

Pré-processamento de medidas para estimação de estado em condições adversas de monitoramento

José Paulo R. Fernandes* Etiane O. P. de Carvalho*
Júlio A. D. Massignan** João Bosco A. London Jr.*

* *Departamento de Engenharia Elétrica e Computação, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos - SP (e-mail: joseprf@usp.br, etiane@usp.br, jbalj@sc.usp.br).*

** *Siemens AG, Smart Infrastructure - Digital Grid, Curitiba-PR (e-mail: juliomassignan@gmail.com).*

Abstract: Static state estimators are still widely used for power systems monitoring and operation. This type of estimator makes use of data related only to the instant under estimation, ignoring past information. This paper proposes a measurement pre-processing method using past data and regression techniques for error detection and correction, providing improved data to static state estimators, specially under adverse situations. This approach requires no knowledge of the system dynamics, relying simply on data analysis, and is able to work with constrained datasets and high granularity. Unlike forecast aided state estimators, the proposed method operates over measurement data and does not output state variables, providing processed measurements instead. Results obtained from simulations with the 30 Bus IEEE test system indicates the proposed method can detect and correct errors in measurements, improve its accuracy and reduce effects of data losses by providing high quality pseudomeasurements.

Resumo: Os estimadores de estado estáticos ainda são amplamente utilizados para monitoramento e operação de sistemas elétricos. Este tipo de estimador faz uso somente de medidas obtidas no instante de operação, ignorando informações disponíveis de estados anteriores. Este artigo apresenta um método de pré-processamento de dados com base nas medições passadas, através de técnicas de regressão para detecção e correção de erros, fornecendo dados aprimorados para os estimadores de estado estáticos, especialmente em situações operacionais adversas. Essa abordagem não requer conhecimento da dinâmica do sistema, dependendo somente da disponibilidade de dados anteriores e pode lidar com amostras de dados pequenas e de alta granularidade. Ao contrário dos estimadores auxiliados por preditores, o método proposto opera sobre os valores medidos e fornece também uma saída de medições e não de estados. Os resultados obtidos com o sistema teste do IEEE de 30 barras mostram que o método proposto pode detectar e corrigir erros nas medições, melhorando a precisão do estimador de estado, com destaque para a capacidade de lidar com a perdas de dados fornecendo pseudomedidas de alta qualidade.

Keywords: Load Forecasting; State Estimation; Pre-Filtering.

Palavras-chaves: Previsão de Carga; Estimação de Estado; Pré-Filtragem.

1. INTRODUÇÃO

Embora existam tipos de estimadores de estado que lidam com problemas de dados e redundância, diferentes implementações variam em desempenho e complexidade. Estimadores de estado estáticos são simples de implementar e não requerem informações sobre a dinâmica dos sistemas, mas não são capazes de capturar efeitos relacionados ao tempo, diferentemente dos estimadores de estado dinâmicos, os quais requerem matrizes de transição (Zhao et al., 2019). Apesar das vantagens, estimadores dinâmicos podem sofrer com problemas de infraestrutura dos sistemas elétricos, uma vez que requerem medições mais frequentes

e precisas. Devido às dificuldades observadas, os modelos estáticos ainda são comuns em aplicações reais.

Os estimadores de estado são métodos estatísticos de análise de sistemas elétricos com base em medições ruidosas. Ainda que sejam capazes de lidar com ruídos, valores muito distorcidos e perdas de dados podem afetar consideravelmente sua precisão, especialmente quando ocorrem em medidas consideradas pontos de alavancamento, ou seja, que possuem grande impacto na convergência do estimador. Erros em medidas deste tipo são capazes de afetar todo o processo de estimação.

Uma das maneiras de lidar com erros em pontos de alavancamento sem modificar a estrutura do estimador é através do uso de técnicas de pré-processamento, as quais usualmente não tem seu desempenho afetado pela topologia do sistema analisado. Não somente erros em

* Os autores agradecem à Universidade de São Paulo (USP) e às agências de fomento CAPES (processo n° 88887.464768/2019-00 e n° 88887.510860/2020-00) e CNPq (processo n° 134387/2018-0).

pontos de alavancamento, mas também erros grosseiros em medidas diversas e dados perdidos podem ser tratados na etapa de pré-processamento, mesmo que ocorram em medidas críticas (Saran and Miranda, 2018).

O tratamento de possíveis problemas relacionados às medições por algoritmos de pré-processamento apresenta ainda potencial vantagem computacional, especialmente no que diz respeito ao tratamento de erros grosseiros (Coutto Filho and Stacchini de Souza, 2009). Ao contrário do estimador tradicional, que após detectar e remover erros precisa repetir o processo de estimação, o pré-processamento não precisa ser reexecutado a cada erro detectado. Entretanto, não há garantia que o pré-processamento será capaz de tratar todos os possíveis erros ou mesmo que não provocará o surgimento de outros, o que torna o ganho computacional variável.

Existe uma variedade de métodos que pode ser aplicada no pré-processamento, desde análises clássicas até técnicas de inteligência artificial. Alguns tipos de erros, como medidas com erros grotescos, nome que será utilizado para referir-se a medidas com erros óbvios devido a diferenças absurdas em relação aos valores usuais, podem ser facilmente identificados com o uso de regressões. Já a detecção de erros grosseiros e a recuperação de medidas requer o uso de técnicas especializadas.

Neste artigo propõe-se o uso de ferramentas de previsão e filtragem de medidas para realizar as operações necessárias sempre que novas medidas forem disponibilizadas pelo sistema de monitoramento, podendo ainda fornecer informações prévias do comportamento esperado da rede. A ferramenta apresentada neste artigo, inicialmente proposta em Fernandes et al. (2021), foi aprimorada para lidar com o problema de estimação, adicionando novas etapas de correção de dados com base no conjunto de medidas fornecido pelo sistema de monitoramento. Os prazos e precisão considerados foram baseados em dispositivos SCADA, ainda muito comuns nos sistemas de monitoramento, mas não há impedimentos para que a metodologia seja aplicada em conjunto com PMUs ou sistemas híbridos.

Simulações no sistema teste de 30 barras do IEEE com a metodologia proposta apresentaram resultados promissores para aplicação em tempo real, uma vez que a proposta foi capaz de identificar problemas no conjunto de medição com impacto computacional compatível com as necessidades operacionais de sistemas reais.

2. APRESENTAÇÃO DO PROBLEMA

O processo de estimação de estado em sistemas elétricos de potência consiste em encontrar as variáveis de interesse para monitoramento da rede com base em amostras ruidosas obtidas por dispositivos de medição. A partir dessa informação, pode ser feita a consideração de que o modelo Z de medição é resultado da adição de um vetor de erro de medidas w com os valores calculados através das funções de fluxo de potência $h(x)$, conforme a Equação 1.

$$Z = h(x) + w \quad (1)$$

O foco deste trabalho não é o problema de estimação em si, e sim a etapa de pré-processamento de dados que

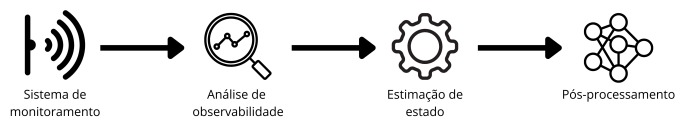


Figura 1. Visão simplificada do fluxo de dados e etapas do processo de estimação de estado.

são utilizados pelo estimador, especialmente analógicos, ou seja, valores obtidos pelas unidades de medições com informações de tensão, potência ativa, reativa e respectivos fluxos.

Erros na estimação de estado podem se manifestar em diferentes formas, tais como erros grosseiros (EG), inversões de sinais, perda de medidas ou outras distorções, sobre as quais a proposta a ser apresentada opera. Portanto, o problema discutido neste trabalho é um problema de dados, cuja aplicação é específica à estimação de estado.

De modo simplificado, o processo de estimação consiste em passar um conjunto de medidas ruidosas por um método matemático a fim obter uma figura confiável da situação operacional do sistema elétrico sob monitoramento. Para que este procedimento seja bem sucedido, é necessário garantir a observabilidade do sistema e a qualidade do conjunto de medidas, ou seja, as medidas disponíveis devem ser suficientes para solução das equações envolvidas e possuir precisão adequada para que o resultado seja confiável. Quando necessário, também é realizada uma etapa de pós-processamento para remoção de erros. O fluxo de dados e etapas envolvidas no processo tradicional de estimação de estado é mostrado no Figura 1.

A proposta deste trabalho atua imediatamente após a primeira etapa, na qual são obtidas as medições. Para garantir que os dados atendam às condições necessárias para o bom funcionamento dos estimadores de estado, são utilizados algoritmos de previsão e filtragem de dados, com adaptações específicas à aplicação desejada.

Algoritmos de previsão já são amplamente utilizados no monitoramento, operação e planejamento de redes elétricas há décadas, com diversas técnicas já estabelecidas aplicadas com sucesso (Fallah et al., 2018). Dependendo do horizonte de aplicação do algoritmo - curto, médio ou longo prazo - são feitas abordagens específicas, que afetam desde a implementação do método até o tipo de amostra necessária para seu funcionamento adequado. É comum a utilização de grandes amostras do histórico de operação, que incluam ainda dados climáticos, no treinamento dos algoritmos de previsão, questão a qual foi tratada em Fernandes et al. (2021). Neste trabalho, o problema de previsão é apresentado como uma função dependente de hábitos sociais H e do tempo T , de modo que os preditores são funções que seguem o modelo apresentado na Equação 2, sendo z_P a medida a ser prevista.

$$z_P = f(H, T) \quad (2)$$

Sozinhos, os algoritmos de previsão estão sujeitos a falhas especialmente em momentos de transição ou mudanças de perfil de carga da rede. Este problema já foi observado nos trabalhos de Khotanzad et al. (1995) e Mahdavi et al. (2006), que buscam criar sistemas especializados

em lidar com este problema. Apesar de ter um bom desempenho nesses casos devido ao uso de regressores múltiplos e fatores de ajuste, a proposta de Fernandes et al. (2021) fica restrita aos modelos de regressão obtidos antes das mudanças. Por outro lado, é possível criar relações entre os próprios dados de medições para captar possíveis mudanças em pontos específicos da rede com base no comportamento das variáveis ao seu redor. Ou seja, pode-se dizer que o comportamento de uma variável da rede é uma função das variáveis em sua vizinhança, conforme mostra a Equação 3, onde z_F é a medida que deseja-se calcular e z_V o conjunto de medidas advindas de sua vizinhança.

$$z_F = f(z_V) \quad (3)$$

Esta abordagem permite fazer uma análise complementar à preditiva. Ao contrário da previsão, esta pode ser utilizada após a chegada do conjunto de medições para verificar a existência de incoerência entre os valores. O problema de aplicar este método, considerando a disponibilidade de dados, está em delimitar quais os dados devem ser utilizados para estimar cada valor, mantendo custo computacional e precisão viáveis à aplicação.

Do ponto de vista do problema de estimação, a análise de conjunto de medidas recebidas deve ser robusta para não falhar em ocasiões de erros generalizados na mesma vizinhança. Observa-se que, nesta situação, a Equação 3 está sujeita a erros já na entrada, o que levaria à propagação destes erros na saída, afetando fortemente o desempenho do estimador de estado a ser executado em sequência.

3. METODOLOGIA PROPOSTA

A metodologia proposta pode ser separada em três etapas a serem realizadas:

- (1) Previsão de medidas: são previstos valores esperados com base no histórico de medição, hábitos sociais e informações de data e hora;
- (2) Validação da previsão: etapa de comparação entre os valores previstos, os valores recebidos do sistema de medição e os valores esperados conforme a vizinhança de cada medida;
- (3) Correção de medidas: atualização dos valores de medidas consideradas imprecisas na etapa anterior através de análise iterativa da vizinhança.

3.1 Previsão de medidas

A primeira etapa é a previsão de medidas, realizada enquanto não há novas informações do sistema de medição. Essas medidas são obtidas por um preditor com entradas baseadas nas informações temporais e de hábitos conhecidos da população local. A utilização de informações temporais já é comum na literatura, enquanto a adição de hábitos sociais tem sido discutida em algumas propostas (Javed et al. (2012) e Mahdavi et al. (2006)). Em Fernandes et al. (2021) foi mostrado que dados sociais de hábitos comuns, como horários normais de almoço, café da manhã, trabalho e finais de semana podem ser utilizados juntamente com informações horárias na previsão. A proposta do trabalho

de utilizar múltiplas curvas do algoritmo *Stochastic Gradient Boosting* (SGB) treinada a partir do uso de diferentes quantis de dados, cuja formulação não será discutida aqui, mas pode ser melhor entendida a partir dos trabalhos de Friedman (2002) e Fernandes et al. (2021), foi mantida, sendo escolhida a curva com menor erro percentual médio nas previsões recentes, mas com mudanças no método de ajuste das curvas.

Enquanto a proposta original considera a multiplicação dos regressores por um fator α obtido em função do menor erro percentual médio ϵ , obtido ao comparar o desempenho da curva selecionada, que gera valores z_P , com os valores de fato utilizados pelo estimador z_R em um passado recente, conforme mostrado nas equações 4 e 5, em qualquer ocasião, neste trabalho foram acrescentadas restrições para melhor tratamento de períodos de transição. Essas restrições visam coibir ajustes muito grandes, casos em que um valor alto de α pode levar o sistema de previsão a cometer erros graves. Neste trabalho, α para qualquer $\epsilon > 0, 1$ foi considerado inapropriado, fazendo-se uso direto da Equação 2 nestes casos (ou seja, considera-se $\alpha = 1$). Tal medida é essencial para evitar problemas quando há mudanças nos horários de transição. A escolha do regressor tem o critério mantido, sendo o com menor ϵ selecionado, mesmo se este for maior que 0, 1.

$$\alpha = \frac{1}{1 + \epsilon} \quad (4)$$

sendo $\epsilon_m = \sum \frac{z_P - z_R}{z_R}$

$$\begin{aligned} z_{P'} &= \alpha f(H, T) \text{ se } \epsilon < 0, 1 \\ z_{P'} &= f(H, T) \text{ caso contrário} \end{aligned} \quad (5)$$

Observe que o uso de α visa tornar o erro médio do preditor em relação às medidas de um passado recente igual a zero ao deslocar a curva do regressor selecionado proporcional à média dos erros observados, de modo que este simule o comportamento teórico de um medidor que siga um modelo da Equação 1, no qual a distribuição do erro tenha média zero.

Os valores obtidos nesta etapa são utilizados para preencher valores de medições perdidas ou, quando identificados que são aproximadamente iguais a um valor multiplicado por -1 , substituir valores com provável erro de sinalização, desde que estes atendam às restrições de distância mínima de n desvios padrão do valor original, sendo n dado pelo menor valor entre 50 desvios padrão (σ) e 10 vezes a tolerância a erros δ_ϵ , a qual é dada por um valor n vezes o desvio padrão de cada medidor σ , desejada pelo operador. Esta tolerância também é utilizada para definir na etapa seguinte quais medições continuarão a ser processadas. Destaca-se que os valores previstos podem ser utilizados para a obtenção de um estado previsto, o qual não é o objetivo deste trabalho, mas pode ser útil em outros casos, atuando de modo semelhante ao estimador FASE Zhao et al. (2019).

3.2 Validação da previsão

Uma vez que os valores previstos já são conhecidos, aguarda-se a chegada de um novo conjunto de medições, o

qual deve ser validado com base nos valores obtidos pelo preditor e pelo filtro, que tem como entradas as medidas da vizinhança. Este filtro é dado pela Equação 3. Para que seu funcionamento seja computacionalmente viável são feitas restrições ao tamanho da vizinhança utilizada através de um valor máximo de distância elétrica entre as grandezas dado por d_{max} , sendo considerado um incremento de 1 para cada "afastamento" do ponto de medição. Por exemplo, para uma medição em uma Barra X, todas as medições na mesma barra ou conectadas no início de linhas que saem dela tem distância $d = 0$, já medições no final de linhas diretamente conectadas a ela possuem $d = 1$, assim como das barras ao final da linha, e assim por diante.

Nem todas as medidas precisam passar pelo filtro, somente aquelas consideradas "suspeitas" de serem imprecisas. A suspeição, dada pela variável s , e cujo valor será 1 para medidas suspeitas e 0 para medidas não suspeitas, é definida conforme parâmetros de entradas que indicam a diferença máxima tolerada entre o valor previsto para uma medida e o valor recebido do sistema de medição. Assim, sempre que a condição apresentada na Equação 6 for verdadeira, uma medida receberá $s = 1$.

$$\begin{aligned} & \text{se } |z_P - z_R| > \delta_\epsilon \\ & \text{então } s(z) = 1 \\ & \text{caso contrário } s(z) = 0 \end{aligned} \quad (6)$$

Todas as variáveis cujo valor de s forem 1 são testadas por um regressor obtido utilizando o modelo apresentado na Equação 3, o qual, neste trabalho, foi obtido utilizando o mesmo método dos preditores, o algoritmo SGB, mas baseando-se em uma curva única ao invés de múltiplos regressores. Compara-se então os valores fornecidos pelo sistema de medição com os valores calculados pelo regressor do filtro para cada medida suspeita, removendo da lista de suspeitas todas aquelas medidas cuja diferença entre o valor estimado pelo filtro e o valor aferido pelo medidor esteja dentro do tolerado pelo operador, conforme mostra a Equação 7.

$$\begin{aligned} & \text{se } |z_F - z_R| < \delta_\epsilon \\ & \text{então } s(z) = 0 \\ & \text{caso contrário } s(z) = 1 \end{aligned} \quad (7)$$

O conjunto de medidas cujo valor de s continuar indicando suspeição desta medida seque para a etapa seguinte, na qual estas podem ser atualizadas ou não.

3.3 Correção de medidas

A terceira e última etapa do método proposto é responsável por tentar corrigir erros menores, não detectados pelas etapas anteriores, e busca aumentar a precisão das pseudo-medidas geradas na primeira etapa para substituir valores perdidos ou suspeitos de inversão de sinal. Esta etapa é realizada através do uso do mesmo regressor responsável pela filtragem, mas segue um processo iterativo desenvolvido para lidar com a atualização de valores com erros na mesma vizinhança. O problema do processamento de erros múltiplos e fortemente correlacionados já foi observado por Falcao et al. (1982), sendo esta etapa desenvolvida

especialmente para aumentar a robustez do algoritmo em situações similares.

Antes de se iniciar a correção de medidas, é calculado um índice de suspeição S para cada medida remanescente no grupo de suspeitas da etapa anterior. O cálculo é realizado conforme a Equação 8, onde $d(z_V)$ indica a distância entre a medida a ser processada e suas vizinhas. Este índice pode variar de 0, quando nenhuma medida na vizinhança é suspeita, até 1, quando todas as medidas da vizinhança continuam suspeitas após as duas primeiras etapas.

$$S = \frac{\sum [s(z_V)[d_{max} + 1 - d(z_V)]]}{\sum [d_{max} + 1 - d(z_V)]} \quad (8)$$

O índice de suspeição é utilizado para definir a ordem em que as medidas serão processadas, iniciando-se da medida com menor S e seguindo a ordem crescente, somente atualizando os valores cuja variação em relação ao valor de entrada, seja este originário do sistema de medição ou de processamento anterior, seja maior que a tolerância δ_ϵ . Ressalta-se que sempre que um valor é atualizado, esta atualização é carregada para a correção das demais medidas que a tenham em sua vizinhança, em um processo similar ao método de Gauss-Seidel para resolução de sistemas de equações.

Sempre que for identificado que a variação esperada para uma variável em uma iteração é menor que a tolerância δ_ϵ , esta é removida da lista de suspeitos. Quando uma variável sai da lista de suspeitas, é mantido o valor pré-iteração. O processo é repetido iterativamente até que todas as medidas saiam da lista de suspeitas ou o número limite de iterações seja alcançado.

A seqüência e as restrições aplicadas tem dois objetivos:

- Minimizar a propagação de erro no filtro ao incluir inicialmente o mínimo de medidas suspeitas possível;
- Minimizar a quantidade de alterações promovidas pelo filtro.

3.4 Visão geral da proposta

Conhecidas as três etapas de processamento, é possível ter uma visão geral de como elas interagem entre si e onde se encaixam no processo de estimação de estado. Para este fim, é apresentada a Figura 2.

4. TESTES, RESULTADOS E ANÁLISE

4.1 Metodologia de testes

O método proposto foi avaliado quanto ao desempenho considerando a previsão direta de medidas e a capacidade de correção. Os valores obtidos foram comparados com valores de referência e com uma amostra ruidosa. Para obtenção dos dados de testes, foram geradas curvas do comportamento do sistema testes de 30 barras do IEEE, o qual é mostrado na Figura 3, a partir de dados reais de um sistema elétrico. As informações das curvas foram validadas através da análise utilizando um algoritmo de fluxo de potencial. Foram aplicados ruídos gaussianos sobre os valores de referência para obter a amostra ruidosa,

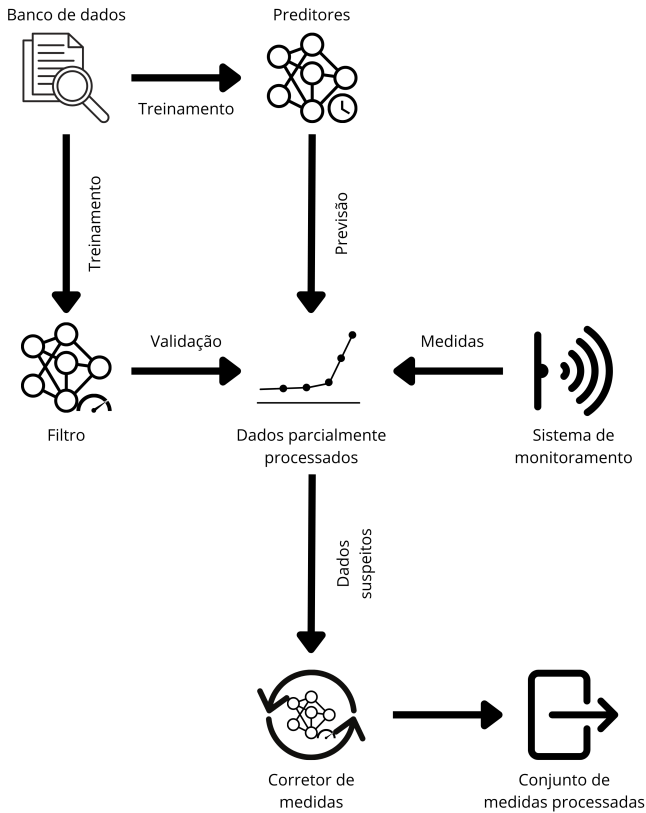


Figura 2. Visão geral da metodologia proposta mostrando a interação entre as diferentes etapas de processamento e o sistema de monitoramento.

Tabela 1. Tipos de medidas e respectivos desvios padrão adotados nos testes.

Tipo de medida	Desvio Padrão
Fluxo	0.002
Injeção	0.002
Tensão	0.02

a qual também serviu para a execução das etapas de treinamento dos regressores.

Ao todo, as curvas simulam um total de 22 dias sequenciais de operação do sistema, sendo 17 utilizados para treinamento do método proposto e parte dos 5 demais para validação. As amostras possuem uma granularidade de 5 minutos, o que significa que 4896 medições de cada variável estavam disponíveis para treinamento e 1440 para validação, das quais 1151 foram utilizadas.

O sistema de medições considerado para testes possui o seguinte sistema de monitoramento:

- Medidas de fluxo de ativo e reativo: 2-1, 2-6, 4-12, 6-7, 9-11, 20-10, 10-22, 22-10, 12-13, 15-23, 18-19, 19-20, 22-24, 24-25, 25-27, 29-27, 30-29;
- Medidas de injeção de ativo e reativo: 1, 2, 5, 6, 8, 10, 12, 15, 16, 18, 19, 24, 25, 26, 27, 30;
- Tensões nas barras: 1, 2, 5, 8, 9, 10, 11, 12, 13, 24, 27.

Sendo os desvios padrão para cada tipo de medida mostrados na Tabela 1.

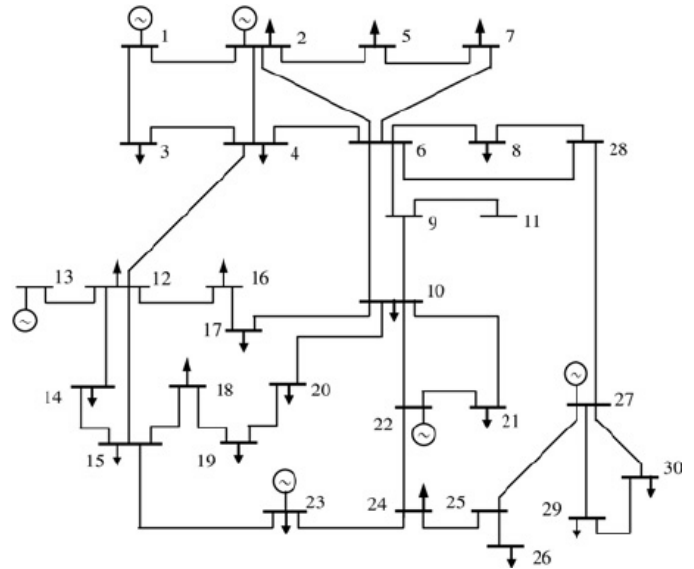


Figura 3. Diagrama do sistema teste do IEEE de 30 barras.

Foram adotados como parâmetros do método proposto $\delta_\epsilon = 3\sigma$, valor usualmente adotado para detecção de erros grosseiros, $d_{max} = 2$ e a utilização de 5 instantes anteriores na etapa de seleção e ajustes dos preditores, os quais foram obtidos através da análise dos quantis de 5% a 95%, em intervalos de 5%, da amostra de testes. Os parâmetros adotados para os regressores SGB seguiram as sugestões apresentadas por Friedman (2002).

Para que pudessem ser simulados problemas no sistema de medição, foram aplicados aleatoriamente de 0 a 5 erros em cada amostra de validação, as quais são retiradas do conjunto de amostras com ruído. Os erros aleatórios possíveis de ocorrer, todos com a mesma probabilidade, são:

- Variação aleatória de até 5σ no valor da medida;
- Inversão do sinal da medida;
- Perda completa do valor da medida.

Feitas todas estas considerações foram avaliadas três situações de operação do método proposto: (I) atuação sobre a amostra ruidosa sem ocorrência dos erros aleatórios adicionais; (II) atuação sobre a amostra ruidosa com ocorrência dos erros aleatórios adicionais; e (III) utilização do método proposto para previsão de todas as medidas do instante seguinte (agindo exclusivamente como preditor). Em todos os casos, os valores processados foram comparados com a amostra de referência, na qual os dados não apresentaram qualquer ruído.

Avaliou-se assim a atuação do método proposto no processamento de um total de 88627 medidas em cada situação, considerando o intervalo de tempo e o sistema de medição testados. Os dados de saída do método proposto foram posteriormente fornecidos a um estimador de estados MQP para avaliação do impacto de sua aplicação.

4.2 Resultados obtidos

Atuação sobre amostra ruidosa sem ocorrência dos erros aleatórios adicionais: A aplicação do método proposto sobre uma medida ruidosa, mas sem ocorrência de qualquer anormalidade, ou seja, sem perdas, inversões de sinais

Tabela 2. Comparação entre amostra ruidosa e amostra processada na Situação I.

	Amostra ruidosa	Amostra processada
Erros de tensão	32	9
Erros de injeção	89	301
Erros de fluxo	99	350
Dados perdidos	0	0
Erros grosseiros	220	660
Erros totais	220	660
Mediana do ruído normalizado absoluto	0,5665	0,4930
Média do ruído normalizado absoluto	0,7027	0,6642

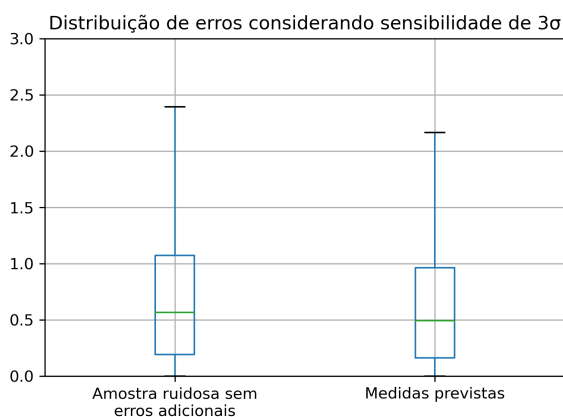


Figura 4. Comparação da distribuição de ruídos da amostra ruidosa sem erros adicionais com os dados de saída para a situação II.

ou erros de grande magnitude, tem a finalidade de avaliar a capacidade deste de não interferir em dados considerados bons e de reduzir o ruído quando possível. A amostra ruidosa originalmente contava com um total de 220 erros grosseiros nas 88627 medidas, subindo para 660 ao final do processo. Porém, houve redução na mediana do ruído normalizado absoluto de 0,5665 para 0,4930 e também na média de 0,7027 para 0,6418.

Ao todo, o método proposto promoveu 23817 mudanças, das quais 14748 resultaram na redução de ruído se comparado à amostra de referência, em relação à amostra ruidosa. A grande maioria dos erros observados após o processamento (651 dos 660, ou seja, quase 99%) foram em medidas de injeção ou fluxo de potência, cujas magnitudes muitas vezes eram similares às dos desvios padrão dos próprios medidores nas condições de testes. Mais detalhes dos resultados são mostrados na Tabela 2 e na Figura 4.

Atuação sobre amostra ruidosa com ocorrência dos erros aleatórios adicionais: A aplicação do método proposto sobre uma medida ruidosa, com ocorrência de anormalidades na comunicação, as quais podem ocasionar em perdas de dados, inversões de sinais ou erros de grande magnitude, visa avaliar a capacidade deste de recuperar e/ou corrigir dados, reduzindo o ruído quando possível. Nesta situação, a aplicação de erros aleatórios gerou uma amostra com 4263 erros, sendo 1982 valores perdidos, 1623 inversões de sinais e 2281 erros grosseiros causados pela adição de ruído extra ou inversões de sinal.

Tabela 3. Comparação entre amostra de entrada e amostra processada na Situação II.

	Amostra de entrada	Amostra processada
Erros de tensão	653	29
Erros de injeção	1710	547
Erros de fluxo	1912	592
Dados perdidos	2036	0
Erros grosseiros	2662	1168
Erros totais	4275	1168
Mediana do ruído normalizado absoluto	0,6150	0,5160

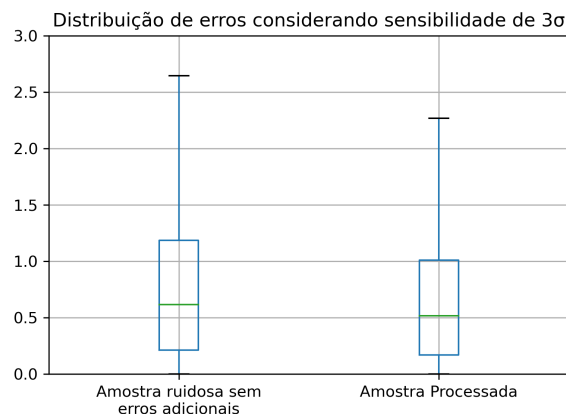


Figura 5. Comparação da distribuição de ruídos da amostra ruidosa sem erros adicionais com os dados de saída para a situação II.

Neste caso, foi possível observar considerável redução na quantidade de erros após o pré-processamento, conforme mostram os dados da Tabela 3. Não somente houve recuperação dos dados, como os dados de saída apresentaram significativa melhora na distribuição de ruídos em relação mesmo à amostra ruidosa sem erros adicionais, conforme mostra a Figura 5.

Ao todo, foram promovidas 27.131 alterações pelo método proposto, das quais 17.549 resultaram em medidas com ruído reduzido em relação à amostra ruidosa sem erros não processados quando comparadas às medidas de referência. Não houve alteração significativa da mediana do ruído normalizado pelo desvio padrão, sendo nos dois casos a mediana aproximadamente 0 (considerando valores absolutos, houve redução de 0,6150 para 0,5160). Não foi feita comparação da média devido à perda de dados simulada. Apesar das melhorias, o maior ruído observado no conjunto de dados processados foi de 144,6 desvios padrão, em uma medida recuperada.

Previsão de todas as medidas do instante seguinte: Esta situação considera que o método proposto é responsável por fornecer previsões de todo o conjunto de medidas, sempre conhecendo o instante anterior, mas sem execução das etapas de validação e correção de medidas, que ocorrem somente após a chegada de dados do sistema de monitoramento. Nestas circunstâncias, todas as 88627 medidas são fornecidas diretamente pelo preditor, o qual gerou 40049 medidas com redução de ruído ao comparar com a amostra ruidosa, 9208 medidas com ruído idêntico, 39367 medidas com aumento de ruído e 3 medidas com ruído acima de

Tabela 4. Comparação entre amostra ruidosa e medidas previstas na Situação III.

	Amostra ruidosa	Medidas previstas
Erros de tensão	32	0
Erros de injeção	89	3813
Erros de fluxo	99	4109
Dados perdidos	0	3
Erros grosseiros	220	7919
Erros totais	220	7922
Mediana do ruído normalizado absoluto	0,5665	0,5013

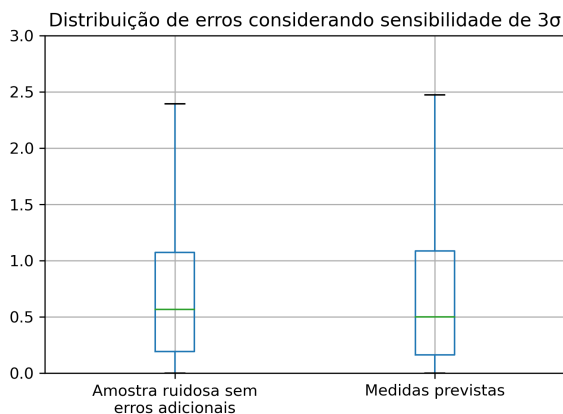


Figura 6. Comparação da distribuição de ruídos da amostra ruidosa sem erros adicionais com os dados de saída para a situação III.

500 desvios padrão, as quais foram consideradas medidas perdidas.

Os resultados detalhados desta situação são mostrados na Tabela 4 e na Figura 6. Observa-se que mesmo nesta situação o preditor teve mediana dos ruídos inferior à mediana da amostra ruidosa, apesar de apresentar uma quantidade alta de EGs.

4.3 Aplicação em conjunto com estimador de estado

Os dados obtidos nas três situações foram processados por um estimador de estado MQP, sem a aplicação da etapa de processamento de erros grosseiros, e comparados aos valores obtidos utilizando a amostra de referência (sem qualquer ruído). Os resultados obtidos são mostrados na Tabela 5. Nas Figuras 7 e Figuras 8 são mostrados o comportamento da estimativa de potência na barra de referência, cujo valor é o mais suscetível a variações, e as distribuições de erros de saída do estimador observados em relação à referência. É importante lembrar que as estimativas da amostra processada são fruto de medidas recuperadas de uma amostra ruidosa que sofreu outras falhas, não sendo diretamente comparável com a utilização da amostra ruidosa sem estas falhas. Não foi utilizada a quantidade de erros grosseiros detectados pelo estimador, pois a presença ou não de EG está relacionada ao conjunto de medidas e não somente a uma medida, não sendo a forma mais adequada de verificar a precisão da saída do estimador para a proposta deste trabalho.

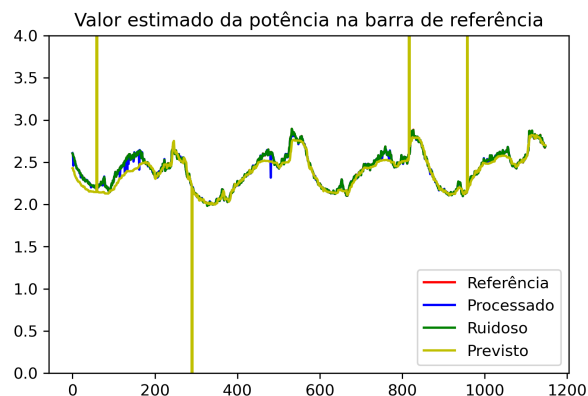


Figura 7. Curvas estimadas considerando cada situação avaliada.

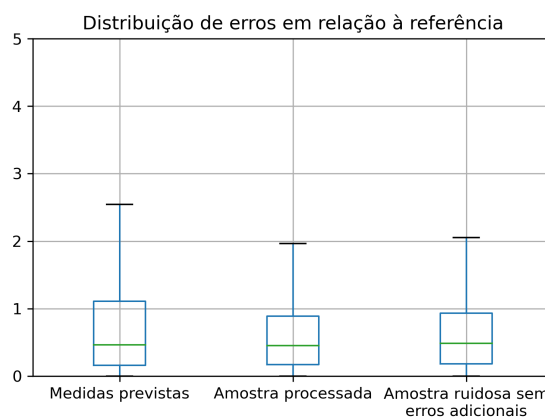


Figura 8. Comparação da distribuição de ruídos em relação à referência após a execução do estimador, sem processamento de erros grosseiros.

Tabela 5. Comparação de desempenhos do estimador utilizando a amostra ruidosa sem erros adicionais, a amostra com erros adicionais processada e o conjunto de medidas previstas.

	Amostra ruidosa	Amostra processada	Medidas previstas
Erros maiores que 3σ	982	1512	4262
Mediana do resíduo normalizado absoluto	0,4867	0,4527	0,4635
Mediana do resíduo normalizado	0	0	0,0003

4.4 Análise dos resultados

Os resultados obtidos mostram o potencial do método proposto para lidar com situações adversas nos sistemas de monitoramento que alimentam estimadores de estado. Destaca-se que os resultados obtidos após o processamento dos dados na situação II, que é a principal condição a qual este tipo de processamento é destinada, são compatíveis com os resultados da utilização de uma amostra ruidosa comum, apesar da quantidade de erros grosseiros superior à da amostra antes da ocorrência dos erros e de haver algumas medidas recuperadas com erros altos em instantes de transição. A eficiência da proposta para lidar com situ-

ações problemáticas pode ser percebida pela semelhança das curvas estimadas mostradas na Figura 7.

A redução de ruído é útil para fornecer estimativas com maior precisão, mas observa-se pela situação I que a aplicação exclusivamente com esta finalidade pode impactar a quantidade de erros grosseiros, especialmente em medidas cujo valor seja de magnitude similar ao desvio do próprio medidor. Neste caso, vale acrescentar que os preditores foram considerados de precisão idêntica aos próprios medidores, tendo sido treinados com amostra ruidosa. Um tratamento diferente, aumentando o desvio esperado nas pseudomedidas geradas pelo processamento, similar ao que é feito no uso de pseudomedidas de baixa precisão, poderia eliminar este efeito, mas é necessário realizar um estudo de qual seria este impacto e qual a melhor forma de estabelecer o desvio adequado para estas medidas.

Por fim, a situação III, na qual o pré-processamento atua de forma similar aos estimadores FASE, mostra que esta pode ser uma aplicação viável da proposta. Novamente, cabe a discussão de como tratar o desvio de medidas obtidas exclusivamente pelo preditor nesta aplicação. Acrescenta-se ainda que as etapas de validação e correção de medidas podem ser adaptadas para lidar com somente o conjunto de pseudomedidas geradas pelo preditor, o que poderia evitar os erros observados na curva prevista da Figura 7, mas trata-se de outra condição que requer estudos específicos.

5. CONCLUSÃO

A proposta deste trabalho mostra-se promissora para lidar com sistemas de monitoramento problemáticos em redes elétricas. Apesar de terem sido observadas ocasiões em que a proposta pode afetar negativamente a quantidade de erros, este impacto foi pequeno comparado ao benefício de se evitar grandes impactos quando há a ocorrência de falhas no sistema de monitoramento.

Além de cumprir seu objetivo de fornecer pseudomedidas de alta qualidade e corrigir valores defeituosos, a proposta foi desenvolvida considerando forte restrição de dados e alta granularidade, compatíveis com a situação operacional de diversos sistemas elétricos. Estes objetivos já foram tratados em Fernandes et al. (2021) e este trabalho visa dar continuidade e diversificar as aplicações da proposição original. Os erros de previsão observados continuam compatíveis com aqueles esperados na literatura para previsões de curto prazo e alta granularidade (Nagumo et al., 2017).

Em trabalhos futuros, espera-se ampliar o escopo dos testes realizados neste trabalho, avaliando o desempenho com variações nos parâmetros e sistemas de maior porte. Apesar de este trabalho ser focado na aplicação do conjunto de preditor e filtro na etapa de pré-processamento do problema de estimação de estado, estas técnicas, em conjunto, poderiam ser aplicadas em outras atividades de operação e monitoramento de sistemas elétricos, como restabelecimento de energia e coordenação de manobras de segurança. Nestas situações, no entanto, o horizonte de análise, a granularidade e a precisão requeridas podem diferir.

AGRADECIMENTOS

Os autores agradecem à Universidade de São Paulo (USP) e às agências de fomento CAPES (processo n° 88887.464768/2019-00 e n° 88887.510860/2020-00) e CNPq (processo n° 134387/2018-0).

REFERÊNCIAS

- Coutto Filho, M.B. and Stacchini de Souza, J.C. (2009). Forecasting-aided state estimation—part i: Panorama. *IEEE Transactions on Power Systems*, 24(4), 1667–1677.
- Falcao, D., Brameller, A., and Cooke, P. (1982). Power system tracking state estimation and bad data processing. *IEEE Power Engineering Review*, PER-2(2), 20–20.
- Fallah, S.N., Deo, R.C., Shojafar, M., Conti, M., and Shamshirband, S. (2018). Computational intelligence approaches for energy load forecasting in smart energy management grids: State of the art, future challenges, and research directions. *Energies*, 11(3).
- Fernandes, J.P.R., Massignan, J.A.D., London, J.B.A., and Fanucchi, R.Z. (2021). Very short-term current and load forecasting for distribution systems in data constrained situations. In *2021 IEEE Madrid PowerTech*, 1–6.
- Friedman, J.H. (2002). Stochastic gradient boosting. *Computational Statistics Data Analysis*, 38(4), 367–378. Nonlinear Methods and Data Mining.
- Javed, F., Arshad, N., Wallin, F., Vassileva, I., and Dahlquist, E. (2012). Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting. *Applied Energy*, 96, 150–160. Smart Grids.
- Khotanzad, A., Hwang, R.C., Abaye, A., and Maratukulam, D. (1995). An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities. *IEEE Transactions on Power Systems*, 10(3), 1716–1722.
- Mahdavi, N., Menhaj, M.B., and Barghinia, S. (2006). Short-term load forecasting for special days using bayesian neural networks. In *2006 IEEE PES Power Systems Conference and Exposition*, 1518–1522.
- Nagumo, T., Ito, H., and Sano, T. (2017). Load current forecasting using statistical analysis. *CIREN - Open Access Proceedings Journal*, 2017(1), 1274–1277.
- Saran, M.A. and Miranda, V. (2018). State estimation pre-filtering with overlapping tiling of autoencoders. *Electric Power Systems Research*, 157, 261–271.
- Zhao, J., Gómez-Expósito, A., Netto, M., Mili, L., Abur, A., Terzija, V., Kamwa, I., Pal, B., Singh, A.K., Qi, J., Huang, Z., and Meliopoulos, A.P.S. (2019). Power system dynamic state estimation: Motivations, definitions, methodologies, and future work. *IEEE Transactions on Power Systems*, 34(4), 3188–3198.