

Estimação, Identificação e Localização de Perdas Técnicas e Comerciais em Sistemas de Distribuição: Método Híbrido Baseado em Estimador de Estados e Reconhecimento de Padrões

Aquiles Rossoni, Rodrigo D. Trevisan, Arturo S. Bretas, Daniel da S. Gazzana, Arlan L. Bettiol, Antônio Carniato, Luis F. N. Passos e Rodolfo P. Martin

Resumo – Este trabalho propõe um sistema computacional baseado em uma metodologia híbrida para estimar, identificar e localizar Perdas Técnicas (PT) e Perdas Comerciais (PC) em sistemas de distribuição. A metodologia utiliza o Estimador de Estados por Mínimos Quadrados Ponderados (EE-MQP) e técnicas de reconhecimento de padrões (Floresta de Caminhos Ótimos (FCO) e Detecção de Anomalias (DA)). A metodologia de reconhecimento de padrões é selecionada de acordo com a disponibilidade de dados da concessionária. Inicialmente, a metodologia foi validada em um sistema teste com clientes com PC conhecidas, sendo este sistema baseado no *IEEE 123 Bus Test Feeder* e no comportamento de clientes residenciais brasileiros. A seguir, a metodologia foi aplicada em um alimentador da Companhia Hidroelétrica São Patrício (CHESP). As principais funcionalidades do sistema integrado são a aquisição automática de informações de bancos de dados e a execução da metodologia proposta por meio de ferramenta computacional e interface homem-máquina otimizada.

Palavras-chave – Estimador de Estados, Perdas Comerciais, Perdas Técnicas, Reconhecimento de Padrões.

I. INTRODUÇÃO

As perdas de energia elétrica nas empresas de distribuição são um problema com impacto elevado. Em 2000, as perdas

Este trabalho foi desenvolvido no âmbito do Programa de Pesquisa e Desenvolvimento Tecnológico do Setor de Energia Elétrica regulado pela ANEEL e consta dos Anais do VIII Congresso de Inovação Tecnológica em Energia Elétrica (VIII CITENEL), realizado na Costa do Sauípe/BA, no período de 17 a 19 de agosto de 2015.

Os autores gostariam de agradecer ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pelo financiamento das bolsas dos mestrandos Aquiles Rossoni e Rodrigo D. Trevisan, à Universidade Federal do Rio Grande do Sul (UFRGS) pelas facilidades oferecidas e às empresas CHESP, CERRP, CERPRO, CERNHE, CERIPA, CERAL-DIS, CETRIL, CERIM, CERMC, CERIS, CEDRI, CERES, CEDRAP, ELFSM, EFLJC, COOPERALIANÇA e CERCOS pelo suporte financeiro para o desenvolvimento deste trabalho mediante a realização de um projeto de P&D ANEEL (projeto 0103-0002/2011).

A. Rossoni, R. D. Trevisan, A. S. Bretas e D. S. Gazzana estão com a Universidade Federal do Rio Grande do Sul (e-mails: arossoni@ece.ufrgs.br; rodtrevisan@ece.ufrgs.br; abretas@ece.ufrgs.br; ds Gazzana@gmail.com).

A. L. Bettiol, A. Carniato, L. F. N. Passos trabalham na Neo Domino Pesquisa em Sistemas Elétricos Ltda. (e-mails: arlan@neodomino.com.br; carniato@neodomino.com.br; luis@neodomino.com.br).

R. P. Martin trabalha na Companhia Hidrelétrica São Patrício (e-mail: roldolfo@chesp.com.br).

em distribuição e transmissão foram equivalentes a 16 % de toda a energia mundialmente produzida, representando um aumento com o passar dos anos, sendo que em 1980, o nível de perdas era de 11% [1]. No Brasil, as perdas de energia elétricas na distribuição representam 15% da energia total consumida [2].

As perdas de energia elétrica são classificadas em técnicas e comerciais. A PT é a perda inerente ao processo de transporte, transformação e medição de energia elétrica e as PC é a energia entregue, porém não faturada, principalmente devido a furtos, fraudes ou problemas de faturamento. No Brasil, as PC geram uma perda anual superior a sete bilhões de reais [2]. Com o objetivo de melhorar seu desempenho técnico e financeiro, as empresas do setor elétrico buscam adotar ações para mitigar as perdas de energia elétrica.

A redução de PC é dada, principalmente, através duas abordagens: ações preventivas e/ou identificação de consumo irregular [3]. Ações como o emprego de condutores anti-fraude, linhas subterrâneas no sistema secundário e medidores robustos à prova de fraudes são algumas das medidas que tornam mais difícil o consumo irregular [3]. Quando há irregularidades em consumidores, estas devem ser constatadas por auditorias realizadas por equipes de inspeção para que as mesmas sejam corrigidas. Considerando que a realização de inspeções em todos os consumidores é financeiramente inviável, métodos são aplicados para identificar os consumidores suspeitos e direcionar estas inspeções, aprimorando a relação entre recuperação de receitas perdidas por PC e o investimento na sua mitigação. Na identificação de consumidores irregulares, alguns dos métodos utilizados são a adoção de medidores com alarmes antifraude, o uso de medidores totalizadores na baixa tensão de transformadores de distribuição e a utilização de metodologias baseadas em reconhecimento de padrões de clientes irregulares. A última família de métodos tem a vantagem de não exigir amplos investimentos na renovação ou compra de novos equipamentos, além de ser flexível e poder ser constantemente aprimorada a um baixo custo [4].

A redução das PT é normalmente realizada através da manutenção ótima dos elementos de distribuição (linhas, transformadores, entre outros) ou adição de elementos ao sistema para compensar a potência reativa (bancos de capacitores), a queda de tensão (regulador de tensão) ou fluxo de

potência (geração distribuída) [3]. Os métodos mais usuais são baseadas em técnicas de otimização. Estes métodos consideram a possível redução das PT levando-se em conta o nível de inserção, a localização e os custos dos elementos mencionados [5]-[7].

Conforme descrito, dado que as ações de redução de PT e PC são diferentes, elas costumam ser tratadas de forma separada. Considerando que essas ações envolvem investimento de capital financeiro e humano, torna-se claro a importância de corretamente estimar as PT e PC para decidir qual é mais relevante à concessionária. Adicionalmente, na literatura, os métodos de reconhecimento de padrões para identificação PC estão desvinculados de métodos de estimação de perdas. Portanto, este trabalho propõe uma nova metodologia para estimar, identificar e localizar PT e PC com o objetivo de direcionar as ações das concessionárias na redução de perdas.

Este artigo inicialmente apresenta o estado da arte na estimação e identificação de perdas e a metodologia proposta. Posteriormente são apresentados os resultados para um sistema teste proposto, sendo descrita a maneira como o método pode ser aplicado em sistemas de distribuição reais. Como produto final, são apresentados os resultados para um alimentador urbano da CHESP, apresentando-se também a interface gráfica desenvolvida no projeto de P&D.

II. ESTADO DA ARTE

A. Reconhecimento de Padrões Aplicados a PC

Sistemas de reconhecimento de padrões podem ser capazes de detectar automaticamente novos padrões, bem como buscar padrões já conhecidos por experiência humana. A taxa de sucesso em inspeções visando encontrar clientes irregulares é a variável a ser maximizada por esses métodos, quando aplicados ao problema de identificação de PC. A tarefa desses sistemas deve ser classificar unidades consumidoras em duas classes: suspeitas de irregularidade e normal. Inicialmente, o algoritmo de classificação deve ser treinado para aprender os padrões dessas duas classes. O aprendizado dos métodos aqui expostos pode ocorrer segundo dois paradigmas: supervisionado e não-supervisionado. O aprendizado de classificadores supervisionados se assemelha ao aprendizado com um professor e requer a existência de amostras rotuladas, ou seja, resultados de inspeções em unidades consumidoras, permitindo que o classificador seja apresentado a elementos de cada classe. Já o aprendizado não-supervisionado não precisa de amostras rotuladas, mas sim de um modelo para representação de dados, como por exemplo o modelo da distribuição dos dados no espaço amostral.

Em [8] é utilizado o classificador supervisionado Máquina de Vetor de Suporte (MVS) para fazer a identificação de PC otimizando os parâmetros do classificador por meio de algoritmos genéticos. Em [9] é utilizado um algoritmo de análise conglomerados para agrupar consumidores com perfis de carga semelhantes e separar os valores aberrantes. O classificador supervisionado *Online Sequential Extreme Learning Machine* (OS-ELM) é utilizado para classificar de perfis de carga como normais ou anormais. Em [10] e [11]

um método de análise de conglomerados é utilizado para identificar comportamento anormal de consumidores. Nestes métodos, a classificação dos consumidores é feita individualmente, com vários classificadores diferentes, como árvores de decisão, redes bayesianas e modelos de detecção de quedas bruscas de consumo baseados no coeficiente de Pearson. Em [4], é utilizado o classificador FCO com diferentes algoritmos de aprendizado. Os bons resultados apresentados na literatura utilizando o FCO, indicam este método como o mais promissor no reconhecimento de padrões, incluindo consumidores suspeitos de possuírem PC.

B. Estimação de PT e PC

Na literatura, as PT e PC são usualmente estimadas através de métodos baseados em fluxo de carga e previsões de carga. Em [12], considerando previsões de cargas dadas de acordo com [13], um fluxo de carga é utilizado para estimar as PT em cada segmento do sistema de distribuição. A partir do cálculo das PT, as PC são estimadas pela diferença entre a energia entregue no sistema e a soma da energia faturada com as PT. Portanto, em [12], as perdas são estimadas, porém, somente as PT são identificadas por segmentos e, adicionalmente, o cálculo das PC é, na verdade, uma estimação das PC mais as PT geradas por estas, conforme ressalta [14].

Com o objetivo de melhorar a estimação das perdas, [14] propõe uma metodologia utilizando fatores de correção. A metodologia aplica a estimação de PT de acordo com [12] e considera que a diferença entre a energia medida e a soma da energia faturada com as PT total a jusante desta medição representam não só as PC, mas PC e as PT geradas por estas PC. Assumindo algumas aproximações em [14], é possível aplicar fatores de correção em todas as cargas a jusante da medição. Este processo de aplicação de fatores de correção pode ser repetido até os valores de correção se tornarem inferiores a uma dada tolerância. Portanto, o método estima e identifica as PT e PC de forma mais precisa, porém, o fluxo de carga é realizado diversas vezes, aumentando o custo computacional e fazendo com que as PC sejam distribuídas em igual proporção às cargas a jusante desta medição, conforme ressalta [15]. Em [16], as PC são identificadas através da análise de resíduos do EE-MQP e análise de variância nas curvas de cargas diárias dos consumidores. O método considera a existência de uma rede de distribuição inteligente com uma infraestrutura avançada de medição, onde o consumo de todos os consumidores é medido em tempo real. Por outro lado, é necessário ressaltar que sistemas típicos de distribuição de energia possuem um número restrito de medições. A aplicação do estimador de estados, nestes casos, é realizado através da combinação de medições e previsões de cargas, chamadas no estimador de estados de pseudomedições. Conforme é exposto em [17], a análise dos resíduos do EE-MQP torna-se muito falha para estes casos de baixa redundância de medição (relação entre número de medições (ou pseudomedições) e estados a estimar).

Como será melhor explicado na metodologia proposta, este trabalho pretende utilizar dados de sistemas típicos de distribuição de energia, ou seja, com um número restrito de medidores, segregando as PT por segmentos linhas, transformadores e medidores e as PC por transformadores. Para isso, é proposta uma metodologia híbrida.

III. METODOLOGIA PROPOSTA

A metodologia desenvolvida consiste da junção de duas abordagens distintas para o monitoramento de perdas em sistemas de distribuição. Para isso, torna-se necessário conhecer os estados do sistema, o que é realizado por meio da utilização de um estimador de estados trifásico. O grande desafio desta abordagem é obter resultados confiáveis em um sistema cujo número de medições é inferior ao número de estados a determinar, além de possuir PC, que é consumo de energia não faturado. Para a obtenção do número de medições necessárias são utilizadas pseudomedições, que são valores estimados de potência consumida baseada em projeções de consumo ou comportamento típico dos clientes do sistema analisado.

As PC são tratadas como resíduos das medições e pseudomedições. Para melhorar a qualidade da estimação de estados e possibilitar a localização precisa de PC, um método de reconhecimento de padrões é empregado. Este método analisa a base de dados dos clientes da empresa e busca por padrões considerados anômalos ou correlacionados com padrões de clientes conhecidamente irregulares para indicar UC candidatas à inspeção. Supondo que este método seja capaz de indicar com grande taxa de acerto a maior parte desses clientes irregulares, essa informação adicional pode ser passada para o estimador de estados, refinando o seu desempenho na estimação dos estados, que tem como consequência a contabilização de PC e PT de forma mais precisa. A variável que contabiliza essa informação é a proporção de clientes suspeitos que cada transformador possui. Considerando que cada transformador possui uma estimativa de consumo, a proporção de clientes irregulares é utilizada como fator para ponderar a confiabilidade destas pseudomedições.

A metodologia é resumida na figura 1. A base de dados fornece informações para o sistema de reconhecimento de padrões e para o estimador de estados trifásicos. O sistema de reconhecimento de padrões identifica os clientes suspeitos, sugerindo inspeções, e fornece os pesos para o estimador de estados.

A. Módulo de Reconhecimento de Padrões

As empresas em que a metodologia pode ser aplicada, provavelmente não possuem a mesma disponibilidade de dados, portanto, o sistema de reconhecimento de padrões foi desenvolvido para ser adaptável à disponibilidade de dados.

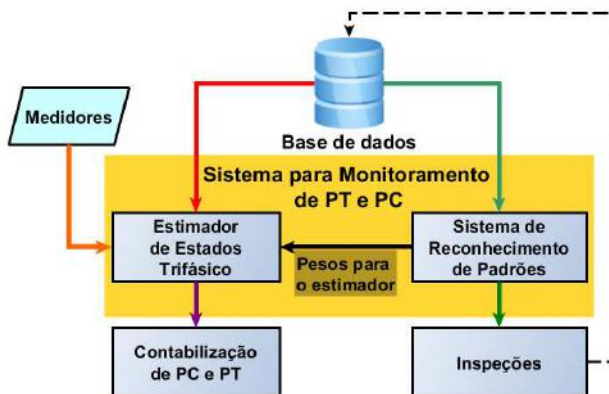


Figura 1. Fluxograma do método proposto.

Uma das informações é a proveniente dos resultados de inspeções realizados no campo. Se esta informação está disponível, ensina-se o sistema a identificar padrões parecidos com aqueles já detectados em inspeções, sendo esta abordagem conhecida como aprendizado supervisionado. Neste caso, o método escolhido é o classificador FCO. Se esta informação não é disponível, utiliza-se o método de DA. O resultado de ambos métodos é uma lista de clientes suspeitos de fraude. A figura 2 mostra de forma sintética as etapas empregadas pelo módulo de reconhecimento de padrões proposto.

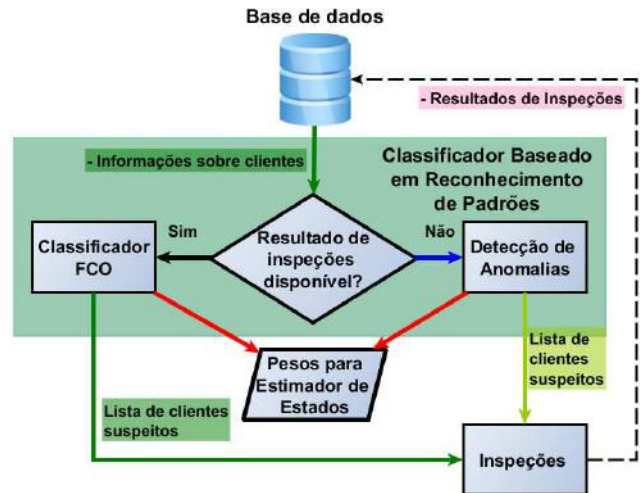


Figura 2. Sistema de reconhecimento de padrões.

B. Floresta de Caminhos Ótimos

O FCO é um classificador baseado em grafos, existindo nas formas supervisionada e não-supervisionada [4]. Na sua forma supervisionada, a partir de um conjunto de treinamento cuja classe é conhecida *a priori*, o FCO é capaz de aprender os padrões contidos neste conjunto e prever a classe de amostras cujo rótulo é desconhecido com grande precisão e baixo custo computacional [4]. Outra vantagem do FCO é que o mesmo é implicitamente multiclasse, o que, no caso da sua aplicação ao problema dicotômico de localização de PC, permite o aprendizado de subclasses das classes normal e irregular.

O projeto do classificador FCO em sua versão supervisionada é realizado em duas etapas: obtenção de protótipos e treinamento do classificador. A primeira etapa tem como objetivo identificar quais amostras do grupo de treinamento, chamadas Z_i , são as mais relevantes, ou seja, aquelas que melhor sintetizam o padrão dos elementos pertencentes a cada classe de amostras. Estas amostras são chamadas de protótipos, que formam o conjunto chamado S . As amostras são representadas no espaço de características por um vetor de características v , cujo número de dimensões é n . A métrica para definir dissimilaridade entre duas amostras é a distância euclidiana d entre os vetores de características de cada amostra.

O conjunto de treinamento é representado por um grafo completo e os protótipos são obtidos a partir da relação teórica entre a floresta de caminhos ótimos e árvore geradora mínima de um grafo [4]. Segundo essa relação, os protótipos são todos os nós da árvore geradora mínima que estão conectados a um nó de classe diferente da sua.

Depois de obtido o conjunto S , pode-se realizar o treinamento do classificador por meio do algoritmo de Floresta de Caminhos Ótimos [18]. O objetivo deste algoritmo é encontrar, a partir de um grafo completo, uma floresta de caminhos de custo mínimo que, por sua vez, são enraizadas nos nós protótipos e possuem apenas elementos de uma mesma classe de Z_I . O critério para a conquista de amostras é a menor função de custo de caminho. O FCO utiliza f_{max} como função de custo de caminho, dada por (1).

$$f_{max}(\langle s \rangle) = \begin{cases} 0, & s \in S \\ +\infty, & \text{caso contrário} \end{cases} \quad (1)$$

$$f_{max}(\pi \cdot \langle s, t \rangle) = \max\{f_{max}(\pi), d(s, t)\}$$

Nesta expressão, π é um caminho, s é uma amostra e $\pi \cdot \langle s, t \rangle$ é a concatenação do caminho π , que termina na amostra t e uma amostra s . Um caminho é dito ótimo quando sua função custo é mínima quando comparada a qualquer outro caminho que termina na mesma amostra.

O algoritmo FCO atribui um caminho ótimo $P^*(s)$ de S para cada amostra s pertencente a Z_I , formando uma floresta de caminhos ótimos P . Cada nó do grafo herda o rótulo do nó raiz do caminho ao qual pertencem. $P(s)$ contém o predecessor de cada nó no caminho, e recebe um marcador nulo quando o nó é uma raiz. $C(s)$ contém o custo de cada amostra, conforme calculado pelo algoritmo. O rótulo fornecido pelo FCO a cada amostra é dado pelo mapa de rótulos $L(s)$. Ao final do treinamento do FCO, obtém-se os mapas de custo ($C(s)$), rótulo ($L(s)$) e precedência ($P(s)$) para cada elemento s de Z_I .

A classificação dos elementos do conjunto de teste é feita elemento a elemento, inserindo uma amostra no grafo floresta de caminhos ótimos. Esta amostra é conectada por uma aresta a cada elemento da floresta de caminhos ótimos. Deseja-se então descobrir a qual protótipo esta amostra é mais fortemente conectada. Para isso recorremos outra vez à função f_{max} , calculando o seu valor para cada elemento da floresta de caminhos ótimos obtida na etapa de treinamento. Escolhe-se então aquela que possui menor valor de f_{max} , que conquistará a amostra classificada, passando a esta seu rótulo. O valor de f_{max} em questão será o valor da função custo da nova amostra. De maneira mais formal, deve-se calcular (2) para obter o custo da amostra t a classificar, e dar a esta o rótulo da amostra s cujo f_{max} é mínimo.

$$C(t) = \min\{\max\{C(s), d(s, t)\}\}, \forall s \in Z_I \quad (2)$$

Para identificação de consumidores suspeitos de possuírem PC, por exemplo, o vetor de características pode ser formado pelos dados de fatura de energia dos consumidores, pois esta informação é diretamente afetada por roubo ou fraude de energia. Outras informações podem ser utilizadas, dependendo da disponibilidade de dados, como a localização geográfica, por exemplo. Em [19]-[20] estes fatores são melhor explorados.

C. Detecção de Anomalias

O método de detecção de anomalias tem o objetivo de encontrar padrões em um conjunto de dados que não estão

conformes com o comportamento esperado [21]. Como outros métodos de reconhecimento de padrões utilizados para identificação de perdas não comerciais, o objetivo do DA é reconhecer os padrões de consumidores irregulares no banco de dados do sistema de distribuição. Da mesma forma que FCO, a análise é feita sobre os dados de fatura de energia dos consumidores ou outras informações disponíveis.

No problema de identificação de PC, a maioria dos consumidores têm um comportamento normal, enquanto que um pequeno número têm algum tipo de irregularidade associada à uma PC. A hipótese feita usando o método de DA é que estes consumidores irregulares divergem da distribuição estatística esperada para os dados, ou seja, possuem comportamento anômalo. Com o objetivo de considerar múltiplos comportamentos irregulares, os consumidores são divididos em grupos relacionados com as suas atividades: residencial, industrial, rural e comercial.

Na metodologia proposta, é aplicada uma DA estatística e paramétrica baseada em um modelo Gaussiano multivariado. O método supõe que a maioria dos dados dos consumidores formam uma distribuição normal de múltiplas variáveis, os consumidores com comportamento próximo da média da distribuição são considerados regulares e aqueles com comportamento distante da média são considerados suspeitos. Esta informação é modelada matematicamente pela função densidade de probabilidade, apresentada em (3).

$$f(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{2\pi^d |\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \quad (3)$$

Onde \mathbf{x} é o ponto no qual a função densidade de probabilidade é calculada, $\boldsymbol{\mu}$ é o vetor que contém os valores médios da distribuição para cada dimensão, $\boldsymbol{\Sigma}$ é a matriz de covariância e d é o número considerado de dimensões. Os consumidores com uma função densidade de probabilidade abaixo de um certo limiar são considerados anômalos. A expressão (4) modela matematicamente o limiar de detecção de anomalias.

$$f(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) < \varepsilon \quad (4)$$

Onde ε é o limiar de decisão. Logo, toda amostra que possuir densidade de probabilidade inferior a ε é classificada como suspeita. As demais possuem um comportamento estatisticamente semelhante ao do cliente médio da sua classe e portanto são consideradas normais. A seleção do limiar envolve um compromisso entre cobrir todos os consumidores anômalos e corretamente classificar aqueles que são irregulares.

D. Estimador de Estados e Perdas usando Metodologia Híbrida

O EE-MQP trifásico aplicado a sistemas de distribuição é apresentado na figura 3.

Os elementos do sistema provenientes de um banco de dados do sistema de distribuição são modelados como impedâncias e pseudomedições, como será descrito posteriormente, e são utilizados juntamente com as medições para a realização do EE-MQP trifásico.

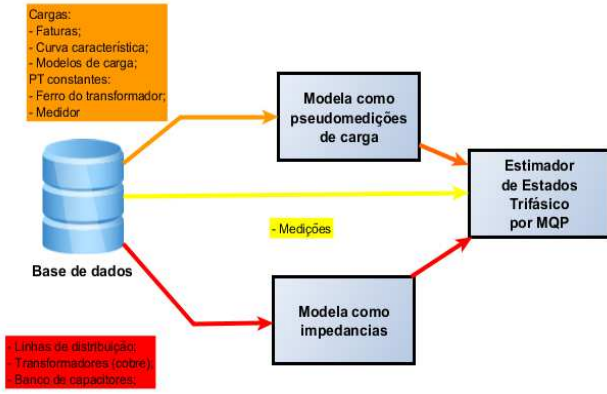


Figura 3. Estimação de estados através do EE-MQP trifásico.

O problema do EE-MQP é dado em (5). O objetivo do método é estimar o vetor de estados \mathbf{x} (magnitude e ângulo das tensões) e, como consequência, estimar as demais grandezas elétricas (fluxos e injeções de potência) de forma que as estimações sejam semelhantes as grandezas elétricas medidas.

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) \quad (5)$$

O vetor de medições \mathbf{z} é composto por medições reais nos alimentadores ou transformadores (fluxos de potência, injeção de potência, magnitudes de tensão) e predições de carga nos transformadores. O vetor de estimações $\mathbf{h}(\mathbf{x})$ é dado pelas equações, demonstradas em (6)-(9):

$$P_{k-i} = V_{k-i} \sum_{j \in \phi} \sum_{m \in K} V_{m-j} (G_{km-ij} \cos \theta_{km-ij} + B_{km-ij} \sin \theta_{km-ij}) \quad (6)$$

$$Q_{k-i} = V_{k-i} \sum_{j \in \phi} \sum_{m \in K} V_{m-j} (G_{km-ij} \sin \theta_{km-ij} - B_{km-ij} \cos \theta_{km-ij}) \quad (7)$$

$$P_{km-i} = V_{k-i} \sum_{j \in \phi} (-V_{k-j} (G_{km-ij} \cos \theta_{kk-ij} + B_{km-ij} \sin \theta_{kk-ij}) + V_{m-j} (G_{km-ij} \cos \theta_{km-ij} + B_{km-ij} \sin \theta_{km-ij})) \quad (8)$$

$$Q_{km-i} = V_{k-i} \sum_{j \in \phi} (-V_{k-j} (G_{km-ij} \sin \theta_{kk-ij} - B_{km-ij} \cos \theta_{kk-ij}) + V_{m-j} (G_{km-ij} \sin \theta_{km-ij} - B_{km-ij} \cos \theta_{km-ij})) \quad (9)$$

Onde V_k é a tensão na barra k . P_{k-i} , Q_{k-i} , P_{km-i} e Q_{km-i} são, respectivamente, as injeções de potência ativa e reativa na fase i do nó k e os fluxos de potência ativa e reativa na i do nó k para o nó m . θ_{km-ij} é a diferença angular entre o ângulo da fase i do nó k e a fase j do nó m . ϕ é o conjunto de fases existentes no nó k e K é o conjunto de nós conectados a barra k , incluindo o próprio nó k . G_{ij-km} e B_{ij-km} são elementos da matriz admitância de barras: $Y_{ij-km} = G_{ij-km} + jB_{ij-km}$.

Nos transformadores onde existem medições, o vetor \mathbf{z} vai conter um valor de potência injetada medida, por outro lado, quando as injeções são dadas por predições de carga, através [22], é necessário considerar um modelo de carga. Importante ressaltar que os modelos de carga influenciam a

estimação dos estados e das perdas, como apresentado em [23]-[25].

Em estudos de regime permanente, a variação na frequência é normalmente desconsiderada e a carga é dependente da tensão [26]. Tradicionalmente, a carga é modelada como potência constante, corrente constante, impedância constante ou uma combinação destes modelos, conhecido como modelo ZIP [27]. A carga modelada como potência constante não altera com a variação da tensão, a carga modelada como corrente constante varia linearmente com a corrente e na carga dada como impedância constante, a potência tem uma relação quadrática com a variação de tensão.

Analizando de forma mais ampla, os modelos de carga tradicionais são casos específicos dos modelos de carga exponencial [28]. No modelo de carga exponencial, coeficientes são utilizados para representar a sensibilidade da carga com a variação de tensão. Os coeficientes podem ser obtidos com base nos tipos de equipamentos [28], tipo de consumidores [26] ou através de medições com variações nos *taps* dos transformadores [29]. Portanto, a carga em uma barra ou transformador pode ser modelado com uma soma de parcelas, cada uma delas relacionada com um coeficiente de carga, demonstrado em (10) e (11). A Tabela I mostra os coeficientes de carga relacionados com os tipos de consumidores [26] e com os modelos de carga típicos [27].

$$P_k^{\text{predita}} = P_k^{\text{nom}} \sum_{i=1}^n a_i \left(\frac{V_k}{V_k^{\text{nom}}} \right)^{np_i} \quad (10)$$

$$Q_k^{\text{predita}} = Q_k^{\text{nom}} \sum_{i=1}^n b_i \left(\frac{V_k}{V_k^{\text{nom}}} \right)^{nq_i} \quad (11)$$

Onde V_k^{nom} é a tensão nominal da barra k . a_i e b_i são as parcelas de potência ativa e reativa nominais na barra k (P_k^{nom} e Q_k^{nom}), obtidas através de [22], relacionadas com os coeficientes dos modelos de carga para a parte ativa e reativa (np_i e nq_i). Como consequência, $\sum_{i=1}^n a_i = 1$ e $\sum_{i=1}^n b_i = 1$. n é o número de diferentes modelos de carga exponencial, e também, o número de diferentes parcelas de carga. Na expressão (10) e (11) as fases das injeções de potência são omitidas, pois as equações se aplicam a todas as fases. A Tabela I mostra que os coeficientes de carga possuem uma variação elevada, os motivos são as diversidades de cada região, como características socioeconômicas e climáticas.

TABELA I. COEFICIENTES DOS MODELOS DE CARGA

Modelo de Carga	nq	np
Potência Constante	0	0
Corrente Constante	1	1
Impedância Constante	2	2
Consumidor Residencial	0,9 a 1,7	2,4 a 3,1
Consumidor Comercial	0,5 a 0,8	2,4 a 2,5
Consumidor Industrial	0,1 a 1,8	0,6 a 2,2

Considerando as medições reais e as predições de carga (pseudomedições), o número de medições é superior que o número de estados a estimar. Considerando este fator e dado o fato que $\mathbf{h}(\mathbf{x})$ é composta por equações contínuas não

lineares (conforme (6)-(9)), o sistema apresentado em (5) pode ser resolvido iterativamente por mínimos quadrados ponderados utilizando equações normais.

Portanto, considerando o uso do processo iterativo, uma condição inicial é considerada (tensão de todas as barras iguais a tensão de referência) e em cada iteração v , o sistema linearizado apresentado em (12) é solucionado para obter as variações nos estados $\Delta \mathbf{x}^v$ e os estados são atualizados por (13). O processo iterativo finaliza quando as variações dos estados são menores que uma dada tolerância.

$$\left(\mathbf{H}(\mathbf{x}^v)^T \mathbf{W} \mathbf{H}(\mathbf{x}^v) \right) \Delta \mathbf{x}^v = \mathbf{H}(\mathbf{x}^v)^T \mathbf{W} \left[\mathbf{z} - \mathbf{h}(\mathbf{x}^v) \right] \quad (12)$$

$$\mathbf{x}^{v+1} = \mathbf{x}^v + \Delta \mathbf{x}^v \quad (13)$$

Onde \mathbf{H} é a matriz de derivadas parciais de $\mathbf{h}(\mathbf{x})$ e \mathbf{W} é a matriz de pesos. \mathbf{W} é uma matriz diagonal, onde cada elemento da diagonal principal é o inverso da covariância de medição, em outras palavras, $w_{ii} = \sigma_i^{-2}$, onde σ_i é o desvio padrão da medição z_i .

Usualmente, para medições reais, é feita uma relação entre o desvio padrão e a precisão do medidor [25]. Para pseudomedições, normalmente é considerado que todos os pesos são iguais, pois eles são dados pelo mesmo procedimento de predição de carga.

Neste ponto, que os autores apresentam uma contribuição clara: ao invés de utilizar os mesmos pesos para todas as pseudomedições, é proposto utilizar o resultado do reconhecimento de padrões para ponderar as pseudomedições de cada transformador de medição. Considerando que as predições de carga são baseadas nas faturas de energia conectadas a este transformador, torna-se claro que a precisão desta pseudomedição está relacionada a precisão da informação representada pela fatura de energia.

No reconhecimento de padrões, como resultado, é obtido uma relação entre o número de consumidores suspeitos e o número de consumidores irregulares em cada transformador k (r_k). A formulação híbrida propõe utilizar este resultado e o valor de cada pseudomedição para calcular os pesos. Na expressão (14), é dado peso para a pseudomedição z_i . Em outras palavras, quanto mais suspeitos um transformador tem, menor o seu peso.

$$w_{ii} = c \frac{1}{z_i (0.001 + r_k)^{-2}} \quad (14)$$

Utilizando o reconhecimento de padrões, alguns transformadores não têm consumidores suspeitos, sendo assim, $r_k = 0$, portanto, para evitar problemas numéricos, é considerado um valor mínimo de $r_k = 0,001$ (0,1 % de suspeitos). c é uma constante para evitar que as pseudomedições de baixa magnitude e, como consequência, pequeno r_k , obtenham pesos superiores aos adotados para os pesos das medições reais. Outra maneira é considerar a constante $c = 1$ e limitar o peso a um valor máximo para as pseudomedições (inferior ao peso das medições reais), caso esse valor seja superado, adota-se esse valor máximo de peso para pseudomedições.

Como pode ser observado, o método proposto utiliza uma forma heurística de considerar os pesos, mas baseado em um método inteligente de reconhecimento de padrões.

Após obter os estados, a estimação das PT nos segmentos do sistema, representados pelas impedâncias, é dada pela diferença entre a energia injetada e entregue pelo segmento (linhas e cobre dos transformadores), como é apresentado em (15). Essa diferença é apresentada como uma soma em (15), pois os fluxos k a m e m a k têm direções opostas. Na aplicação em sistemas de distribuição reais, algumas PT são aproximadamente constantes, por exemplo, a perda nos medidores de energia e no ferro dos transformadores, e são modeladas como cargas com potência constante no estimador de estados. A estimação de PC em cada barra é dada pela diferença entre a potência estimada e a carga prevista (P_{k-i}^{predita} e Q_{k-i}^{predita}) para este transformador, como é apresentado em (16).

$$TL_{km} = \sqrt{\left(\sum_{i \in \phi} (P_{km-i} + P_{mk-i}) \right)^2 + \left(\sum_{i \in \phi} (Q_{km-i} + Q_{mk-i}) \right)^2} \quad (15)$$

$$NTL_k = \sqrt{\left(\sum_{i \in \phi} (P_{k-i} - P_{k-i}^{\text{predita}}) \right)^2 + \left(\sum_{i \in \phi} (Q_{k-i} - Q_{k-i}^{\text{predita}}) \right)^2} \quad (16)$$

Observe que as estimções de perdas estão em potência. Para obter estes valores em energia, é necessário multiplicar pelo período (em horas) que estes valores de potência são válidos.

IV. RESULTADOS EM SISTEMA TESTE

A. Sistema Proposto

Um sistema-teste foi desenvolvido para avaliar a metodologia proposta. O sistema é baseado no Sistema de 123 Barras da IEEE (<http://ewh.ieee.org/soc/pes/dsacom>) considerando-se cada carga como uma soma de consumidores residenciais cuja potência média é descrita em [30] e a distribuição estatística do consumo mensal é descrita em [22], resultando em mais de 12 mil consumidores. A sazonalidade foi representada por mudanças no consumo mensal e a variação considerada foi baseada no comportamento de consumidores residenciais de um alimentador da empresa CHESP em Ceres (GO). Um histórico de consumo mensal de 12 meses foi considerado nos testes realizados.

Adicionalmente, foi adotado um número de oito medições para o EE-MQP, um número plausível para um sistema de distribuição típico no Brasil. Além disso, o sistema foi dividido em quatro regiões geográficas. Os pontos onde as medições foram realizadas são mostrados na figura 4.

Primeiramente, um caso base foi desenvolvido para simular um sistema de distribuição com perdas. Neste caso base, um fluxo de carga trifásico foi aplicado considerando-se as potências de cada mês, que variam de acordo com a sazonalidade considerada.

Depois, PC foram adicionadas nos consumidores. As PC foram consideradas como fraude, dada por uma redução parcial na fatura de energia, e como furto, dado para uma redução a zero da fatura de energia, e como consequência,

para os dois casos, uma alteração nos valores de cargas previstas.

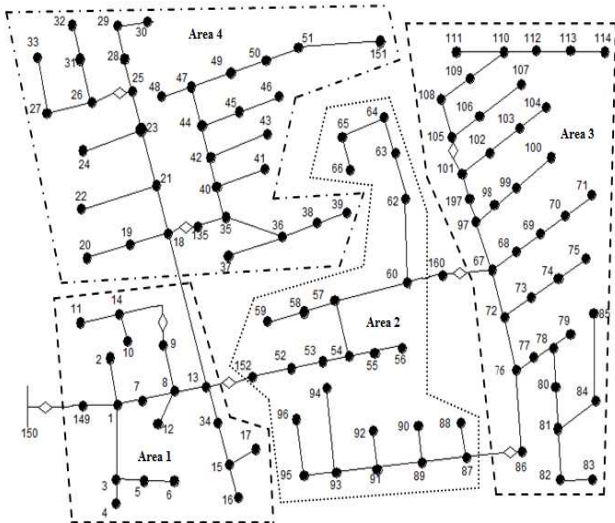


Figura 4. Sistema baseado no Sistema de 123 Barras da IEEE com quatro áreas e oito medidores (representados por losangos).

As PC consideradas representam 7 % do total da energia, número semelhante à média no Brasil, e são dadas pela redução nas faturas de energia. Os consumidores com PC foram escolhidos aleatoriamente, mas considerando que 99 % dos consumidores com PC estão localizados nas doze primeiras barras, ou seja, estão concentradas na área 1. O nível de redução na fatura de energia é dado de forma aleatória para 50 % dos consumidores ilegais, para os outros 50 % é considerada a redução para zero. O mês que a fraude no consumidor inicia é dada aleatoriamente e permanece para todas as meses seguintes a este.

B. Resultados do Reconhecimento de Padrões

Para avaliar o desempenho do classificador FCO foram escolhidas duas métricas. A primeira delas é o Valor Preditivo Positivo (VPP), que indica a parcela dos clientes corretamente considerados suspeitos sobre o total de clientes apontados como possivelmente irregulares. A segunda métrica é chamada de Sensibilidade (SENS), que mede a porcentagem do total de clientes irregulares que foram incluídos na lista de suspeitos. Cada unidade consumidora foi representada por um vetor contendo o seu histórico de consumo mensal do último ano analisado e também pelos dados categóricos provenientes das regiões geográficas definidas pela figura 4. Estes foram incorporados ao vetor de características por um método conhecido como *dummy encoding* [31]. Para analisar a influência do tamanho do conjunto de treinamento, foram considerados 4 diferentes tamanhos. O algoritmo foi executado dez vezes para cada variante. A média dos resultados do FCO para cada variação se encontra na Tabela II.

TABELA II. RESULTADOS DOS TESTES DO FCO NO SISTEMA-TESTE

Treinamento [%]	VPP [%]	SENS [%]
0,5	92,21	25,77
1,0	78,55	35,02
2,0	85,60	46,66
5,0	92,28	55,15

O DA foi desenvolvido considerando os doze meses de consumo dados pelas faturas de energia com PC e foi utilizado o parâmetro $\epsilon = 0,0001$. Os dados de consumo foram normalizados, dividindo os valores de consumo de cada mês pela média de consumo do período considerado (um ano). Uma transformação linear é aplicada ao valor da função densidade de probabilidade e seus valores variam de 0 a 1.

A metodologia de DA corretamente classificou 98,35 % dos consumidores. Entre os considerados anômalos, 77,54 % foram precisamente rotulados como consumidores irregulares. Isto significa que 84,03 % do total de consumidores irregulares pode ser localizada, isto na primeira iteração da aplicação da metodologia no sistema teste.

C. Resultados do Método Híbrido

O EE-MQP considerou as faturas de energia com erros e os valores reais de medição, obtidos através dos valores de referência. O estimador de estados foi executado doze vezes, considerando a potência prevista dada através das faturas de cada mês. Os pesos para os medidores reais foram considerados como $w_{ii} = 1000$ e a constante das pseudomedições adotada foi $c = (10000)^{-1}$. Os valores de r_k utilizados foram fornecidos pelo DA

A figura 5 mostra os valores de referência e os valores estimados através do método para as PT. Esta figura apresenta os valores de perdas em energia média mensal para todos os segmentos do sistema teste. Cada número no eixo horizontal corresponde a um segmento de linha do sistema teste.

A figura 6 mostra os valores de referência e estimados das PC. A figura mostra a perda média mensal de energia em cada nó do sistema. Cada número no eixo horizontal corresponde a um nó do sistema teste.

O método proposto obteve um bom desempenho na estimação de PT e PC como podem ser vistas nas figuras 5 e 6. O valor médio e máximo de erro para a estimação de PT foram de 0,002 MVAh e 0,13 MVAh, respectivamente. Para a estimação de PC, os erros médio e máximo foram de 0,12 MVAh e 4,29 MVAh.

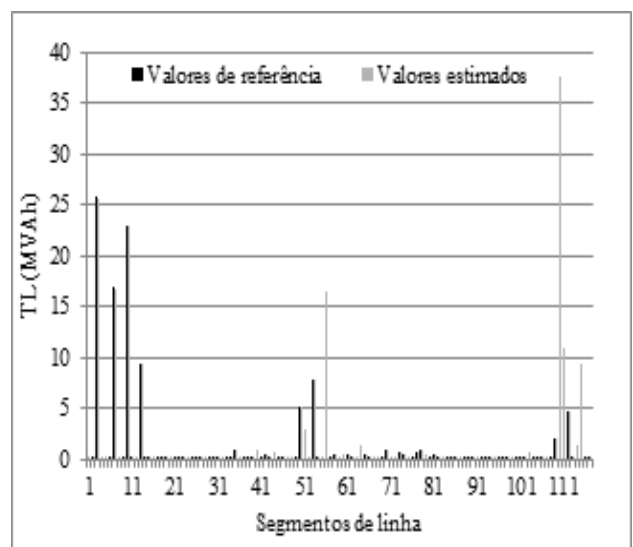


Figura 5. Valores de referência e estimados de PT nos segmentos de linha do sistema teste.

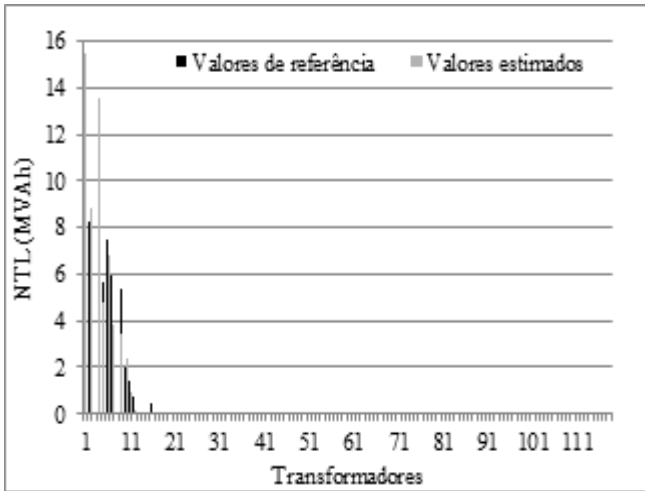


Figura 6. Valores de referência e estimados para PC nos nós do sistema teste.

V. APLICAÇÃO DO MÉTODO EM SISTEMAS REAIS

Este item tem o objetivo de apresentar algumas premissas adotadas para a aplicação do método em sistemas de distribuição reais. Algumas questões apresentadas são decorrentes da aplicação da metodologia nas empresas envolvidas no projeto.

A. Medidores e Predições de Carga

Para o sistema teste foram realizadas doze estimações, uma para cada mês. Porém, o medidor de energia desenvolvido no âmbito deste projeto pode realizar medições com um intervalo de tempo de minutos e enviar estas informações para o centro de operações através de um modem GPRS ou armazenar esta informação em um cartão SD para ser pós-processada. Este medidor é instalado no lado de baixa tensão dos transformadores e mede grandezas como tensão, injeção de corrente e potência, conforme é melhor descrito em [32]. Neste aspecto, o método híbrido pode ser realizado para cada horário, resultando em uma melhor precisão.

Porém, a maior parte dos transformadores não possui medidores, portanto os valores horários de injeção de potência são dados por uma metodologia de predição de carga, conforme [22]. Com base nesta referência, as curvas diárias de injeção de potência ativa e reativa nos transformadores são obtidas com base nas faturas dos clientes conectados ao transformador e nas campanhas de medição [33], onde o perfil diário de consumo e o fator de potência são mensurados. Na inexistência de campanhas de medição para alguns transformadores, é possível utilizar tanto os valores de campanha de medição de transformadores semelhantes ou as curvas obtidas dos transformadores com medidores.

O modelo de carga com respeito a variação com a tensão pode ser alterado de acordo com o conhecimento da concessionária ou utilizando valores médios padrões de acordo com o tipo de consumidor (residencial, comercial e industrial) de acordo com a Tabela I.

Portanto, o método híbrido pode estimar os estados e as perdas de forma horária considerando este conjunto de medições e pseudomedições.

B. Impedância das Linhas e Transformadores

Os sistemas de gerenciamento de algumas empresas trabalham com a informações das linhas em impedâncias de sequência. Como o método é desenvolvido em impedâncias de fases, uma aproximação pode ser feita através de (17).

Para considerar a PT no cobre do transformador, o transformador é considerado como um segmento do sistema através da sua impedância, dada conforme (18).

$$\mathbf{z}_{km} = \begin{bmatrix} z_{aa-km} & z_{ab-km} & z_{ac-km} \\ z_{ab-km} & z_{bb-km} & z_{bc-km} \\ z_{ca-km} & z_{cb-km} & z_{cc-km} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2z_1 + z_0 & z_0 - z_1 & z_0 - z_1 \\ z_0 - z_1 & 2z_1 + z_0 & z_0 - z_1 \\ z_0 - z_1 & z_0 - z_1 & 2z_1 + z_0 \end{bmatrix} \quad (17)$$

Onde z_1 e z_0 são as impedâncias de sequência positiva e negativa.

$$z_{1-km} = \frac{(V^{\text{nom}}/\sqrt{3})^2}{S^{\text{nom}}/3} z(\%) \quad (18)$$

Onde z_{1-km} e $z(\%)$ são, respectivamente, a impedância de fase do transformador e a impedância característica do transformador e S^{nom} é a potência aparente nominal.

C. PT nos Medidores e no Ferro dos Transformadores

Considerando que as PT nos medidores de energia e no ferro dos transformadores é aproximadamente constante [12], elas são consideradas no estimador de estados como cargas com potência constante e são somadas as cargas do transformador correspondentes aos consumidores. As PT nos transformadores devido aos medidores de energia são dados por (19).

$$PTM_k = n(ptm) \quad (19)$$

Onde PTM_k é a PT total dos medidores no transformador k , ptm é a PT característica do medidor de energia e n é o número de medidores (ou consumidores) conectados ao transformador k . Caso o valor característico não seja conhecido, pode-se utilizar um valor típico [12], por exemplo, $ptm = 1,2 W$. As PT nos transformadores devido as perdas no ferro dos transformadores, conforme (20).

$$PTT_k = ptt \quad (20)$$

Onde PTT_k é a PT devido a perda no ferro do transformador k e ptt é a perda característica no ferro deste transformador. Em alguns casos a perda característica não é fornecida, portanto, se utiliza um valor típico [14], dado pela expressão (21).

$$ptt = 15,71 \frac{Sn^{0,74}}{1000} W \quad (21)$$

D. Estimação de PC: Não Faturamento de Potência Reativa

A maioria das empresas de distribuição não fatura energia reativa de boa parte de seus consumidores. Portanto, apesar de esse consumo poder ser estimado, com base no uso de medições em transformadores ou através de predições de carga baseadas em campanhas de medição, esta é uma energia não faturada. Portanto, a PC é extratificada em duas parcelas: a perda devido ao não faturamento de energia reativa

e a perda devido a irregularidades. A segunda, é obtida através de (16), e a primeira pela diferença entre a potência aparente estimada e a potência real estimada, em cada transformador.

E. Redução do Sistema

Como as impedâncias são normalmente dadas pelos segmentos de linhas de poste em poste, o objetivo da redução do sistema é reduzir o número de postes considerados como barras para o estimador de estados, reduzindo o esforço computacional. Inicialmente, o método de redução elimina todos os postes terminais que não possuem cargas (consumidores solicitaram desligamento ou são utilizados somente em eventos). Posteriormente, são eliminados todos os postes que não são terminais, não possuem cargas e não originam ramificações no sistema de distribuição, as impedâncias destes segmentos são somadas e interpretadas como um único segmento de linha para o estimador de estados.

F. Reconhecimento de Padrões

As metodologias estudadas baseiam seu funcionamento na premissa de que pode-se inferir um comportamento dito suspeito ou anômalo a partir de dados relativos aos causadores desta anomalia, que neste caso são as unidades consumidoras. Dentre os dados obtidos, aqueles que mais se enquadram neste quesito são os dados de consumo mensal dos clientes. A aplicação descrita abaixo se refere aos dados apresentados pelo alimentador da CHESP, descrito no item VI, a aplicação em demais sistemas de distribuição é análoga.

Os clientes são classificados em 7 grupos: “Comercial”, “Consumo Próprio”, “Industrial”, “Poderes Públicos”, “Residencial”, “Rural” e “Serviço Público”. Há 7 consumidores pertencentes ao grupo tarifário A e 3789 ao grupo tarifário B.

Nenhum dado foi fornecido quanto à autuação por fraude, o que é necessário para a utilização de qualquer método supervisionado de classificação, como o FCO. Com relação aos pontos apresentados algumas premissas foram adotadas para aplicação da Detecção de Anomalias. Cada grupo de clientes foi representado por um modelo gaussiano distinto. Ou seja, a base de dados foi subdividida nas 7 classes consideradas. No entanto, os clientes das classes “Consumo Próprio”, “Poderes Públicos” e “Serviço Público” foram ignoradas por possuírem poucos clientes, da mesma maneira dos os clientes da classe tarifária A. O critério de avaliação utilizado foi o perfil de consumo de cada cliente durante o período de Janeiro de 2013 a Dezembro de 2013. Os dados de consumo mensal foram normalizados pelo maior consumo considerado no período avaliado. O parâmetro ϵ foi ajustado para o valor 0,0001, tornando a detecção de anomalias seletiva.

VI. RESULTADOS PARA ALIMENTADOR DE DISTRIBUIÇÃO REAL

Os resultados apresentados são para um alimentador do sistema de distribuição da CHESP. O sistema é radial em 13,8 kV, possui linhas trifásicas (área urbana) e monofásicas (área rural), mais de duzentos transformadores e mais de

três mil consumidores. No atual momento, existe uma medição no alimentador principal e mais dez medidores, conforme [32], em transformadores de distribuição.

A. Reconhecimento de Padrões

É importante destacar que, diferentemente do que ocorre com métodos baseados em fluxo de carga, a quantificação das perdas, tanto técnicas quanto não-técnicas, foge do escopo das metodologias baseadas em reconhecimento de padrões. No sistema analisado não havia disponibilidade de dados rotulados, portanto o método de DA foi utilizado. Os resultados se encontram na Tabela III.

TABELA III. RESULTADOS DA APLICAÇÃO DE DA NO SISTEMA-CHESP

Grupo	Analisados	Suspeitos	Ignorados
Residencial	3082	749	117
Rural	174	41	6
Industrial	19	0	0
Comercial	285	77	31
Outros	0	0	82
Total	3560	867	236

B. Metodologia Híbrida

O método foi aplicado considerando os pesos para os medidores reais como $w_{ii} = 1000$, a constante das pseudomedições adotada foi $c = 1$, porém, os valores dos pesos das pseudomedições foram limitados em $w_{ii} = 100$. Os valores de r_k utilizados foram fornecidos pelo DA.

Conforme descrito na metodologia proposta, o método estima as PT, extratificando-as em perdas nos segmentos de linhas, ferro e cobre dos transformadores e medidores de energia, e as PC, extratificando-as por transformadores, considerando as perdas por irregularidade e por não faturamento de energia reativa. Adicionalmente, as estimativas foram realizadas considerando o comportamento de consumo horário. Alguns resultados são apresentados, dada a quantidade grande de informações.

A figura 7 apresenta a estimativa de PT em um transformador de distribuição (PT elevadas), para um dia de semana. Nas PT estão contabilizadas as perdas no cobre e ferro do transformador e as perdas nos medidores de energia. A figura deixa clara a importância de considerar o comportamento horário, adicionando precisão ao método. Como pode ser observado, a curva característica de PT no transformador tem um comportamento semelhante a curva característica de consumo do mesmo.

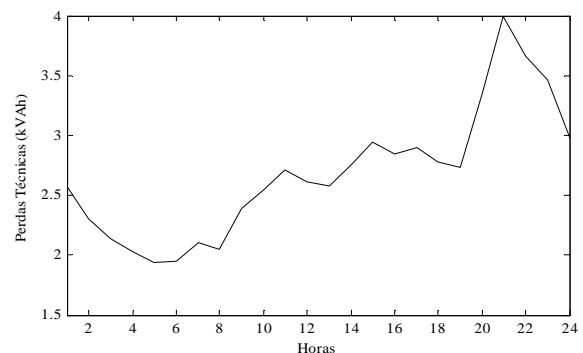


Figura 7. Comportamento horário de PT em transformador de distribuição, considerando um dia útil da semana.

A figura 8, de forma análoga, apresenta os valores de PC em um dia útil para um transformador de distribuição. Os valores apresentados contemplam as perdas por irregularidades e por não faturamento de potência reativa. Este transformador foi escolhido para demonstrar que, em alguns casos, temos uma PC com valor negativo. Este fato ocorre quando o valor estimado de potência para este horário é inferior ao valor dado pela predição de carga. Porém, ao considerar a PC ao longo de todo dia, temos um valor positivo. Portanto, a variação ao longo do dia é significativa, e a sua consideração é significativa na precisão do método.

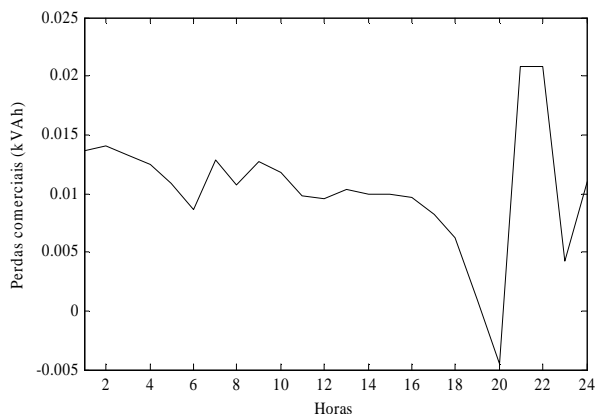


Figura 8. Comportamento horário de PC em transformador de distribuição, considerando um dia útil da semana.

A figura 9, apresenta as PT extratificadas para cada tipo de elemento do sistema de distribuição. A figura mostra o somatório das perdas para cada tipo de elemento, por exemplo, a soma das perdas em todos segmentos de linhas do sistema de distribuição. Observe como as perdas nos segmentos de linha e no cobre do transformador variam de acordo com o comportamento de consumo, enquanto as perdas no ferro do transformador e nos medidores têm comportamento constante.

A figura 10 apresenta as PC para todo o alimentador de distribuição. Na figura, são apresentados os somatórios das perdas em todos os transformadores, extratificando as perdas devido a irregularidades e devido ao não faturamento de potência reativa.

A variação do comportamento ao longo do dia é significativo, porém, com o objetivo de guiar ações de combate para reduzir as PT e PC, primeiramente é necessário analisar quais PT e PC são mais significativas para o alimentador de distribuição, isto é apresentado na Tabela IV, onde temos o somatório das perdas ao longo de um mês e o seu valor percentual com relação as perdas globais (PG) deste mesmo mês.

Com base nestes resultados, ações podem ser tomadas baseadas em análises como: a PC é mais significativa que a PT para este alimentador, sendo que sua parcela mais significativa é decorrente de irregularidades; dentro das PT, os valores mais significativos ocorrem nos segmentos de linha do sistema. Adicionalmente, o método estima as PT e PC em todos segmentos de linhas e transformadores.

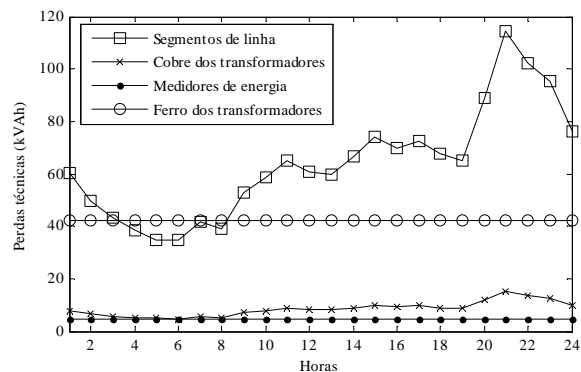


Figura 9. Comportamento horário das PT, considerando o somatório de todos os elementos do alimentador de distribuição em um dia útil da semana.

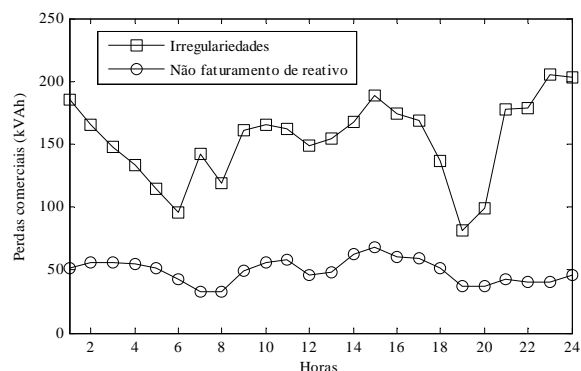


Figura 10. Comportamento horário das PC, considerando o somatório de todos os transformadores do alimentador de distribuição em um dia útil da semana.

As figuras 11 e 12 mostram as PT e PC mensais totais para cada transformador do alimentador de distribuição.

TABELA IV. EXTRATIFICAÇÃO DAS PERDAS NO ALIMENTADOR ESTIMADAS AO LONGO DE UM MÊS

Tipo de Perda	kVAh	Percentual da PG
PT nos segmentos de linha	45908	20%
PT no cobre dos transformadores	6127	3%
PT nos medidores de energia	3238	1%
PT no ferro dos transformadores	30317	13%
PT totais	85590	37%
PC devido a irregularidades	110322	48%
PC devido ao não faturamento de reativo	35329	15%
PC totais	145651	63%
Perdas globais (PG)	231242	100%

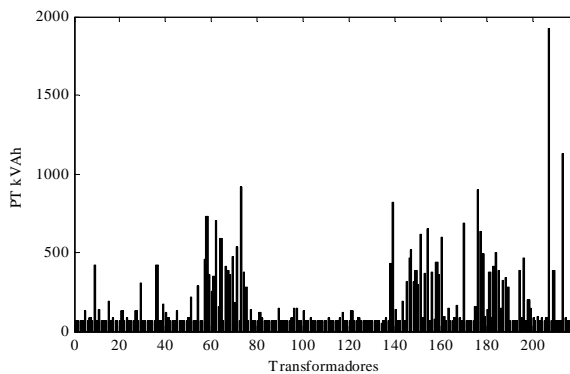


Figura 11. PT totais em cada transformador do alimentador de distribuição, considerando o período de um mês.

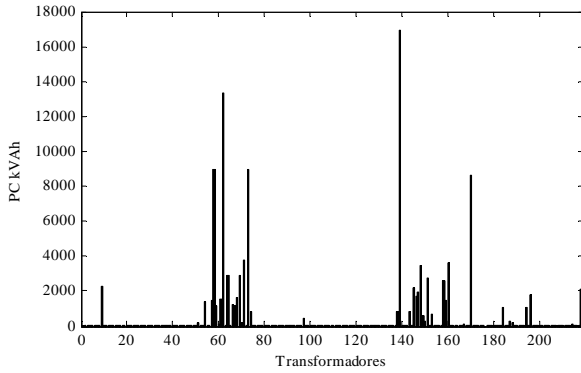


Figura 12. PC totais em cada transformador do alimentador de distribuição, considerando o período de um mês.

C. Interface Homem-Máquina

No sistema integrado, o programa SQLITE gerencia as entradas e saídas e fornece uma interface homem-máquina simples com os resultados da estimação e identificação de PT e PC em todos os transformadores do sistema de distribuição. O programa acessa o banco de dados das empresas (por exemplo, plataforma USEALL) e os dados obtidos dos medidores, transformando estas informações em arquivos textos que são processados na metodologia híbrida, desenvolvida em MATLAB. A metodologia fornece arquivos textos como saída, apresentados em uma interface homem-máquina através do SQLITE.

Os resultados da metodologia são apresentados em uma interface homem-máquina. Esta interface apresenta o mapa do sistema de distribuição, seus postes, linhas e transformadores. Na interface é possível selecionar a grandeza que se deseja analisar, PT, PC ou PG. A Figura 13 apresenta os resultados para o alimentador de distribuição analisado.

A interface apresenta de forma gráfica, através de cores, quais são os transformadores críticos, ou seja, que apresentam perdas estimadas mais elevadas. A criticidade pode ser definida pelo usuário, baseando-se em critérios como: PT, PC ou PG em kVAh ou em sua porcentagem com relação a potência nominal do transformador; número de clientes suspeitos e porcentagem de clientes suspeitos com relação ao total de clientes. A figura 14 apresenta esta tela de definição de critérios.

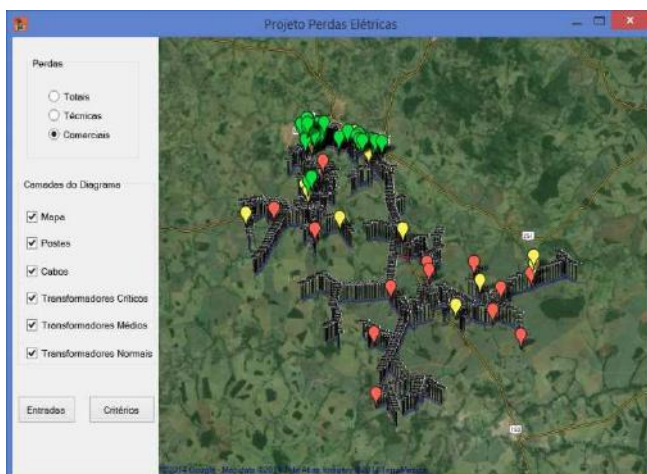


Figura 13. Interface homem-máquina apresentando os transformadores críticos (perdas altas ou número elevado de suspeitos).

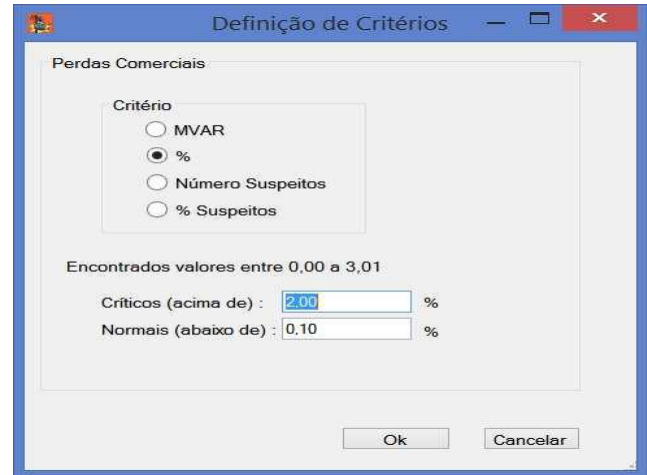


Figura 14. Opções de definição de criticidade para os transformadores.

Adicionalmente, é possível selecionar um dos transformadores, conforme figura 15, e visualizar os detalhes de perdas estimadas e os clientes suspeitos de acordo com o método de reconhecimento de padrões. Portanto, a interface homem-máquina é visual e de fácil utilização, permitindo que decisões possam ser tomadas para combater as perdas nos sistemas de distribuição.

Unidade Consumidora	Fases	Classe	Grupo
50100054	ABC	Comercial	B
50100516	ABC	Comercial	B
50300704	ABC	Residencial	B
50301312	ABC	Residencial	B
50301405	ABC	Residencial	B
50304682	ABC	Residencial	B
50305864	ABC	Residencial	B
50305939	ABC	Residencial	B
50105859	ABC	Residencial	B
50105933	ABC	Residencial	B
50301091	ABC	Residencial	B
50301092	ABC	Residencial	B
50302487	ABC	Residencial	B
50304182	ABC	Residencial	B
50105675	ABC	Residencial	B
50300705	ABC	Residencial	B

Figura 15. Detalhes de perdas e consumidores suspeitos em transformador.

VII. CONCLUSÕES

Este trabalho apresenta um sistema integrado desenvolvido no âmbito de um projeto de P&D para estimar e identificar PT e PC, de forma que seja possível tomar ações de combate para reduzi-las. O sistema integrado contempla uma metodologia híbrida nova, baseada em EE-MQP e métodos de reconhecimento de padrão (FCO e DA), cujos resultados são promissores e superiores as demais metodologias de estimação de perdas da literatura. Adicionalmente, estes resultados são apresentados em uma interface homem-máquina visual e simples, porém, que contempla as informações mais importantes.

VIII. AGRADECIMENTOS

Os autores gostariam de agradecer à Universidade Federal do Rio Grande do Sul, ao Conselho Nacional de Desen-

volvimento Científico e Tecnológico e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes). Os autores agradecem igualmente às empresas CHESP, CERRP, CERPRO, CERNHE, CERIPA, CERAL-DIS, CETRIL, CERIM, CERMC, CERIS, CEDRI, CERES, CEDRAP, ELFSM, EFLJC, COOPERALIANÇA e CERCOS pelo suporte na realização do projeto de P&D ANEEL 0103-0002/2011.

IX. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] T. B. Smith, "Electricity theft: a comparative analysis," *Energy Policy*, vol. 32, no. 18, pp. 2067-2076, Dec. 2004.
- [2] R. Vidinich and G. A. L. Nery, "Pesquisa e desenvolvimento contra o furto de energia," *Revista de Pesquisa e Desenvolvimento da ANEEL*, vol. 3, pp. 15, Jun. 2009.
- [3] J. Aguero, "Improving the efficiency of power distribution systems through technical and non-technical losses reduction", in *Proc. 2012 IEEE PES Transmission and Distribution Conf. and Exposition*, pp.1-8.
- [4] C. C. O. Ramos, A. N. de Souza, J. P. Papa e A. X. Falcão, "A new approach for nontechnical losses detection based on optimum-path forest," *IEEE Trans. Power Syst.*, vol.26, no.1, pp. 181-189, Feb. 2011.
- [5] M. F. Shaaban, Y. M. Atwa and E. F. El-Saadany, "DG allocation for benefit maximization in distribution networks," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 639-649, May 2013.
- [6] R. S. Zulpo, R. C. Leborgne and A. S. Bretas, "Optimal siting and sizing of distributed generation through power losses and voltage deviation," in *Proc. 2014 IEEE International Conf. on Harmonics and Quality of Power*, pp. 871-875.
- [7] V. H. M. Quezada, J. R. Abbad and T. G. S. Román, "Assessment of energy distribution losses for increasing penetration of distributed generation," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 533-540, May 2006.
- [8] J. Nagi, S. K. Yap, S. K. Tiong, S. K. Ahmed and A. M. Mohammad, "Detection of abnormalities and electricity theft using genetic support vector machines," in *Proc. 2008 IEEE Region 10 Conf.*, pp. 1-6.
- [9] A. H. Nizar and Z. Y. Dong, "Identification and detection of electricity customer behavior irregularities," in *Proc. 2009 IEEE/PES Power System Conference and Exposition*, pp. 1-10.
- [10] C. León *et al.*, "Variability and trend-based generalized rule induction model to NTL detection in power companies," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1798-1807, Nov. 2011.
- [11] I. Monedero *et al.*, "Detection of frauds and other non-technical losses in a power utility using Pearson coefficient, Bayesian networks and decision trees," *Int. Journal of Electrical Power & Energy Systems*, vol. 34, pp. 90-98, Apr. 2012.
- [12] C. C. B. Oliveira *et al.*, "A new method for the computation of technical losses in electrical power distribution system," in *Proc. 2001 Int. Conf. on Electricity Distribution*, pp. 5.
- [13] J. A. Jardini, C. M. V. Tahan, M. R. Gouvea, S. U. Ahn, and F. M. Figueiredo, "Daily load profiles for residential, commercial and industrial low voltage consumers," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 375-380, Jan. 2000.
- [14] A. Meffé. "Metodologia para cálculo de perdas técnicas por segmento do sistema de distribuição," Dissertação de Mestrado, Escola Politécnica da Universidade de São Paulo, 2001.
- [15] A. Rossoni, "Estimação de perdas técnicas e comerciais: métodos baseados em fluxo de carga e estimador de estados," Dep. Eng. Elétrica, Univ. Federal do Rio Grande do Sul, 2014.
- [16] S. Huang, Y. Lo and C. Lu, "Non-Technical Loss Detection Using State Estimation and Analysis of Variance," *IEEE Trans. Power Syst.*, vol. 28, pp. 2959-2966, Aug. 2013.
- [17] C. Gómez-Quiles, A. Gómez-Exposito and A. V. Jaén, "State Estimation for Smart Distribution Substations," *IEEE Trans. Smart Grid*, vol. 3, pp. 986-995, June 2012.
- [18] J. P. Papa, A. X. Falcão, P. V. M. Miranda, C. T. N. Suzuki and N. D. A. Mascarenhas, "Design of robust pattern classifiers based on optimum-path forests," in *Proc. 2007 International Symposium on Mathematical Morphology*, pp. 337-348.
- [19] R. D. Trevisan *et al.* "Detecção e Identificação de Perdas Comerciais usando Curvas de Carga e Classificador Floresta De Caminhos Ótimos," in *Congreso Internacional de Distribución Eléctrica 2014*, pp. 1-5.
- [20] R. D. Trevisan, "Detecção e identificação de perdas comerciais em sistemas de distribuição: metodologia baseada em floresta de caminhos ótimos," Dep. Eng. Elétrica, Univ. Federal do Rio Grande do Sul, 2014.
- [21] V. Chandola, A. Banerjee and V. Kumar, 2009, "Anomaly detection: a survey," *ACM Comput. Surv.*, vol. 41, pp. 15:1-15:58, Jul. 2009.
- [22] J. A. Jardini and R. P. Casolari. (2014, Nov. 14). *Curvas de carga de consumidores e aplicações na engenharia da distribuição*. [Online]. Disponível: <http://social.stoa.usp.br/jajardini/livros>
- [23] A. Rossoni *et al.*, "Load models effects on distribution system losses estimation: a numerical study," in *Proc. 2013 IEEE Power and Energy Soc. General Meeting*, pp. 1-5.
- [24] R. de F. V. Dresch, "Análise do efeito da modelagem da carga nas estimativas de perdas elétricas em sistemas de distribuição," Dep. Eng. Elétrica, Univ. Federal do Rio Grande do Sul, 2014.
- [25] M. Karimi *et al.*, "Impact of Load Modeling in Distribution State Estimation," in *Proc. 2012 Int. Power Eng. and Optimization Conf.*, pp. 67-71.
- [26] J. Machowski, J. Bialek and J. Bumby, *Power system dynamics: Stability and Control*, 2.ed. New York: John Wiley & Sons, 2008. 688p.
- [27] W. H. Kersting, *Distribution system modeling and analysis*, 1.ed. New York: CRC Press LLC, 2002. 329p.
- [28] IEEE Task Force on Load Representation for Dynamics Performance, "Load Representation for Dynamics Performance Analysis," *IEEE Trans. on Power Syst.*, vol. 8, no. 2, May 1993.
- [29] T. Vinnal, K. Janson and H. Kalda, "Analysis of power consumption and losses in relation to supply voltage quality," in *Proc. 2009 European Conf. Power Electronics and Applicat.*, pp. 1-9.
- [30] IEEE/PES Distribution System Analysis Subcommittee (2004, Sept.). IEEE 123 Node Test Feeder. [Online]. Disponível: <http://ewh.ieee.org/soc/pes/dsacom/testfeeders/>.
- [31] K. P. Murphy, "Probability" in *Machine Learning: a Probabilistic Perspective*, Cambridge: The MIT Press, 2012, ch. 2, pp. 35.
- [32] N. Penner *et al.*, "Equipment for monitoring and combating of non-technical losses in distribution networks: design and preliminary results," in *Proc. 2014 Int. Universities. Power Eng. Conf.*, pp.1,4.
- [33] *Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – Módulo 2: Planejamento da Expansão do Sistema de Distribuição*. Agência Nacional de Energia Elétrica Disponível: http://www.aneel.gov.br/arquivos/PDF/Módulo2_Revisao_5.pdf