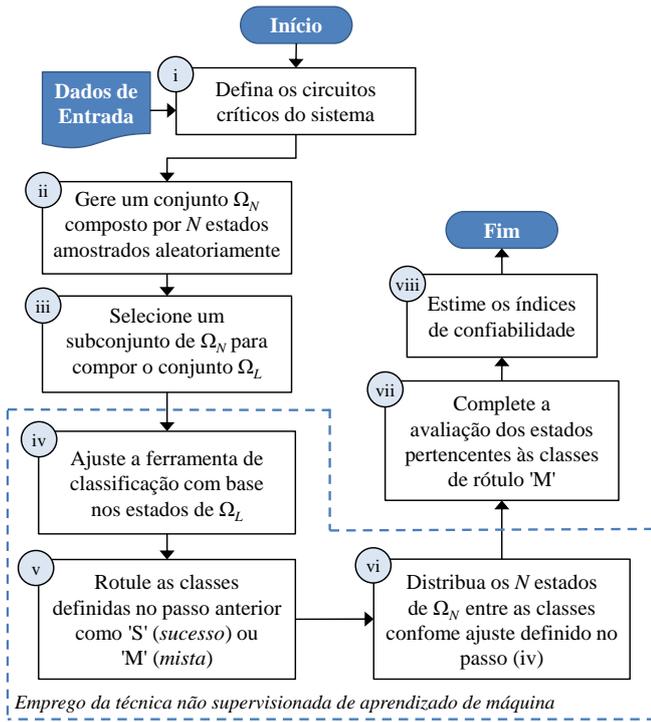


Fig. 1 Fluxograma da metodologia proposta.

### 3.1. Algoritmo da Metodologia Proposta

Para dar início ao processo de avaliação da confiabilidade de acordo com a metodologia proposta neste trabalho, é necessário conhecer alguns dados gerais de entrada, os quais são comuns à avaliação convencional de confiabilidade composta. Basicamente, os dados de entrada compreendem informações gerais relacionadas à topologia geral do sistema, à curva horária de carga para o período de estudo e à natureza estocástica de disponibilidade de seus equipamentos (i.e., taxas de falha e de reparo dos equipamentos). Definidas essas informações iniciais, a metodologia se desenvolve conforme o algoritmo representado na Fig. 1.

**Passo (i)** – Identificação de circuitos críticos do sistema:

O primeiro passo do algoritmo consiste em identificar para o sistema em estudo os circuitos críticos do ponto de vista da avaliação de confiabilidade. Basicamente, consideram-se como críticos aqueles circuitos cujas eventuais indisponibilidades fragilizam a operação do sistema. A ocorrência de indisponibilidade de circuitos deste tipo são informações importantes que são consideradas nas etapas de ajuste/treinamento e de classificação das técnicas não supervisionadas de aprendizado de máquinas conforme metodologia proposta.

A definição em relação à criticidade dos circuitos pode ser realizada previamente pelo planejador, com base em sua experiência na operação do sistema. Adicionalmente, critérios mais determinísticos podem ser empregados. Na metodologia proposta neste trabalho, são considerados críticos os circuitos do sistema que apresentam níveis mais elevados de carregamento (e.g., superior a 80% ou 90%) em

situação de rede intacta no atendimento da carga pico. Para tanto, uma avaliação de fluxo de potência ótimo do sistema é requerida na execução do passo (i) do algoritmo (Fig. 1). Além disso, como um segundo critério, consideram-se como críticos aqueles circuitos que possuem pelo menos uma das barras terminais interligadas ao sistema por apenas dois circuitos em série (o próprio circuito e mais um) ou apenas pelo próprio circuito. No caso de indisponibilidade deste tipo de circuito, a barra se torna terminal ou fica isolada do restante da rede e, portanto, fragiliza a operação do sistema.

**Passo (ii)** – Geração de estados para avaliação:

No passo (ii) do algoritmo, são gerados (amostrados)  $N$  estados do sistema, os quais devem ser armazenados no conjunto  $\Omega_N$ . A geração de estados neste passo é realizada como no método de SMC não sequencial, sendo que cada estado do sistema é composto pelos estados individuais dos equipamentos de geração e transmissão (disponíveis ou indisponíveis) e por um nível de carga amostrado. A quantidade  $N$  de estados a serem gerados deve estar relacionada ao nível de confiabilidade do sistema, ou seja, quanto mais robusto em termos de adequação estática for o sistema, maior o número de amostras requeridas para uma boa estimativa dos índices de confiabilidade.

**Passo (iii)** – Seleção de estados para ajuste/treinamento:

Um subconjunto de estados de  $\Omega_N$  é selecionado no passo (iii) para efetuar o ajuste do classificador empregado pela metodologia (i.e., para ajustar/treinar a técnica de classificação não supervisionada, o que é realizado no passo (iv)). Os estados selecionados são armazenados no conjunto  $\Omega_L$ . De forma geral, para ajustar/treinar a ferramenta de classificação é essencial que apenas um número bastante reduzido de estados seja utilizado, uma vez que o tempo (esforço computacional) envolvido neste processo está diretamente atrelado ao tamanho da amostra utilizada. Portanto, o número de estados selecionados de  $\Omega_N$  deve representar apenas uma pequena parcela (e.g., <1%) do total de  $N$  estados gerados. Além disso, o processo de seleção deve fornecer à técnica de classificação uma amostra que seja potencialmente representativa das situações de falha e de sucesso do sistema.

Portanto, na metodologia proposta, três informações relacionadas a cada estado são observadas no processo de seleção do passo (iii), sendo estas mesmas informações também utilizadas para o ajuste/treinamento da ferramenta de classificação no passo (iv):

- *nível de carga*: nível de carga sorteado para o estado conforme curva de carga do sistema;
- *reserva geração*: diferença entre capacidade de geração disponível e montante de carga do estado, com valor normalizado em relação ao maior valor de reserva observado para os  $N$  estados gerados;
- *número de circuitos críticos indisponíveis*: total de circuitos definidos como críticos no passo (i) que estão indisponíveis no estado.

A definição das informações a serem observadas foi baseada em testes realizados, onde se buscou conciliar velocidade de processamento na fase de ajuste (ligada ao número de informações observadas) e desempenho da ferramenta na classificação dos estados. Salienta-se, contudo, que a depender das características do sistema em estudo, outras informações de estado podem substituir ou complementar as informações aqui consideradas.

De forma geral, até completar a quantidade de estados necessária para o conjunto  $\Omega_L$ , o filtro de seleção captura dentre os  $N$  estados gerados no passo (ii):

- a) estados que apresentem *nível de carga* igual ou superior a um certo limite percentual definido (e.g., 80% ou 90%) e, ao mesmo tempo, *reserva de geração* igual ou inferior a um limite percentual também definido (e.g., 5% ou 10%). Espera-se que estes estados sejam potencialmente representativos para o diagnóstico de falha do sistema;
- b) estados que apresentem *nível de carga* igual ou inferior a um certo limite percentual definido (e.g., 70% ou 80%) e, ao mesmo tempo, *reserva de geração* igual ou superior a um limite percentual também definido (e.g., 10% ou 15%). Espera-se que estes estados sejam potencialmente representativos para o diagnóstico de sucesso do sistema;
- c) estados que apresentem *número de circuitos críticos indisponíveis* iguais ou superiores a 1 (um). Estes estados são, também, potencialmente representativos do diagnóstico de falha do sistema.

Ressalta-se que a seleção de estados com base no filtro definido não demanda a realização de avaliações de estados com relação à execução de algoritmos de fluxo de potência com medidas remediadoras (fluxo de potência ótimo).

**Passo (iv)** – Ajuste/treinamento da técnica de classificação:

Formado o conjunto  $\Omega_L$ , a técnica de classificação empregada deve então ser ajustada no cumprimento do passo (iv). Uma vez que ainda não se conhece o diagnóstico (falha ou sucesso) do sistema para cada estado selecionado, este ajuste/treinamento deve ser realizado de forma não supervisionada, com base apenas nas informações observadas dos estados de  $\Omega_L$ , aquelas listadas na apresentação do passo (iii). Nesta etapa, cada tipo de técnica apresenta uma estratégia própria de manipulação das informações. No caso de técnicas de redes neurais artificiais, por exemplo, as informações de entrada são manipuladas para ajustar coeficientes de polinômios e/ou de um conjunto de pesos. Em geral, o ajuste/treinamento de técnicas de aprendizado de máquina para classificação corresponde à realização de processos iterativos, que depois de convergidos, determinam  $K$  classes para os dados de entrada. A depender da técnica empregada, o número  $K$  de classes é um parâmetro que deve ser definido na execução do passo (iv).

**Passo (v)** – Rotulagem das classes definidas:

Em (v), as  $K$  classes definidas no passo anterior para os estados de  $\Omega_L$  devem ser rotuladas como 'S' (sucesso) ou 'M'

(mista). Para tanto, o sistema é avaliado em termos de capacidade de atendimento à carga para cada estado do conjunto  $\Omega_L$ . Ou seja, os estados de  $\Omega_L$  são, neste ponto, definidos como de sucesso ou de falha a partir de análises de fluxo de potência ótimo. Em seguida, na definição dos rótulos, se para uma dada classe todos os estados são de sucesso, seu rótulo é o 'S'. Caso contrário, ou seja, se pelo menos um estado da classe for de falha, o rótulo dado à classe é o 'M'.

**Passo (vi)** – Classificação de todos os estados gerados:

O passo (vi) compreende classificar entre as  $K$  classes definidas todos os  $N$  estados gerados e armazenados em  $\Omega_N$ . Estes estados são classificados conforme característica própria da técnica de classificação empregada e observando as mesmas informações de estado definidas para o filtro de seleção no passo (ii) e para ajuste/treinamento no passo (iii). Na verdade, apenas os estados de  $\Omega_N$  não selecionados para compor o conjunto  $\Omega_L$  necessitam ainda serem classificados. Ao fim da execução do passo (vi), cada um dos  $N$  estados inicialmente gerados pertence a uma das  $K$  classes definidas e rotuladas (i.e., cada estado pertence a uma única classe que pode ser do tipo 'S' ou do tipo 'M').

**Passo (vii)** – Avaliação dos estados das classes do tipo 'M':

No passo (vii), todos os estados que foram atribuídos às classes rotuladas como 'M' devem ser avaliados em relação à capacidade de atendimento à carga. Na prática, parte dos estados dessas classes já foi avaliada no passo (v), no processo de rotulagem de classes. É possível verificar, portanto, que apenas nos passos (v) e (vii) da metodologia proposta, avaliações de fluxo de potência com medidas remediadoras precisam ser realizadas. Inclusive, para as classes rotuladas como 'S', as quais devem representar a maior parcela do total dos  $N$  estados amostrados, apenas uma pequena parte dos estados precisou ser avaliada. Neste ponto, é verificado o benefício apresentado pela metodologia em termos de redução do esforço computacional na avaliação da confiabilidade composta.

**Passo (viii)** – Estimativa dos índices de confiabilidade:

Finalmente, no passo (viii), os índices de confiabilidade LOLP, LOLE, EPNS e EENS são estimados. A estimativa destes índices é realizada como no método de SMC não sequencial, utilizando as equações de (1) a (5) apresentadas na Seção II. Destaca-se que, neste caso, as funções teste (2) e (3) recebem zero (valor nulo) para todos os estados atribuídos às classes rotuladas como 'S', mesmo que estes estados não tenham sido efetivamente avaliados por meio de análises de fluxo de potência. Ou seja, todos os estados das classes do tipo 'S' são presumidos como de sucesso.

Conforme já apontado, diferentes técnicas não supervisionadas de aprendizado de máquina podem ser empregadas no cumprimento dos passos de (iv) a (vi) relativos à metodologia proposta (Fig. 1). A fim de avaliar o desempenho da metodologia em relação a este quesito, são exploradas no presente trabalho três diferentes técnicas não supervisionadas de classificação, as quais são sucintamente apresentadas nas próximas subseções.

### 3.2. Kohonen Self-Organizing Map (SOM)

Os mapas auto-organizáveis de Kohonen – *Kohonen self-organizing map* (SOM) (Kohonen, 1982) constituem uma técnica de redes neurais artificiais comumente empregada para a tarefa de classificação não supervisionada de dados. Uma rede SOM pode, basicamente, ser formada por uma camada de saída com neurônios dispostos em uma matriz 2D, conforme estrutura representada pela Fig. 2. No exemplo de estrutura da Fig. 2, a camada de entrada possui dois vetores ( $x_1$  e  $x_2$ ) de informações e a camada de saída possui 16 neurônios (mapa de saída 4×4). Na rede SOM, toda informação de entrada  $i$  possui uma conexão com cada neurônio  $j$  da camada de saída. A cada uma dessas conexões, está associado um peso sináptico  $w_{ij}$ , que em conjunto, formam uma matriz de pesos sinápticos que relaciona a entrada com a saída da rede.

A execução do algoritmo SOM pode ser resumida em duas etapas principais: treinamento e classificação. O processo de treinamento (ajuste realizado no passo (iv) da metodologia proposta) consiste em um processo iterativo e competitivo entre os neurônios para adaptação dos pesos sinápticos da rede, o que é realizado com base nos padrões verificados nas informações de entrada. Alguns parâmetros relacionados ao treinamento da rede devem ser definidos no seu emprego, basicamente, os coeficientes inicial e final de aprendizagem e o raio de vizinhança inicial (Kohonen, 1982).

Com os pesos ajustados, a etapa de classificação do SOM (passo (vi)) consiste em identificar, na camada de saída, o neurônio mais representativo para qualquer nova amostra de entrada apresentada à rede. Neste caso, o neurônio, cujos pesos sinápticos adaptados apresentam maior nível de similaridade em relação às características de uma amostra de entrada, representa esta amostra. Para a metodologia proposta (Fig. 1), quando o SOM é empregado, o número de neurônios da camada de saída da rede representa o número  $K$  de classes de estados. Para medidas de similaridade nos processos competitivos da rede, é considerada neste trabalho a distância Euclidiana, dada por:

$$d_{ij} = \sqrt{\sum_{u=1}^p (x_{iu} - x_{ju})^2} \quad (8)$$

onde  $d_{ij}$  representa a distância Euclidiana entre os elementos  $i$  e  $j$ ;  $x_{iu}$  e  $x_{ju}$  representam, respectivamente, a  $u$ -ésima variável das  $p$  variáveis que formam os elementos  $i$  e  $j$ .

Maiores detalhes relativos aos processos de treinamento e de classificação do SOM podem ser verificados em Kohonen (1982); Luo et al. (1999); e Luo et al. (2000).

### 3.3. K-means

O *K-means* (Jain, 2010) é uma técnica simples e eficiente para classificação de dados, que objetiva particionar um conjunto de observações em  $K$  grupos com características similares. Basicamente, para o processo de classificação, o algoritmo *K-means* é iniciado com a seleção aleatória de  $K$  elementos dentro do conjunto de observações (na metodologia proposta,  $\Omega_L$ ) para representar pontos de referência das  $K$  classes a serem definidas no processo. Em

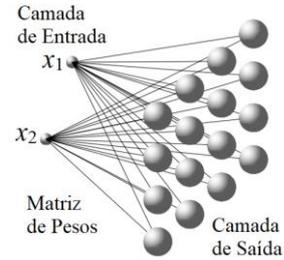


Fig. 2 Estrutura ilustrativa da rede SOM.

seguida, cada elemento do conjunto de observações é atribuído a uma dessas classes de acordo com sua similaridade em relação aos  $K$  pontos de referência definidos. A similaridade entre os elementos pode ser medida, por exemplo, por meio da distância Euclidiana definida em (8).

Na sequência, um processo iterativo é iniciado para substituição dos pontos de referência. Neste processo, cada classe passa a ser representada pelo valor médio verificado entre todos os seus elementos atuais, denominado *centróide*. Com novos  $K$  *centróides* definidos a cada iteração, todos os elementos do conjunto de observações são novamente classificados. Pode-se considerar que há a convergência deste processo iterativo de ajuste (ou seja, fim do passo (iv) na metodologia proposta) quando, entre duas iterações consecutivas, não são mais verificadas modificações das classes para o conjunto de dados de entrada (conjunto  $\Omega_L$ ).

Com o *K-means*, a classificação de estados no passo (vi) da metodologia proposta é realizada por meio da verificação de similaridade entre os  $N$  estados do conjunto  $\Omega_N$  em relação aos  $K$  *centróides* finais do processo iterativo de ajuste, o que define  $K$  classes para estes estados. Mais informações sobre a técnica *K-means* podem ser encontradas em Jain (2010).

### 3.4. K-medoids

De forma geral, a técnica *K-medoids* é bastante semelhante à técnica *K-means*. No processo de classificação, o algoritmo do *K-medoids* também trabalha com  $K$  elementos para representar as  $K$  classes que se deseja definir. Em um processo iterativo, esses elementos vão sendo atualizados e as observações de entrada reclassificadas. A diferença básica entre as técnicas *K-means* e *K-medoids* reside no fato de que, na segunda, o elemento que representa uma dada classe deixa de ser o valor médio (*centróide*) das observações dessa classe e passa a ser objeto localizado na posição mais central, chamado de *medoid* (Park & Jun, 2009). Devido a essa diferença, a técnica *K-medoids* pode ser considerada mais robusta em termos de ruído que a técnica *K-means*. Por outro lado, a depender do tamanho dos dados de entrada, o *K-medoids* pode demandar para o ajuste um maior tempo de processamento.

No presente trabalho, o *K-medoids* empregado para o passo (iv) da metodologia proposta foi implementado como considerado por Park & Jun (2009), com *medoids* iniciais definidos de forma aleatória entre os elementos do conjunto  $\Omega_L$ . A classificação final dos  $N$  estados de  $\Omega_N$  no passo (vi) é realizada como para o *K-means*, a partir da verificação de similaridade de cada estado com os *medoids* finais definidos no processo de ajuste do *K-medoids*.

Tabela 1. Resultados para o sistema IEEE-RTS

	SMC-NSC	K = 9			K = 16			K = 25		
		SOM	K-means	K-medoids	SOM	K-means	K-medoids	SOM	K-means	K-medoids
LOLP	$1,119 \times 10^{-3}$	$1,091 \times 10^{-3}$ (2,43%)	$1,090 \times 10^{-3}$ (2,55%)	$1,117 \times 10^{-3}$ (0,13%)	$1,117 \times 10^{-3}$ (0,13%)	$1,089 \times 10^{-3}$ (2,68%)	$1,117 \times 10^{-3}$ (0,13%)	$1,117 \times 10^{-3}$ (0,13%)	$1,117 \times 10^{-3}$ (0,13%)	$1,117 \times 10^{-3}$ (0,13%)
EPNS (MW)	$1,320 \times 10^{-1}$	$1,303 \times 10^{-1}$ (1,27%)	$1,305 \times 10^{-1}$ (1,08%)	$1,319 \times 10^{-1}$ (0,04%)	$1,319 \times 10^{-1}$ (0,04%)	$1,305 \times 10^{-1}$ (1,12%)	$1,319 \times 10^{-1}$ (0,04%)	$1,319 \times 10^{-1}$ (0,04%)	$1,319 \times 10^{-1}$ (0,04%)	$1,319 \times 10^{-1}$ (0,04%)
N <sub>AVAIL</sub>	700.000	154.189	14.617	51.284	86.264	11.252	40.724	127.229	26.306	34.400
T <sub>AJ/TR</sub> (s)	–	13,72	1,43	2,42	25,64	2,14	3,03	42,30	2,91	3,05
T <sub>TOTAL</sub> (s)	191,56	66,15	18,34	28,99	61,11	18,18	26,91	88,18	23,12	17,60
Speed-up	–	2,90	10,45	6,61	3,13	10,54	7,12	2,17	8,28	10,89

#### 4. RESULTADOS NUMÉRICOS

A fim de analisar o desempenho da metodologia proposta neste artigo, com investigação do emprego de diferentes técnicas não supervisionadas de aprendizado de máquina (SOM, *K-means* e *K-medoids*), é avaliada a confiabilidade composta do sistema IEEE-RTS (*IEEE Reliability Test System*) (IEEE APM, 1979), para dois diferentes cenários de geração e carga. Os resultados obtidos pela metodologia proposta para cada caso são apresentados nesta seção e comparados aos resultados obtidos pela avaliação convencional de confiabilidade composta, realizada via SMC não sequencial convencional, aqui denominado SMC-NSC.

##### 4.1 Sistema IEEE-RTS Original

O sistema IEEE-RTS é formado por um conjunto de 32 geradores, distribuídos em 14 usinas, cuja capacidade total de geração original é de 3405 MW. A rede de transmissão do sistema é composta de 38 circuitos, que conectam 24 barras em dois níveis de tensão, 138 e 230 kV. A carga do sistema é representada por uma curva horária anual, composta de 8736 pontos, cujo pico é, originalmente, de 2850 MW. Considera-se que a carga distribuída ao longo da rede varia de forma correlacionada de acordo com a curva horária.

Para o cenário original de geração e carga do sistema IEEE-RTS, na execução do passo (i) da metodologia proposta (Fig. 1), foram definidos como críticos os circuitos da rede de transmissão que apresentaram carregamento superior ou igual a 80% após avaliação de fluxo de potência ótimo, como a minimização do custo de produção. Além disso, foram também considerados como críticos os circuitos com carregamento igual ou superior a 50% e que possuem pelo menos uma de suas barras terminais interligadas ao sistema por até dois circuitos, conforme estratégia determinística proposta na Subseção 3.1. De acordo com esses critérios, 6 circuitos foram classificados como críticos para o desempenho do sistema em termos de adequação estática. A indisponibilidade desses circuitos é monitorada no processo de avaliação da confiabilidade via metodologia proposta.

Na execução do passo (ii) da metodologia, foram gerados  $N = 700.000$  estados para o conjunto  $\Omega_N$ , sendo que este mesmo conjunto de estados é também utilizado para avaliação da confiabilidade via SMC-NSC. Os valores dos índices LOLP e EPNS obtidos, bem como o tempo total ( $T_{TOTAL}$ ) necessário para avaliação dos  $N$  estados via SMC-NSC, estão expostos

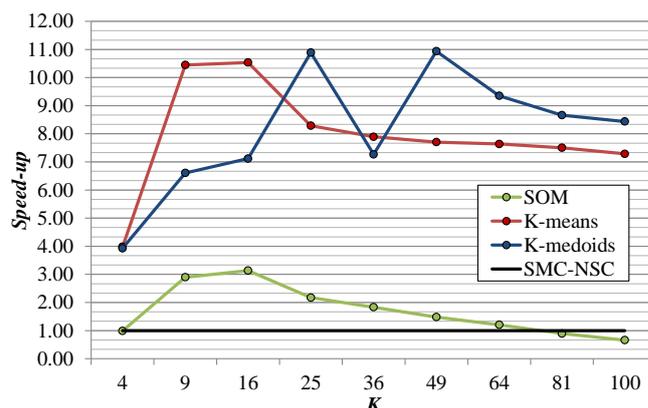


Fig. 3 Speed-up  $\times$  K – Sistema IEEE-RTS.

na Tabela 1. As incertezas de estimativa verificadas para estes índices, obtidos via SMC-NSC, são de  $\beta_{LOLP} = 3,57\%$  e  $\beta_{EPNS} = 4,88\%$ . Destaca-se, que os índices LOLE e EENS não são apresentados uma vez que podem ser facilmente calculados utilizando (4) e (5), respectivamente.

Na obtenção do conjunto  $\Omega_L$  para avaliação do sistema IEEE-RTS, conjunto utilizado no ajuste/treinamento das técnicas de classificação, definiu-se selecionar 0,25% do total dos  $N$  estados gerados, o que equivale a 1750 estados de  $\Omega_N$ . O filtro proposto na Subseção 3.1 foi empregado, com limites percentuais para o nível de carga e para a reserva de geração iguais, respectivamente, a 80% e 5% na etapa (a) e de 70% e 15% na etapa (b). Em relação à quantidade  $K$  de classes, três diferentes valores foram avaliados: 9, 16, e 25. Finalmente, os parâmetros da técnica SOM foram ajustados como: coeficientes inicial e final de aprendizagem iguais, respectivamente, a 0,1 e 0,0085; raio de vizinhança inicial igual à raiz quadrada de  $K$ .

Os resultados obtidos via metodologia proposta para o sistema IEEE-RTS original, com emprego das três técnicas não supervisionadas de classificação (SOM, *K-means*, e *K-medoids*), são também apresentados na Tabela 1. Os valores entre parênteses que acompanham os índices LOLP e EPNS para estes resultados correspondem às diferenças percentuais em relação aos respectivos valores destes índices obtidos via SMC-NSC. Para as três técnicas de classificação empregadas, são também apresentados na Tabela 1 o número realizado de avaliações de estado ( $N_{AVAIL}$ ), os tempos demandados na etapa de ajuste/treinamento ( $T_{AJ/TR}$ ) e total de avaliação ( $T_{TOTAL}$ ), e os *speed-ups* de cada ferramenta, os quais são calculados com base no tempo total da SMC-NSC.

Tabela 2. Resultados para o sistema IEEE-MRTS

	SMC-NSC	K = 9			K = 16			K = 25		
		SOM	K-means	K-medoids	SOM	K-means	K-medoids	SOM	K-means	K-medoids
LOLP	$4,233 \times 10^{-3}$	$4,153 \times 10^{-3}$ (1,89%)	$4,197 \times 10^{-3}$ (0,87%)	$4,217 \times 10^{-3}$ (0,39%)	$4,220 \times 10^{-3}$ (0,31%)	$4,157 \times 10^{-3}$ (1,81%)	$4,207 \times 10^{-3}$ (0,63%)	$4,220 \times 10^{-3}$ (0,31%)	$4,140 \times 10^{-3}$ (2,20%)	$4,210 \times 10^{-3}$ (0,55%)
EPNS (MW)	$6,349 \times 10^{-1}$	$6,333 \times 10^{-1}$ (0,26%)	$6,338 \times 10^{-1}$ (0,17%)	$6,340 \times 10^{-1}$ (0,14%)	$6,344 \times 10^{-1}$ (0,08%)	$6,304 \times 10^{-1}$ (0,72%)	$6,339 \times 10^{-1}$ (0,16%)	$6,344 \times 10^{-1}$ (0,08%)	$6,302 \times 10^{-1}$ (0,75%)	$6,339 \times 10^{-1}$ (0,16%)
N <sub>AVAIL</sub>	300.000	15.735	19.687	32.256	52.994	20.397	26.625	54.115	15.900	26.712
T <sub>AJ/TR</sub> (s)	–	6,67	0,68	0,67	12,64	1,04	0,89	19,68	1,23	1,57
T <sub>TOTAL</sub> (s)	88,89	29,80	24,92	27,89	44,65	25,26	26,77	51,95	24,34	27,49
Speed-up	–	2,98	3,57	3,19	1,99	3,52	3,32	1,71	3,65	3,23

É possível verificar que para todos os testes apresentados na Tabela 1, os índices LOLP e EPNS obtidos via metodologia proposta (para as três técnicas de classificação empregadas) apresentaram desvios percentuais abaixo das respectivas incertezas de estimativa ( $\beta$ ), com valores bastante reduzidos. Ou seja, foi mantida uma boa precisão na estimativa dos índices de confiabilidade composta em relação à avaliação realizada via SMC-NSC. Verifica-se, ainda, que os desvios percentuais relativos ao índice EPNS são menores que os desvios relacionados ao índice LOLP. Este fato mostra que os estados de falha não capturados com a avaliação realizada via metodologia proposta apresentam, em média, cortes de carga mais reduzidos, pertencendo, portanto, a uma região de operação marginal entre os diagnósticos de falha e de sucesso do sistema. Além da qualidade observada para a precisão dos índices, o processo de avaliação realizado pela metodologia proposta demandou a realização de um número mais reduzido de avaliações do sistema em relação ao  $N$  definido. Na média, para os resultados da Tabela 1, a metodologia proposta precisou avaliar por meio de fluxo de potência apenas 8,67% dos  $N$  estados gerados.

Em relação ao processamento computacional, é possível perceber para a metodologia proposta que excelentes *speed-ups* foram obtidos em relação à SMC-NSC, principalmente com emprego das técnicas *K-means* e *K-medoids*. Os maiores *speed-ups* apresentados por estas duas técnicas, quando comparadas à técnica SOM, são decorrentes, principalmente, dos menores tempos  $T_{AJ/TR}$  demandados por elas. Este tempo corresponde, basicamente, ao cumprimento dos passos (iii), (iv) e (vi) do algoritmo apresentado na Fig. 1. Pela Tabela 1 é possível ainda observar que o aumento do número de classes (valor de  $K$ ) demanda maiores valores de  $T_{AJ/TR}$  para as três técnicas empregadas. Contudo, o tempo  $T_{AJ/TR}$  demandado pela técnica SOM apresenta uma maior taxa de crescimento com o aumento de  $K$  em detrimento das demais técnicas investigadas. Este comportamento pode ser melhor visualizado no gráfico ilustrativo da Fig. 3, que apresenta o *speed-up* obtido pela metodologia proposta em relação à SMC-NSC para mais valores  $K$ , além daqueles inicialmente utilizados.

No gráfico da Fig. 3, é possível verificar que, para o sistema IEEE-RTS com cenário original de geração e carga, os melhores desempenhos obtidos pela metodologia proposta são observados para valores mais reduzidos de  $K$ . Verifica-se, ainda, que a partir de  $K = 81$ , se torna inviável o emprego da técnica SOM, enquanto que as técnicas *K-means* e *K-medoids* continuam a apresentar bons níveis de *speed-up*.

#### 4.2 Sistema IEEE-MRTS

De forma geral, pode-se considerar que o sistema IEEE-RTS com cenário original de geração e carga apresenta uma rede de transmissão “robusta”. Ou seja, em poucos casos, a falha de uma linha de transmissão ou de um transformador leva à necessidade de se realizar o corte de carga no sistema. Portanto, a fim de investigar o desempenho da metodologia proposta na avaliação de confiabilidade composta para sistemas com redes de transmissão mais estressadas, é considerado modificar o sistema IEEE-RTS. Neste sentido, o sistema IEEE-MRTS é obtido após modificação do sistema IEEE-RTS original, em que a capacidade de geração e o pico de carga do sistema são multiplicados por dois. Esta modificação faz com que estados de falha do sistema causados por indisponibilidades de equipamentos de transmissão sejam mais frequentes. Neste caso, também a ocorrência de estados de falha do sistema se torna menos rara, o que acelera a avaliação de confiabilidade mesmo considerando a avaliação via SMC-NSC, sendo assim necessário amostrar um menor número de estados para estimativa dos índices de confiabilidade.

Com o emprego da metodologia proposta, após avaliação do sistema conforme passo (i) do algoritmo apresentado na Fig. 1, e adoção dos mesmos critérios considerados para o sistema IEEE-RTS original, 13 circuitos são considerados como críticos para avaliação da confiabilidade do sistema IEEE-MRTS. O aumento do número de circuitos críticos evidencia, portanto, o maior estresse para a rede de transmissão.

Para um conjunto  $\Omega_N$  formado por  $N = 300.000$  estados amostrados, os índices LOLP e EPNS estimados pela SMC-NSC para o sistema IEEE-MRTS são apresentados na Tabela 2. Como esperado, percebe-se um aumento no valor destes índices em relação ao que foi verificado com o cenário original de geração e carga do sistema. As incertezas associadas às estimativas destes índices, avaliadas via (6), são de  $\beta_{LOLP} = 2,80\%$  e  $\beta_{EPNS} = 4,64\%$ .

Na avaliação da confiabilidade do sistema IEEE-MRTS via metodologia proposta, os parâmetros definidos para a técnica SOM e empregados no estudo anterior foram mantidos. Para composição do conjunto  $\Omega_L$  (passo (iii) da Fig. 1), dos  $N$  estados amostrados, 0,25% (i.e., 750 estados) foram selecionados conforme filtro proposto. Os resultados obtidos para 9, 16 e 25 classes estão presentes na Tabela 2.

Em geral, o desempenho em termos de precisão na estimativa dos índices foi mantido para o novo cenário de geração e carga, evidenciando, assim, que a metodologia proposta se adapta bem à nova característica do problema, de sistemas em que as dificuldades relacionadas à transmissão são mais recorrentes. Neste caso, foram avaliados por meio de fluxo de potência pelas três técnicas de classificação nos testes da Tabela 1, em média, 9,79% dos  $N$  estados amostrados para o problema.

Em relação ao desempenho computacional, um mesmo comportamento em relação ao estudo anterior (para o sistema IEEE-RTS original) foi verificado entre as três ferramentas de classificação empregadas. Neste caso, as técnicas *K-means* e *K-medoids*, por exemplo, apresentaram *speed-ups* superiores à técnica SOM, com valor médio de 3,41. Para todas as três ferramentas de classificação juntas, o *speed-up* médio verificado é de 3,02. É importante enfatizar que já eram esperados menores *speed-ups* para o sistema IEEE-MRTS em comparação com o IEEE-RTS original. Acontece que a proporção de estados de falha em relação ao número total de estados amostrados aumenta com as modificações realizadas (o que pode ser verificado pelo aumento do índice LOLP com o sistema IEEE-MRTS). Dessa forma, também dentre os estados avaliados pelas técnicas de classificação, a proporção de estados de falha é maior que no estudo anterior. Como o tempo demandado para avaliação de fluxo de potência ótimo da rede é maior para um estado de falha que para um estado de sucesso, devido às medidas remediadoras que visam minimizar o corte de carga, esperam-se ganhos mais reduzidos para o segundo sistema. Portanto, é possível concluir que para sistemas com eventos de falha mais raros, maiores benefícios podem ser verificados com emprego da metodologia proposta.

## 5. CONCLUSÕES

Uma metodologia simples e robusta é proposta neste artigo para acelerar o processo de avaliação da confiabilidade composta de sistemas elétricos de potência. Nessa metodologia, técnicas não supervisionadas de aprendizado de máquina são combinadas à simulação Monte Carlo não sequencial a fim de reduzir o esforço computacional envolvido no processo de estimativa dos índices de confiabilidade. A performance da metodologia é analisada para dois diferentes cenários de geração e de carga do sistema IEEE-RTS, sistema comumente empregado para avaliação da confiabilidade composta. Os desempenhos de três diferentes técnicas de classificação não supervisionada foram avaliados.

Os resultados obtidos apontaram um bom desempenho da metodologia proposta na avaliação da confiabilidade composta. Em todas as análises realizadas, foi mantida a qualidade da precisão de estimativa dos índices de confiabilidade em comparação com a avaliação convencional via método de simulação Monte Carlo não sequencial. Ao mesmo tempo, excelentes *speed-ups* em termos de custo computacional envolvido puderam ser verificados. Salienta-se, que excelentes resultados foram atingidos mesmo com a utilização de técnicas simples de classificação não supervisionada, que podem ser facilmente implementadas e empregadas.

Vale destacar, por fim, que os ganhos em termos de tempo computacional apresentados neste trabalho correspondem a uma única avaliação de confiabilidade composta. Contudo, na realização de estudos de planejamento de sistemas elétricos, a realização de um número elevado de avaliações pode se tornar necessária, acentuando, portanto, os benefícios inerentes ao emprego da metodologia proposta neste trabalho.

## REFERÊNCIAS

- Allan R.N. e Billinton, R. (1996). *Reliability evaluation of power systems*. (2nd), USA: Plenum, New York, NY.
- Benidris, M., Elsaiah, S. e Mitra, J. (2015). Power system reliability evaluation using a state space classification technique and particle swarm optimization search method. *IET Gener. Transm. Distrib.*, 9 (14), 1865-1873.
- Borges, C.L.T., Falcao, D.M., Mello, J.C.O. e Melo, A.C.G. (2001). Composite reliability evaluation by sequential Monte Carlo simulation on parallel and distributed processing environments. *IEEE Trans. Power Syst.*, 16 (2), 203-209.
- González-Fernández, R.A., Leite da Silva, A.M., Resende, L.C. e Schilling, M.T. (2013). Composite systems reliability evaluation based on Monte Carlo simulation and cross-entropy methods. *IEEE Trans. Power Syst.*, 28 (4), 4598-4606.
- IEEE APM (1979). IEEE reliability test system. *IEEE Trans. Power App. and Syst.*, PAS-98 (6), 2047-2054.
- Jain, A.K. (2010). Data clustering: 50 years beyond k-means”, *Pattern Recognition Letters*, 31, 651-666.
- Kohonen, T. (1982). Self-Organized formation of topologically correct feature maps. *Biol. Cybern.*, 43, 59-69.
- Leite da Silva, A.M., Rezende, L.S., Manso, L.A.F. e Resende, L.C. (2010). Reliability worth applied to transmission expansion planning based on ant colony. *System. Int. J Elec. Power*, 32 (10), 1077-1084.
- Luo, X., Singh, C. e Patton, A.D. (1999). Loss-of-Load State Identification Using Self-Organizing Map. in *IEEE Power Engineering Society Summer Meeting, Singapore*, 670-675.
- Luo, X., Singh, C. e Patton, A.D. (2000). Power System Reliability Evaluation Using Self Organizing Map. in *Proc. IEEE Power Engineering Society Winter Meeting, Singapore*, 1103-1108.
- Miloca, S.A., Volpi, N.M.P., Yuan, J. e Pinto, C.L.S. (2015). Expansion planning problem in distribution systems with reliability evaluation: an application in real network using georeferenced database. *Int. J Elec. Power*, 70, 9-16.
- Park, H.S. e Jun, C.H. (2009). A simple and fast algorithm for K-medoids clustering. *Expert Systems with Applications*, 36 (2), 3336-3341.
- Pereira, M.V.F. e Balu, N.J. (1992). Composite generation/transmission reliability evaluation. *Proceedings of the IEEE*, 80 (4), 470-491.
- Saraiva, J.T., Miranda V. e Pinto, L.M.V.G. (1996). Generation/transmission power system reliability evaluation by Monte-Carlo simulation assuming a fuzzy load description. *IEEE Trans. Power Syst.*, 11 (2), 690-695.
- Shu, Z., Jirutitijaroen, P., Leite da Silva, A.M. e Singh, C. (2014). Accelerated state evaluation and latin hypercube sequential sampling for composite system reliability assessment. *IEEE Trans. Power Syst.*, 29 (4), 1692-1700.
- Urgun, D. e Singh, C. (2019). A hybrid Monte Carlo simulation and multi label classification method for composite system reliability evaluation. *IEEE Trans. Power Syst.*, 34 (2), 908-917.