

Aplicação de Redes Neurais Artificiais com Tratamento de Erros via *Z-Score* em Sistemas de Distribuição de Energia Elétrica

Lucas Eduardo Silva Braga* Thales Schuabb de Almeida*
Pedro Caruso Fracetti* Leonardo Willer de Oliveira*
Edimar José de Oliveira*

* Faculdade de Engenharia Elétrica, Universidade Federal de Juiz de Fora, MG, (e-mails: braga.lucas@engenharia.ufjf.br, thales.schuabb@engenharia.ufjf.br, caruso.fracetti@engenharia.ufjf.br, leonardo.willer@ufjf.edu.br, edimar.oliveira@ufjf.edu.br)

Abstract: The lack of measurements is a characteristic of Electrical Energy Distribution Systems (EEDS) which requires State Estimation (SE) methods able to check precisely the electrical network operation state, so SE methods have been studied including Artificial Neural Networks (ANN). The ANN is a computational intelligence tool able to have agility in information processing which means more useful information, such as estimated state variables of EEDS. This potential is investigated in this paper, where an ANN provides polar coordinates of nodal tension variables from available measurements and pseudo-measurements. The proposed methodology also includes the *Z-Score* algorithm approach to bad data detection on measures, which will be treated to avoid bad estimation quality.

Resumo: Características de Sistemas de Distribuição de Energia Elétrica (SDEE), como carência de medição, requerem métodos de Estimação de Estados (EE) capazes de aferir o estado da rede elétrica com precisão satisfatória. Com isso, métodos têm sido propostos para estimadores de estados, incluindo Redes Neurais Artificiais (RNA), que são ferramentas de inteligência computacional capazes de processar dados com agilidade a fim de prover informação útil, como estimativas para variáveis de estado em um SDEE. Este potencial é investigado no presente trabalho, cuja aplicação proposta envolve uma RNA que fornece estimativas para as variáveis de tensão nodal em coordenadas polares, a partir de medidas disponíveis e pseudo-medidas. O trabalho inclui a aplicação do algoritmo *Z-Score* para detecção prévia de erros grosseiros nas medições, bem como para tratamento desses erros a fim de evitar perda na qualidade das estimativas obtidas.

Keywords: State Estimation; Artificial Neural Network; Distribution Systems; Bad Data; Z-score.

Palavras-chaves: Estimação de Estados; Redes Neurais Artificiais; Sistemas de Distribuição; Erros Grosseiros; Z-score.

1. INTRODUÇÃO

O monitoramento de sistemas de distribuição de energia elétrica (SDEE), a partir de informações quanto aos níveis de tensão, é importante para a operação e os respectivos processos decisórios, como chaveamentos, reconfiguração, entre outros, visando a uma boa qualidade no fornecimento de energia e aumento da confiabilidade. Com isso, vários Métodos estão sendo estudados com o objetivo de realizar a estimação dos estados (EE) de SDEE. Um dos mais utilizados é o método dos mínimos quadrados (WLS) (Liu et al., 2019), (Pau et al., 2017), que se baseia em processo numérico iterativo. Além do WLS, outras abordagens vêm ganhando importância, como as baseadas em modelos de otimização, conforme a referencia (Duque et al., 2017),

* Os autores deste artigo agradecem ao apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, FAPEMIG, CAPES, INERGE, e aos grupos GOPT e SMART4GRIDS.

que apresenta um Fluxo de Potência Ótimo Estendido (E-OPF), que demonstrou eficiência na estimação, mesmo com uma quantidade baixa de medidas disponíveis. Também é notório o crescente uso de técnicas de inteligência computacional, como as Redes Neurais Artificiais (RNA) (Haykin, 1999) (Zamzam et al., 2019), (Zhang et al., 2019), (Mestav et al., 2019), (Cao et al., 2020), (Basumallik et al., 2019), (Ashraf et al., 2017).

Um dos desafios para aplicação de um estimador de estados em tempo hábil para a operação de um SDEE consiste na dimensão das redes elétricas reais. Quanto maior a dimensão, maior é o número de variáveis a serem estimadas e, conseqüentemente, maior é o esforço computacional para processamento dessas variáveis, podendo afetar a eficiência computacional de métodos de EE. Múltiplas pesquisas abordam alternativas para aumentar ou manter a eficiência mencionada em padrões aceitáveis para os

sistemas gerenciadores de energia. No trabalho de (Pau et al., 2017), por exemplo, é proposta a divisão da rede de distribuição em diversas áreas, o que reduz o esforço computacional.

No trabalho de (Majidi et al., 2017), é proposto um método que utiliza a matriz de impedâncias e os fasores de tensões nodais para obter um conjunto de equações suficiente para o processo de estimação de estados. Segundo (Carquex et al., 2018) é apresentado um algoritmo baseado no método PASE em conjunto com filtro Kalman para realizar a EE de SDEE. No trabalho de (Zhang and Wang, 2020), um algoritmo baseado na Série de Taylor de números complexos é proposto para a EE. De modo geral, questões como a baixa capacidade de adaptação a mudanças topológicas da rede elétrica, comuns em SDEE, bem como a perda de qualidade devido à erros grosseiros nas medidas, são desafios para o processo de estimação de estados.

No cenário descrito anteriormente, as Redes Neurais Artificiais (RNA) são potenciais para aplicação ao problema de EE, devido à sua capacidade de rápido processamento e, portanto, boa eficiência computacional. Além disso, têm a capacidade de generalização a partir de uma base de dados conhecida, o que significa que podem responder adequadamente a dados desconhecidos a partir de uma base amostral previamente utilizada no aprendizado, segundo (Haykin, 1999). Essa característica confere flexibilidade diante de variações topológicas e de condições operativas. No trabalho de (Zamzam et al., 2019), uma RNA é treinada com pseudo-medidas e dados históricos, com o objetivo de aprender a mapear medidas disponíveis para a EE. Em (Zhang et al., 2019), o treinamento de uma RNA com pseudo-medidas confere capacidade de melhoria das estimativas na saída.

O trabalho de (Mestav et al., 2019) utiliza dados de baixa precisão para treinar uma RNA profunda, capaz de detectar dados irrelevantes antes do processo de estimação de estados. Na referência (Cao et al., 2020), uma RNA dedicada a EE passa por um processo de aprendizado com uma série de dados, como os de carga na rede. Em (Basumallik et al., 2019), uma RNA é treinada com amostras que incluem dados falsos, a fim de capacitá-la na detecção de ataques cibernéticos que prejudicam a EE. No trabalho de (Ashraf et al., 2017), o treinamento da RNA é realizado com fasores de tensão, obtidos através de Unidades de Medição Fasorial, do inglês *Phasor Measurement Unit* (PMU), disponíveis em barramentos da rede elétrica, para estimar margens de carregamento do SDEE, visando aumentar a precisão na avaliação de estabilidade do sistema.

O presente trabalho pretende investigar a aplicação de uma RNA ao processo de estimação de estados, a partir de medidas de módulo e fase de tensão obtidas de PMUs, ou de módulo de tensão obtidas de Medidores Inteligentes (MI), no SDEE. Os MI são economicamente mais atrativos do que PMUs para um SDEE e, portanto, é importante avaliar o efeito na EE quando MIs ou PMUs são considerados. A RNA proposta para aplicação é do tipo *MultiLayer Perceptron* e opera com Aprendizado Supervisionado (AS). Como contribuição em relação a (Ahmad et al., 2019), tem-se que a estrutura de RNA proposta não requer pré-processamento de medidas para obtenção

de potências nodais, evitando a propagação de erros neste processamento. Adicionalmente, o trabalho visa o estudo do impacto de erros grosseiros de medidas nas métricas comumente utilizadas para avaliação da acurácia do processo de EE. Por fim, um algoritmo da literatura, denominado *Z-score* (Ashraf et al., 2017), é desenvolvido e aplicado para detecção e tratamento dos erros grosseiros. A contribuição específica do presente trabalho é aplicar RNA para estimação de estados, considerando medidas de tensão com tratamento de erros via *Z-score*, no contexto de SDEE, o que é, a partir do conhecimento dos autores, ainda inexplorado na literatura. A adequação para a distribuição, além de envolver MI entre as opções, também inclui adaptação no tratamento da carga para considerar níveis de incertezas típicos de SDEE. Sistemas teste de distribuição conhecidos são utilizados para a avaliação da aplicação e os resultados permitem verificar a aplicabilidade da abordagem proposta.

Além desta seção introdutória, o presente artigo tem outras quatro seções. A Seção 2 apresenta uma fundamentação teórica sobre o processo de estimação de estados. Já a Seção 3 descreve a aplicação de RNA proposta, envolvendo sua topologia, cenários de carga, conjunto amostral, tipos de medidas consideradas e o método utilizado para o tratamento de erros grosseiros. Na Seção 4, são apresentados os estudos de casos envolvendo sistemas teste da literatura e os resultados obtidos. Por fim, as conclusões do trabalho são abordadas na Seção 5.

2. FUNDAMENTAÇÃO TEÓRICA

O problema de Estimação de Estados é avaliado com o auxílio de índices ou métricas. Uma métrica muito utilizada consiste no erro quadrático formulado em (1), aqui denominado *EEE* (Erro de Estimação de Estados) (Tecchio et al., 2010). Além do índice *EEE*, o presente trabalho também avalia mais três métricas propostas por (Duque et al., 2017):

- *EMT* (Erro de Módulo de Tensão), formulado em (2);
- *EAF* (Erro de Ângulo de Fase), conforme (3);
- *IGE* (Índice Geral de Estimação), dado pelo somatório das métricas *EEE*, *EMT* e *EAF*, conforme (4).

$$EEE = \sqrt{\sum_{i=1}^k (V_i^* - \widehat{V}_i)^2 + \sum_{i=1}^k (\theta_i^* - \widehat{\theta}_i)^2} * 100 \quad (1)$$

$$EMT = \max_k \left| \frac{|V_k^*| - |\widehat{V}_k|}{|V_k^*|} \right| * 100 \quad (2)$$

$$EAF = \max_k \left| |\theta_k^*| - |\widehat{\theta}_k| \right| * 100 \quad (3)$$

$$IGE = EEE + EMT + EAF \quad (4)$$

Em que:

- V_i^* é o módulo de tensão real da barra i .
- \widehat{V}_i é o módulo de tensão estimado da barra i .
- θ_i^* é o ângulo de fase real da tensão da barra i .
- $\widehat{\theta}_i$ é o ângulo de fase estimado da tensão da barra i .

Verifica-se, portanto, que o índice *IGE* representa um somatório de todas as métricas de erros. Esse índice, porém, generaliza o resultado, sem dar sensibilidade sobre o erro individual dos barramentos do SDEE. Logo, são utilizados, de forma complementar, os índices *EMT* e *EAF*, que trazem informação sobre a pior estimativa, que pode ser útil para mapeamento da fonte de erros no processo de EE.

3. METODOLOGIA PROPOSTA

Esta seção apresenta os principais aspectos da rede neural artificial proposta para aplicação, incluindo os dados utilizados para treinamento da RNA e sua arquitetura. As amostras utilizadas para treinamento, validação e teste envolvem diferentes cenários de carga do SDEE, com o propósito de verificar a capacidade de generalização da RNA. Após o treinamento, a RNA é testada com amostras que incluem erros grosseiros em determinadas medidas, a fim de avaliar o impacto desses erros na resposta da RNA, bem como nas métricas de EE anteriormente definidas. Por fim, aplica-se o algoritmo *Z-score* ao conjunto amostral de teste para detectar e tratar os erros introduzidos, com reavaliação das métricas de EE. Estes aspectos são detalhados nas subseções a seguir.

3.1 Cenários de Carga

A obtenção dos dados de treinamento, validação e teste da RNA é feita a partir de um gerador de cenários de carregamento do SDEE e cálculo de fluxo de potência para cada cenário gerado. Os cenários são obtidos a partir da modelagem de funções de distribuição normais de probabilidade para as potências ativas de cada barra do SDEE, cuja média ou esperança é o valor obtido da literatura para a barra em questão, denominado como caso base no presente trabalho. O desvio padrão considerado é de 13,33% em relação à média, de modo que 99,7% dos valores estejam dentro do intervalo de $\pm 40\%$ da média, seguindo a definição de pseudo-medidas de (Duque et al., 2017).

Destaca-se que o trabalho apresentado em (Ashraf et al., 2017) considera o percentual de 30% de incerteza para a demanda de sistema de transmissão de energia elétrica. Portanto, o nível de incerteza para a distribuição é atualizado para 40%, conforme recomendado por (Duque et al., 2017).

Logo, para cada cenário ou amostra, um valor de potência ativa é sorteado dentro da distribuição normal de probabilidades de cada barra, conforme (5). Então, a amostra do valor de potência reativa correspondente é obtida de forma a manter constante o fator de potência de cada barra, conforme (6).

$$P_{i,c} = NP_i \pm 3\sigma \quad \forall c \in Nc \quad (5)$$

$$Q_{i,c} = P_{i,c} \cdot \tan(\arccos(\Psi_i)) \quad \forall c \in Nc \quad (6)$$

Em que:

- $P_{i,c}$ é a potência ativa da barra i em um dado cenário amostral c ;

- $Q_{i,c}$ é a potência reativa da barra i no cenário c ;
- NP_i é a potência ativa nominal da barra i (caso base obtido da literatura);
- σ é o desvio padrão;
- Ψ_i é o fator de potência constante da barra i .

A partir dos valores de $P_{i,c}$ e $Q_{i,c}$ gerados para Nc cenários, são realizados Nc cálculos de fluxo de potência, sendo um para cada cenário c . Dessa forma, são obtidos os valores de $V_{i,c}$ e $\theta_{i,c}$ para cada cenário c entre os Nc cenários, sendo esses os valores de interesse para treinamento, validação e teste da RNA, uma vez que as medidas consideradas são de módulo de tensão e ângulo de fase de tensão, quando provenientes de PMU, e apenas módulo de tensão quando provenientes de MI. Destaca-se que esse procedimento evita o cálculo de potências nodais a partir das medidas de tensão em módulo e fase obtidas de medidores, como é feito em (Menke et al., 2019), evitando a propagação de erros destas medidas devido ao processamento para a obtenção das referidas potências.

3.2 Conjunto Amostral

Os Nc valores de $V_{i,c}$ e $\theta_{i,c}$, para toda barra i do SDEE, conforme descrito na subseção anterior, são divididos em três conjuntos:

- 50% dos Nc cenários - Conjunto de Treino;
- 25% dos Nc cenários - Conjunto de Validação;
- 25% dos Nc cenários - Conjunto de Teste.

O maior conjunto, de treinamento, é aquele a partir do qual a RNA aprende a associação entre os padrões de entrada e de saída do sistema. O segundo conjunto, de validação, é utilizado como critério de convergência para o treinamento, em que um valor N_{vc} é estabelecido como limite de vezes em que a precisão da RNA, durante o treinamento, piora para o conjunto de validação. Caso o limite N_{vc} seja alcançado, o critério mencionado, conhecido como “validação cruzada”, é atendido e, então, o treinamento é finalizado.

O último conjunto definido, de teste, utiliza as 25% amostras restantes, com as quais a RNA não tem contato em seu treinamento ou validação. Esses dados de teste, por serem, inéditos para a RNA, são úteis para a avaliação da capacidade de generalização da rede já treinada. Destaca-se que no presente trabalho, o conjunto de testes é utilizado para avaliar o comportamento da RNA na presença de erros grosseiros nas medições, o que será tratado mais a frente na Seção 4.2. Ou seja, erros grosseiros são introduzidos em algumas amostras selecionadas do conjunto de teste para essa avaliação.

3.3 Tipos de Medidas Consideradas

No presente trabalho, são utilizadas medidas de V_i e/ou θ_i para nas barras i do SDEE em que há disponibilidade de medidores. Os medidores do tipo PMU, apresentam capacidade de medição de módulo e de ângulo de fase da tensão nodal. A precisão de suas medidas, tanto de módulo de tensão, δ_V , quanto de ângulo de fase, δ_θ , é de $\pm 0,4\%$, conforme (Abur and Exposito, 2004). Os medidores do tipo MI, são capazes de medir apenas o módulo de tensão, com uma precisão de $\pm 2\%$ (Peppanen et al., 2014)

Para a avaliação do estimador de estados via RNA proposto, considera-se o pior caso para o erro nas medidas, ou seja, considera-se que todas estejam em um dos limites, de acordo com o tipo de medidor, conforme a seguir:

- PMUs : $\delta_V = \delta_\theta = +0,4\%$;
- MIs : $\delta_V = +2\%$ para MI.

Além de retratar o pior caso, este procedimento evita o ruído branco ou gaussiano que seria obtido caso fossem consideradas medidas com erro de $+\delta\%$ ou $-\delta\%$ (Wang et al., 2017). Com a inclusão dos erros nas medidas, obtêm-se os valores de entrada da RNA proposta, V_i^{err} e θ_i^{err} , conforme formulado em (7) e (8).

$$V_i^{err} = V_i + \delta_V \quad (7)$$

$$\theta_i^{err} = \theta_i + \delta_\theta \quad (8)$$

3.4 Rede Neural Artificial Proposta

O processo de aprendizagem da RNA proposta é feito de forma supervisionada através do algoritmo denominado *backpropagation*. A topologia dessa RNA é conhecida como *MultiLayer Perceptron* (MLP). O conjunto de medidas provenientes dos medidores, passando pela aplicação de erros formulada anteriormente, é utilizado como entrada da RNA. O conjunto de saída consiste em vetores com $2 * N_b$ elementos, em que N_b é o número de barras do sistema, cujos valores incluem os módulos de tensão estimada (\hat{V}_i) e de ângulo de fase estimado ($\hat{\theta}_i$) nas N_b barras i do SDEE.

A arquitetura da RNA proposta é representada pela Figura 1, em que V_{mi} e θ_{mi} são as medidas de módulo de tensão e ângulo de fase, respectivamente, da barra i . Foi utilizada uma arquitetura de RNA presente em toolbox do ambiente de programação *Python*, mais precisamente, a toolbox *Scikit-Learn*. Durante o treinamento da RNA, são checados os seguintes critérios de convergência: número máximo de validações cruzadas N_{vc} e número máximo de iterações ou épocas N_{gm} . O treinamento é encerrado quando pelo menos um desses critérios é atendido. A configuração da RNA, conforme anteriormente descrita, é resumida na Tabela 1.

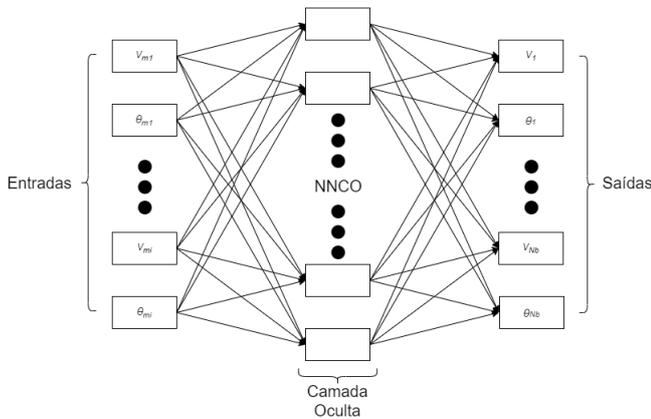


Figura 1. Rede *MultiLayer Perceptron* Utilizada

O número de neurônios da camada oculta ($NNCO$) é obtido através de (9), em que NE representa o número de

Tabela 1. Configuração da RNA

Parâmetro	Configuração
Função de Ativação	Logística
Tolerância	$1e^{-9}$
Taxa de Aprendizagem	$1e^{-3}$
Número Máx de Validações Cruzadas (N_{vc})	500
Número Máximo de Épocas (N_{gm})	100000

entradas da RNA e NS o número de saídas. Esses valores, por sua vez, são dados por (10) e (11), respectivamente, em que N_{bm} é o número de barras com medição e N_b é o número total de barras do sistema, conforme anteriormente definido.

O erro médio quadrático utilizado como métrica para treinamento e validação da RNA proposta corresponde ao índice relativo ao erro de estimação de estados, EEE , formulado em (1).

Destaca-se que o valor de $NNCO$ é diferente do valor proposto em (Menke et al., 2019), pois o número de saídas conforme metodologia proposta, envolvendo todas as tensões nodais em módulo e fase, é maior do que o número de saídas de (Menke et al., 2019) que, por sua vez, fornece apenas os módulos de tensão. Logo, no presente trabalho, o número de saídas NS é dividido por 3 em (9), a fim de evitar número excessivo de neurônios na camada oculta e o comprometimento da eficácia da RNA proposta. Neste caso, o fator '1/3' utilizado em (9) foi determinado de forma empírica.

$$NNCO = \frac{2}{3} * NE + \frac{1}{3} * NS \quad (9)$$

$$NE = 2 * N_{bm} \quad (10)$$

$$NS = 2 * N_b \quad (11)$$

3.5 Tratamento de Erros Grosseiros

A detecção e o tratamento de erros grosseiros em medidas é relevante para evitar que dados ruins prejudiquem o processo de estimação de estados. No presente trabalho, a detecção de erros grosseiros é feita através do algoritmo *Z-Score*. Este algoritmo estatístico tem a função de verificar se um determinado valor está dentro ou fora da média de um conjunto de dados em que está contido. Para realizar o cálculo do *score*, basta tomar um valor da medição, subtrai-lo da média do conjunto de amostras anteriores e dividir o resultado pelo desvio padrão do mesmo conjunto, conforme formulado em (12).

$$Z = \frac{(x - \mu)}{\sigma} \quad (12)$$

Em que:

- Z é a pontuação do *Z-score*;
- x é o valor de determinada medida;
- μ é a média das medidas anteriores a x ;
- σ é o desvio padrão das medidas anteriores a x .

A pontuação Z da medida x analisada é então comparada com um valor limiar pré-estabelecido (Z_l), e caso esteja

acima deste valor, a medida x é caracterizada por apresentar um erro grosseiro. Essa medida é, então, substituída pela média das amostras anteriores (μ) para a realização da EE.

4. ESTUDO DE CASOS

4.1 Descrição dos Casos de Estudo

A metodologia proposta foi aplicada no sistema de teste conhecido na literatura, contendo 14 barras (Civanlar et al., 1988). Esse sistema é ilustrado na Figura 2.

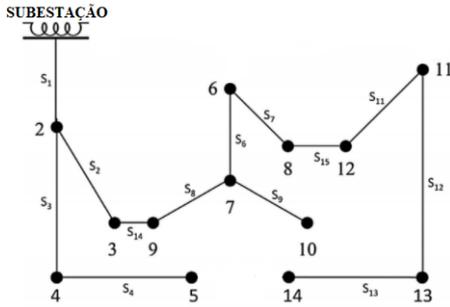


Figura 2. Sistema de 14 Barras

Para esse sistema, foram obtidos 10000 cenários de carga ($N_c = 10000$). As seguintes análises são realizadas:

- **Análise-1:** Condição de monitoramento mínimo considerada, com uma PMU na barra determinada pelo método de alocação ótima de (Duque et al., 2017);
- **Análise-2:** Condição de monitoramento máximo considerada, com quatro PMUs alocadas conforme (Duque et al., 2017);
- **Análise-3:** Similar à Análise-1, porém, considerando MI;
- **Análise-4:** Similar à Análise-2, porém, considerando MI;
- **Análise-5:** Similar à Análise-3, porém, com a inserção de erros grosseiros;
- **Análise-6:** Similar à Análise-4, porém, com a inserção de erros grosseiros;
- **Análise-7:** Similar à Análise-5, porém, com a identificação e o tratamento de erros grosseiros via algoritmo Z -score;
- **Análise-8:** Similar à Análise-6, porém, com a identificação e o tratamento de erros grosseiros via algoritmo Z -score.

Em todas as análises anteriores, as métricas para o processo de EE, descritas no presente artigo, são avaliadas. As Análises-5 e 6 visam avaliar o impacto de erros nas condições de mínimo (1 medição) e máximo monitoramento (4 medições) consideradas, respectivamente, através de MI. As Análises-7 e 8, por sua vez, têm por finalidade identificar a eficácia do tratamento via algoritmo Z -score, nas condições de mínimo e máximo monitoramento, respectivamente, através de MI. O valor limiar de pontuação considerado no algoritmo Z -score (Z_l) é igual a 38. Este valor foi escolhido de forma empírica, sendo suficiente para identificar erros grosseiros. As barras monitoradas por PMUs ou MIs, de acordo com a análise, estão resumidas

na Tabela 2. O monitoramento mínimo conta com apenas um medidor alocado, enquanto o máximo utiliza quatro medidores.

Tabela 2. Alocação da medição

Monitoramento	Barras
Mínimo	11
Máximo	2-10-12-13

4.2 Resultados

Para fins de comparação com a literatura, utilizou-se a referência (Duque et al., 2017). Na avaliação da RNA já treinada, obtém-se a média (\hat{x}) e o desvio padrão (σ) das métricas obtidas para as 250 amostras de teste. Este procedimento é necessário para considerar a incerteza inerente às medidas. Cada valor médio (\hat{x}) é comparado com o respectivo valor de (Duque et al., 2017), que já considera a incerteza das medidas de forma intrínseca.

A Tabela 3 apresenta os resultados obtidos para as Análises-1 a 4, que consideram apenas os erros intrínsecos aos medidores, ou seja, que não consideram presença de erros grosseiros.

Tabela 3. Análises sem erros grosseiros

Análise-1	IGE	EMT	EAF	EEE
Desvio Padrão (σ)	0,35427	0,14135	0,069457	0,20054
Média (\hat{x})	1,26853	0,38246	0,16100	0,72506
(Duque et al., 2017)	1,76873	0,42098	0,03754	1,31020
Análise-2	IGE	EMT	EAF	EEE
Desvio Padrão (σ)	0,45010	0,14653	0,06220	0,28018
Média (\hat{x})	1,25406	0,38209	0,14765	0,72431
(Duque et al., 2017)	1,74393	0,40046	0,02674	1,31673
Análise-3	IGE	EMT	EAF	EEE
Desvio Padrão (σ)	0,37392	0,07911	0,10667	0,30630
Média (\hat{x})	7,38548	2,11659	0,59942	4,66946
Análise-4	IGE	EMT	EAF	EEE
Desvio Padrão (σ)	0,20836	0,06921	0,07870	0,12953
Média (\hat{x})	7,31790	1,91417	0,69445	4,70927

Observa-se da Tabela 3 que a RNA proposta fornece médias (\hat{x}) para o indicador relativo ao erro de estimação de estados, EEE formulado em (1), inferiores aos respectivos valores de (Duque et al., 2017) para as Análises-1 e 2, o que demonstra uma boa característica da RNA. Estas análises são passíveis de comparação com a literatura (Duque et al., 2017), que também utiliza PMUs para a estimação de estados. O índice EEE é utilizado como base de comparação por ter sido a métrica utilizada no treinamento da RNA. No entanto, é possível também fazer uma análise com relação aos demais índices, de onde se pode notar que as médias (\hat{x}) para o índice geral de estimação, IGE , definido em (4), obtidas nas Análises-1 e 2 são inferiores aos respectivos indicadores de (Duque et al., 2017). A Tabela 3 também apresenta os desvios padrão (σ) obtidos nas análises propostas. Conforme esperado, observa-se ainda que o aumento do número de medidores implica em melhoria de qualidade da estimativa, ou seja, redução do índice IGE , como se pode verificar da Análise-1 para a Análise-2, bem como da Análise-3 para a Análise-4.

As Análises-3 e 4, por sua vez, apresentam os índices obtidos com o monitoramento sendo efetuado por MI.

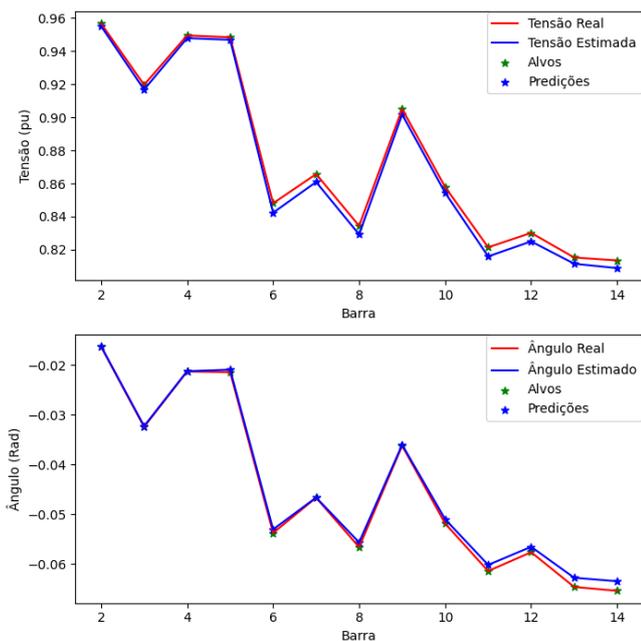


Figura 3. Análise-1: Monitoramento mínimo por PMU

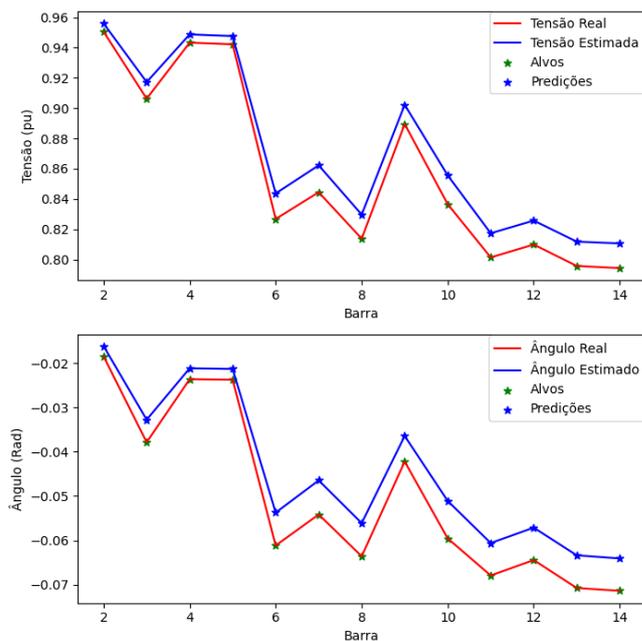


Figura 5. Análise-3: Monitoramento mínimo por MI

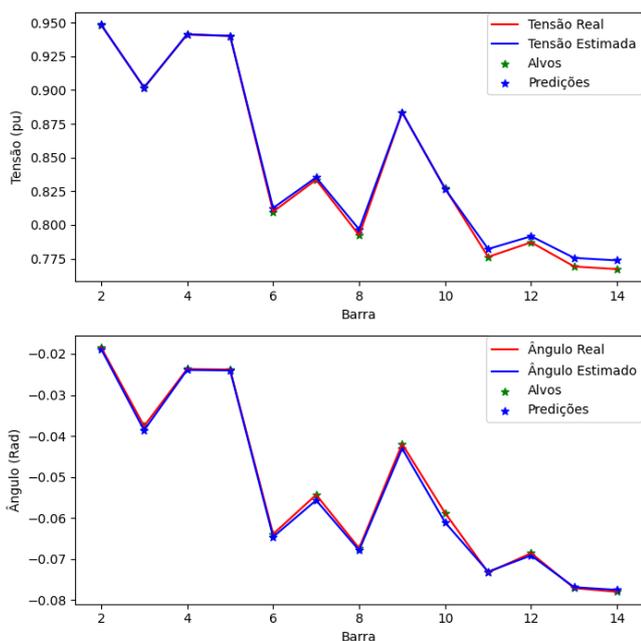


Figura 4. Análise-2: Monitoramento máximo por PMU

Como a precisão destes medidores é mais baixa do que a de PMU, os índices da Análise-3 são piores, ou seja, maiores, do que os respectivos índices da Análise-1 (ambas para condição de monitoramento mínimo). O mesmo pode ser observado comparando-se os indicadores das Análises-2 e 4 para monitoramento máximo.

Embora o comportamento anteriormente descrito já fosse esperado, o presente trabalho tem como objetivo avaliar o impacto da utilização de MI na distribuição, ao invés de PMUs.

Essa avaliação é possível tanto em termos numéricos, através dos resultados da Tabela 3, quanto em termos gráficos,

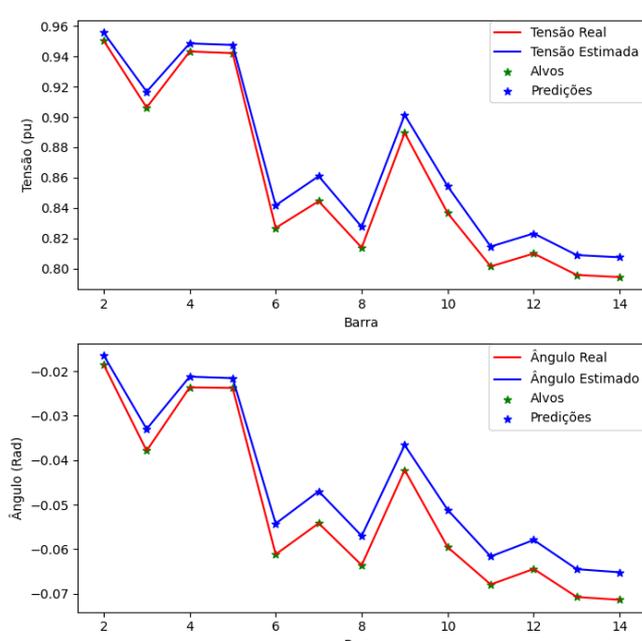


Figura 6. Análise-4: Monitoramento máximo por MI

através das Figuras 3 a 6. Estas figuras correspondem às piores amostras de teste, entre as 250 amostras utilizadas para obter a média (\hat{x}) e o desvio padrão (σ) em cada análise. Ou seja, as figuras ilustram o comportamento dos piores casos, ou amostras de teste, em cada análise.

Das figuras anteriormente citadas, observa-se que as curvas de módulo de tensão estimadas e reais são praticamente coincidentes nas Figuras 3 e 4. O mesmo ocorre com as curvas de ângulo de fase nestas mesmas figuras. Por outro lado, ao se observar as curvas de módulo de tensão nas Figuras 5 e 6, já se pode visualmente verificar uma divergência entre as curvas reais e estimadas, devido à menor

precisão dos MI em relação às PMUs. Essa divergência é mais acentuada para a barra '14' na Figura 6.

Nas Análises-5 e 6, o conjunto de teste é infectado com cinco medições que apresentam erros grosseiros de forma aleatória. O objetivo é analisar o impacto desses erros na EE via RNA. Os resultados da Tabela 4 mostram uma piora, ou seja, aumento significativo dos indicadores de EE, demonstrando o prejuízo causado pela existência de erros grosseiros nas medidas. As amostras com piores métricas para essas análises são ilustradas nas Figuras 7 e 8, de onde se pode visualmente comprovar o impacto substancial de erros grosseiros, a partir da divergência entre as curvas reais e estimadas.

Tabela 4. Análise da detecção e tratamento de erros grosseiros

Análise-5	IGE	EMT	EAF	EEE
Desvio Padrão (σ)	47,47335	14,05151	2,75187	30,95350
Média (\bar{x})	13,76909	3,99043	0,97321	8,80544
Análise-6	IGE	EMT	EAF	EEE
Desvio Padrão (σ)	42,63978	13,59877	5,62211	23,44119
Média (\bar{x})	12,55008	3,56982	1,39057	7,58968
Análise-7	IGE	EMT	EAF	EEE
Desvio Padrão (σ)	0,66252	0,17955	0,11328	0,46289
Média (\bar{x})	7,36097	2,10779	0,59884	4,65433
Análise-8	IGE	EMT	EAF	EEE
Desvio Padrão (σ)	0,63666	0,16331	0,09629	0,42532
Média (\bar{x})	7,26871	1,90353	0,69107	4,67410

Na Análises-7 e 8, o algoritmo *Z-score* de detecção de *outliers*, ou seja, de dados discrepantes, foi capaz de identificar os erros grosseiros e tratá-los, visando manter a integridade do processo de EE. Desta forma, as piores métricas foram similares às Análises-3 e 4, que não apresentam erros grosseiros, Figuras 5 e 6. Na Tabela 4, é possível verificar numericamente o efeito da aplicação do algoritmo *Z-score*, a partir da comparação dos índices da Análise-7 com os índices da Análise 5 (monitoramento mínimo), bem como comparando os indicadores da Análise-8 com aqueles da Análise-6 (monitoramento máximo). Esses resultados demonstram a aplicabilidade do algoritmo de detecção e tratamento de erros *Z-score* grosseiros.

5. CONCLUSÃO

O presente trabalho demonstrou que há aplicabilidade de redes neurais artificiais ao problema de estimação de estados de redes de distribuição, com pouca disponibilidade de medidas. Foram consideradas medidas de módulo e fase de tensão obtidas de unidades de medição fasorial, bem como de medidores inteligentes, de menor custo de investimento e mais adequados ao contexto da distribuição. A rede neural proposta não requer pré processamento das medidas disponíveis e, com isso, evita a propagação de erros. Para o caso do monitoramento com PMU, a rede neural proposta obteve métricas melhores do que a referência (Duque et al., 2017). O uso de MI piora a estimativa em relação às unidades fasoriais, com aumento na média e no desvio padrão do índice de estimação, como esperado, devido ao maior erro associado e ao menor número de informações para o algoritmo. No entanto, este tipo de equipamento é mais frequente em SDEE. É importante destacar que a variação na quantidade de medidores disponíveis não

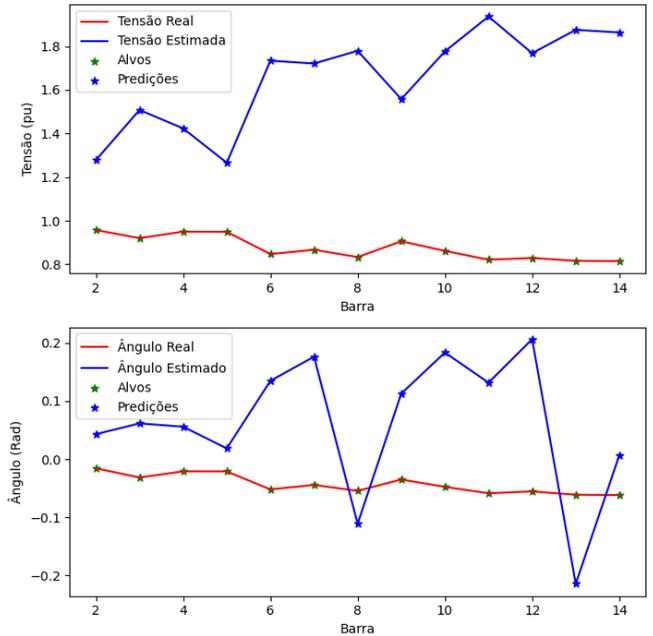


Figura 7. Análise-5: Erro Grosseiro para 1 MI

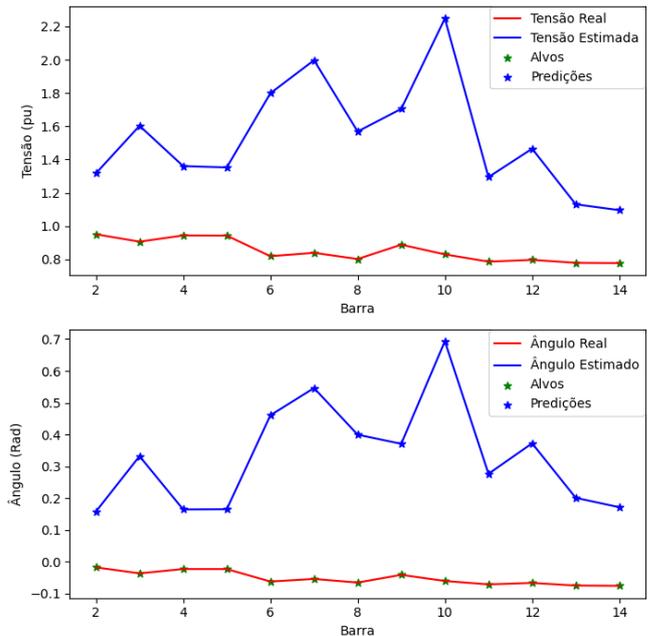


Figura 8. Análise-6: Erro Grosseiro para 4 MI

representou uma melhora significativa nos índices de estimação, com uma pequena variação numérica na média. Os resultados indicam que a RNA é uma boa candidata para a estimação de estados, já que os índices de estimação para o sistema teste conhecido da literatura foram menores que a referência (Duque et al., 2017).

Outra contribuição do trabalho foi mostrar o impacto negativo que erros grosseiros nas medições causam na estimação de estados. O algoritmo *Z-score* detectou com precisão os *outliers*, ou seja, dados discrepantes, e o tratamento de dados utilizado conseguiu evitar o prejuízo decorrente para o estimador proposto.

AGRADECIMENTOS

Os autores deste artigo agradecem ao apoio da FAPEMIG, CNPq, CAPES, INERGE, e ao grupo SMART4GRIDS - Systems Markets, Analytics Renewables Transition for Grids.

REFERÊNCIAS

- Abur, A. and Exposito, A.G. (2004). *Power system state estimation: theory and implementation*. CRC press.
- Ahmad, F., Tariq, M., and Farooq, A. (2019). A novel ann-based distribution network state estimator. *International Journal of Electrical Power & Energy Systems*, 107, 200–212.
- Ashraf, S.M., Gupta, A., Choudhary, D.K., and Chakrabarti, S. (2017). Voltage stability monitoring of power systems using reduced network and artificial neural network. *International Journal of Electrical Power & Energy Systems*, 87, 43–51.
- Basumallik, S., Ma, R., and Eftekharnejad, S. (2019). Packet-data anomaly detection in pmu-based state estimator using convolutional neural network. *International Journal of Electrical Power & Energy Systems*, 107, 690–702.
- Cao, Z., Wang, Y., Chu, C.C., and Gadh, R. (2020). Scalable distribution systems state estimation using long short-term memory networks as surrogates. *IEEE Access*, 8, 23359–23368.
- Carquex, C., Rosenberg, C., and Bhattacharya, K. (2018). State estimation in power distribution systems based on ensemble kalman filtering. *IEEE Transactions on Power Systems*, 33(6), 6600–6610.
- Civanlar, S., Grainger, J., Yin, H., and Lee, S. (1988). Distribution feeder reconfiguration for loss reduction. *IEEE Transactions on Power Delivery*, 3(3), 1217–1223.
- Duque, F.G., de Oliveira, L.W., de Oliveira, E.J., and Augusto, A.A. (2017). State estimator for electrical distribution systems based on an optimization model. *Electric Power Systems Research*, 152, 122–129.
- Haykin, S. (1999). *Neural networks: a comprehensive foundation*. Prentice hall.
- Liu, Y., Li, J., and Wu, L. (2019). State estimation of three-phase four-conductor distribution systems with real-time data from selective smart meters. *IEEE Transactions on Power Systems*, 34(4), 2632–2643.
- Majidi, M., Etezadi-Amoli, M., and Livani, H. (2017). Distribution system state estimation using compressive sensing. *International Journal of Electrical Power & Energy Systems*, 88, 175–186.
- Menke, J.H., Bornhorst, N., and Braun, M. (2019). Distribution system monitoring for smart power grids with distributed generation using artificial neural networks. *International Journal of Electrical Power & Energy Systems*, 113, 472–480.
- Mestav, K.R., Luengo-Rozas, J., and Tong, L. (2019). Bayesian state estimation for unobservable distribution systems via deep learning. *IEEE Transactions on Power Systems*, 34(6), 4910–4920.
- Pau, M., Ponci, F., Monti, A., Sulis, S., Muscas, C., and Pegoraro, P.A. (2017). An efficient and accurate solution for distribution system state estimation with multiarea architecture. *IEEE Transactions on Instrumentation and Measurement*, 66(5), 910–919.
- Peppanen, J., Grimaldo, J., Reno, M.J., Grijalva, S., and Harley, R.G. (2014). Increasing distribution system model accuracy with extensive deployment of smart meters. In *2014 IEEE PES General Meeting| Conference & Exposition*, 1–5. IEEE.
- Tecchio, P.P., Benedito, R.A., and Alberto, L.F. (2010). The behavior of wls state estimator near the maximum loadability point of power systems. In *IEEE PES General Meeting*, 1–6. IEEE.
- Wang, S., Zhao, J., Huang, Z., and Diao, R. (2017). Assessing gaussian assumption of pmu measurement error using field data. *IEEE Transactions on Power Delivery*, 33(6), 3233–3236.
- Zamzam, A.S., Fu, X., and Sidiropoulos, N.D. (2019). Data-driven learning-based optimization for distribution system state estimation. *IEEE Transactions on Power Systems*, 34(6), 4796–4805.
- Zhang, L., Wang, G., and Giannakis, G.B. (2019). Distribution system state estimation via data-driven and physics-aware deep neural networks. In *2019 IEEE Data Science Workshop (DSW)*, 258–262. IEEE.
- Zhang, Y. and Wang, J. (2020). Towards highly efficient state estimation with nonlinear measurements in distribution systems. *IEEE Transactions on Power Systems*, 35(3), 2471–2474.