

Prospecting and Fraud Combat Actions Identification to Commercial Losses

H. O. Henriques, H. L. S. Almeida, M. Z. Fortes, *Member, IEEE*, F. B. Carregosa and W. L. Moura

Abstract—The traditional commercial losses prospect processes have not been effective. This paper presents a network inspection methodology using measurements at some strategic network points. Customers' loads and distribution transformers are provided for a priori, based on existing load typology of a distribution utility and based on a social research considering financial reserve and consumption habits. These predictions are represented as trapezoidal fuzzy numbers. Commercial losses are detected by the difference between power supply output measure reading and the distribution transformers loads sum added with the ohmic losses. The commercial losses are estimated to be solving a linear fuzzy system, which calculates the transformers and distribution customers hourly demand correction initially referred to inaccurately by load typology. The transformers and customers with larger correction values are selected for inspection.

Keywords—Load Curves, Fraud, Intelligent Inspections, Fuzzy Logic.

I. INTRODUÇÃO

A BUSCA por identificação e redução de perdas comerciais é um tema que a muitos anos é pesquisado por concessionárias de distribuição de energia e grupos de pesquisa das academias. Esta temática atinge a diversos países independente de situação econômica. Exemplos de pesquisas nesta área são discutidos em [1] que apresenta alguns dados da Mongólia e [2] no Iraque. Uma visão geral do impacto financeiro destas perdas para a concessionária pode ser mais bem entendido com o texto apresentado por Subramaniam et al. [3].

É importante destacar as pesquisas desenvolvidas na busca de equipamentos de detecção de fraudes, como um amperímetro adaptado para aplicação em sistemas de baixa tensão [4] e radar para identificação de conexões ilegais em eletrodutos enterrados [5].

Na busca por soluções as técnicas de inteligência artificial como ferramenta de suporte a tomada de decisão são amplamente aplicadas e conhecidas. O aprendizado de máquina e mineração de dados são técnicas que têm sido amplamente utilizadas por pesquisadores, para extensa análise inteligente de dados, com a finalidade de reconhecer padrões normais e anormais de comportamento. Em [6] é apresentada a aplicação de uma técnica de aprendizado de máquina para

analisar os dados de consumo de energia para detectar fraude usando redes neurais artificiais e medidor inteligente. Segundo os autores, a abordagem alcança uma taxa de detecção de fraude de energia mais elevada do que métodos semelhantes nesta área.

Alguns esquemas têm sido propostos para as concessionárias detectarem o roubo de energia mas, todos eles exigem que os usuários enviem suas informações privadas tais como arquivos de carga ou leituras do medidor em determinados momentos, invadindo a privacidade dos usuários e suscitando sérias preocupações sobre segurança. Foi desenvolvido em [7] dois algoritmos de detecção de roubo de energia com base em decomposição LU, chamado LUD e LUPD, respectivamente, que podem identificar usuários fraudulentos, sem invadir a privacidade de qualquer usuário.

Em [8] os sistemas de detecção de energia-roubo AMI são discutidos com base em três categorias: classificação, estimativa de estado, e teoria dos jogos, fazendo comparações entre eles. Em [9] foi desenvolvido um algoritmo que aplica técnicas de classificação e agrupamento de consumidores, bem como o uso concomitante de medidores em transformadores e detectores de anomalias, para detectar mudanças no padrão de uso e proporcionar um desempenho elevado e ajustável na detecção de furtos de energia, com uma baixa taxa de amostragem.

Um exemplo de estudo matemático para a obtenção dos valores de perdas é apresentado em Mogilenko e Pavlyuchenko [10] no qual um modelo de regressão *fuzzy* é utilizado para estimação de perdas. Sistemas de gestão e propostas de soluções integradas também estão presentes na literatura como o trabalho de Vieira et al. [11], abordando esta integração em um sistema de Baixa Tensão.

O Brasil em especial, possui uma legislação abordando alguns aspectos deste tema e uma visão geral da legislação sobre regulação e aspectos relacionados esta apresentada em [12].

Para se detectar e calcular as perdas comerciais é necessário obter o valor da energia fornecida às cargas alimentadas pela concessionária e retirar as perdas ôhmicas nas redes e equipamentos e a energia faturada.

Basicamente, a perda comercial pode existir por um dos seguintes motivos: diferenças no processo de faturamento; furto de energia; ligações clandestinas; ou ligações irregulares.

As diferenças entre as energias lidas e as utilizadas no processo de faturamento são devidas a irregularidades na instalação do cliente. Nesse caso, são feitas cobranças por estimativa e, às vezes, defasadas no tempo. Denominam-se ligações clandestinas àquelas que a companhia não tem registro em seu cadastro de consumidores. As ligações irregulares são aquelas que foram adulteradas de alguma

H. O. Henriques, Universidade Federal Fluminense (UFF), Niterói, Brasil, rrique@vm.uff.br.

H. L. S. Almeida, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brasil, heraldo@ufrj.br.

M. Z. Fortes, Universidade Federal Fluminense (UFF), Niterói, Brasil, mzf@vm.uff.br.

F. B. Carregosa, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brasil, felipecarregosa@gmail.com.

W. L. Moura, Companhia Energética de Pernambuco (CELPE), Recife, Brasil, wladimir@celpe.com.br.

forma, geralmente dentro do quadro de entrada ou no próprio medidor.

Este artigo descreve a metodologia desenvolvida e os principais resultados de projeto para identificação e combate a fraudes. O principal produto desse trabalho é um software que pode ser utilizado em uma concessionária de energia para orientar e planejar inspeções inteligentes. No desenvolvimento deste software foram utilizadas técnicas de inteligência computacional e medições feitas em um percentual de até 30% dentre os transformadores de distribuição de um alimentador em um sistema estudo de caso analisado. O artigo propõe um processo para avaliação e localização de perdas comerciais e procura calcular a verdadeira energia consumida pelos consumidores através das diferenças entre esta e a energia faturada, verificando onde a perda acontece.

Conhecendo os principais motivos que levam ao erro nos processos de previsão das curvas de carga dos clientes, os produtos principais deste artigo são:

- Software para localização de fraudes e ligações clandestinas em clientes de baixa tensão;
- Modelo de representação das incertezas da demanda das cargas utilizando lógica *fuzzy*;
- Modelo de agregação de cargas *fuzzy* para compor um alimentador primário;
- Proposta de correção das incertezas das cargas, baseado em um modelo de estimação linear *fuzzy*;
- Modelo de inspeção supervisionada.

II. FUNDAMENTAÇÃO TEÓRICA

O cálculo de perdas comerciais utiliza a demanda do cliente ao longo do período faturado. Então, em primeiro lugar, deve-se estimar a curva de carga de cada consumidor. No presente trabalho, essa estimativa é feita de duas maneiras: com base na energia faturada e na tipologia de cargas; ou com base na tipologia de cargas construída baseada na Pesquisa Social de Posses e Hábitos de Consumo (PSPH).

O processo ideal para identificação das perdas comerciais é o de instalar medidores comparativos ao longo de todo o sistema. Entretanto, o custo desse procedimento é elevado, pois seria necessária uma quantidade muito grande desses medidores.

Stojkov et al.[13] comentam que a identificação e a proposição de metodologias para a medição de perdas comerciais é bem complexa e dependente da topologia das redes e características dos clientes, apresentando um processo de identificação por análise de ondas refletidas.

A demanda dos clientes estimada pelo método baseado na energia faturada pode conter erros grosseiros devido à possibilidade do consumo estar contaminado pelas irregularidades no consumidor. Nesse método, o cálculo da curva de carga é feito após a obtenção da energia faturada do cliente para um determinado período. Neste instante, calcula-se a demanda média e multiplica-se pelos números em p.u. da curva de carga de sua tipologia.

A demanda estimada através do método da PSPH utiliza

dados declarados de uso de aparelhos elétricos por uma amostra de consumidores tipo. Desse modo, a metodologia fica isenta dos erros de irregularidades no faturamento, mas, em geral, possui um desvio padrão bem elevado.

Outro problema que causa erros em ambos os métodos é a defasagem entre o cadastro comercial e o elétrico das empresas. Essa defasagem, na maioria dos casos, é devida a manobras e demora na atualização de obras, da ligação de novos clientes, desligamento ou cortes, juntamente com as falhas de caracterização da classe do consumidor, tais como residencial, comercial, etc.

A solução encontrada para a representação horária da carga de cada cliente, nesse trabalho, devido a todos esses problemas encontrados na maioria das empresas de energia elétrica, é feita por números *fuzzy* trapezoidais seguindo técnica apresentada em [14], cujas incertezas são traduzidas pelos valores limites da base inferior do trapézio. Esses limites são determinados pela manipulação da tipologia de carga e da pesquisa de campo. Nessa pesquisa, constam perguntas sobre, identificação do cliente, caracterização do domicílio, informações sobre energia elétrica, posses e hábitos de consumo de aparelhos e iluminação, dados socioeconômicos, pesquisa de opinião (perguntas sobre fraudes e inadimplências) e satisfação do consumidor. As respostas a esse questionário são devidamente codificadas e os clientes são separados em clusters de máxima similaridade de dados cadastrais. Cada cluster possui características predominantes de uso de aparelhos elétricos e de topologia de carga.

A demanda trapezoidal resultante é formada por quatro valores. As bases superior e inferior do trapézio são formadas pelo valor esperado da demanda mais ou menos o desvio padrão das curvas geradas pela tipologia de carga e pelos hábitos de uso dos aparelhos elétricos,.

As curvas de carga de todos os clientes são agregadas ao transformador de distribuição que os alimenta. Dessa forma, obtém-se a curva de carga *fuzzy* do transformador.

A demanda lida do medidor de saída do alimentador é modelada como uma medida *fuzzy*, capaz de representar uma perda ôhmica estimada. A diferença horária *fuzzy* entre o valor medido na subestação e o somatório dos valores estimados para cada transformador é a perda comercial *fuzzy*.

Essa perda é alocada entre os transformadores. O processo de alocação sugerido neste estudo é o de estimação linear *fuzzy* como sugerido em [15], onde valores trapezoidais de demanda de transformadores de maior base inferior, que representam maiores incertezas, recebem valores maiores de correção ao contrário dos de base menor.

Os transformadores que tiverem maior correção serão listados e *ranqueados*. Os primeiros da lista são inspecionados com maior detalhe, sendo instalados medidores na saída do transformador e em outros locais, segmentando a baixa tensão. O estimador linear determina quais os ramais de serviço que têm maior correção e cujos consumidores serão minuciosamente inspecionados.

III. METODOLOGIA

A metodologia de prospecção de perdas proposta neste projeto consiste essencialmente das seguintes etapas: obtenção e preparo dos dados; formação dos clusters por rede neural; estimativa inicial das curvas de carga *fuzzy*; correção das curvas de carga por estimação *fuzzy*; processo iterativo de localização de medidores e operacionalização do processo.

Estas etapas são descritas detalhadamente nas subseções seguintes:

A. Obtenção e Preparo dos Dados

Os dados de entrada do processo de estimação das curvas de carga *fuzzy* para clientes e para transformadores de distribuição são: dados cadastrais dos clientes e de conexão à rede elétrica; resultados da pesquisa social de posses e hábitos; tabelas de consumo típico de aparelhos elétricos; e curvas de carga obtidas na campanha de medidas.

A.1 Dados Cadastrais e Preparo dos Dados

Os dados cadastrais e as informações de conexão de clientes à rede elétrica foram obtidos de arquivos extraídos das bases de dados de uma concessionária de energia. Esta concessionária já possui estudo de pesquisa relacionado perdas em sistemas de distribuição monofásicos com retorno terra em áreas rurais apresentado em [16].

Os dados cadastrais dos clientes são utilizados como base para a formação de clusters por critérios de afinidade cadastral, enquanto que as informações a respeito da conexão dos clientes e transformadores de distribuição à rede elétrica são utilizadas para construir curvas de carga agregadas por transformador e para calcular o balanço de energia em alimentadores e transformadores.

As informações extraídas da base de dados e utilizadas nos processos de formação de cluster, estimativa de curvas de carga agregadas e geração de listas de inspeção foram: identificação do cliente (contrato); código da subestação; código do alimentador; código do transformador; coordenadas X e Y; classe (residencial, comercial, industrial, etc.); tipo de tarifa; tensão (monofásica, bifásica ou trifásica); consumo no mês de dezembro (referência maior consumo); consumo médio nos 12 meses do ano; situação da dívida (adimplente ou inadimplente); data da última inspeção (quando aplicável); resultado da última inspeção (quando aplicável) e indicativo de fraude (quando houver)

A.2 Resultado da Pesquisa Social e de posse e hábitos

Neste estudo utilizaram-se os dados da PSPH realizada no ano de 2008 com 932 consumidores residenciais, comerciais e industriais.

Além de informações socioeconômicas e de pesquisa de opinião, foram levantadas informações a respeito da posse de aparelhos elétricos e dos hábitos de consumo dos clientes.

Foram levantados na pesquisa de posses e hábitos os tipos de aparelhos utilizados pelos consumidores, as quantidades de cada tipo de aparelho utilizado e os horários do dia em que esses aparelhos são utilizados. A Tabela I enumera as faixas de horário considerados na pesquisa de campo.

TABELA I. FAIXAS DE HORÁRIO CONSIDERADAS NA PESQUISA DE CAMPO.

HORARIO	HORA INICIAL	HORA FINAL
1	00:00	05:00
2	05:00	08:00
3	08:00	11:00
4	11:00	14:00
5	14:00	18:00
6	18:00	21:00
7	21:00	24:00

A.3 Tabela de Consumo Típicos de Aparelhos Elétricos

A estimativa da curva de carga a partir das informações de posse e hábitos de uso de aparelhos elétricos é feita com base em estimativas padronizadas de consumo e tempo de utilização de diferentes aparelhos.

Para isso foram utilizadas tabelas adaptadas com base em informações obtidas no portal do PROCEL/Eletróbrás [17].

Nestas tabelas, são arbitrados, para cada tipo de aparelho, valores típicos para o consumo de potência, o tempo de utilização diária e o número de dias de uso por mês.

A.4 Curva de Cargas obtidas na Campanha de Medidas

Foram utilizadas no processo de estimação de curvas de carga *fuzzy* as curvas de consumo típicas por classe de cliente e por faixa de consumo, obtidas na Terceira Campanha de Medidas da concessionária em estudo, realizada no ano de 2004.

A estrutura dos dados obtidos da Terceira Campanha de Medidas é descrita na Tabela II.

TABELA II. ESTRUTURA DE DADOS DA TERCEIRA CAMPANHA DE MEDIDAS.

CAMPO	DESCRIÇÃO
1	CLASSE DE CLIENTE
2	FAIXA DE CONSUMO MÉDIO
3	IDENTIFICAÇÃO DO CLIENTE
	MEDIDO
4	DIA DA SEMANA
5	HORÁRIO (0 A 23)
6	CONSUMO MEDIDO NO HORÁRIO

B. Formação dos Clusters por Rede Neural

Inicialmente, as informações cadastrais dos clientes foram codificadas em formato binário padronizado, adequado para o posterior agrupamento dos clientes em clusters, com base em critérios de afinidade de parâmetros cadastrais.

Este formato binário padronizado, que identifica as características cadastrais de cada cliente, será doravante referenciado como o “DNA” do cliente.

Para promover o agrupamento inteligente dos consumidores em clusters que representem grupos com tendências comportamentais semelhantes, foi utilizada uma rede neural artificial competitiva do tipo *Self-Organizing Map* (SOM), ou mapa auto-organizável, também conhecida como Mapa de Kohonen, apresentado em [18].

Este tipo de rede neural artificial é capaz de aprender a detectar regularidades e correlações nos dados de entrada, adaptando suas futuras respostas de modo que neurônios geograficamente próximos no espaço multidimensional dos dados de entrada tendam a responder de forma similar.

O processo de formação dos clusters envolve as etapas de: definição dos critérios de agrupamento; determinação do “DNA” dos clientes; treinamento das redes neurais e classificação dos clientes nos clusters.

B.1 Critérios de Agrupamento

Os clientes de duas subestações exemplos, nomeadas neste texto de (PAM) e (JBR), foram agrupados por semelhança de perfil, de acordo com os critérios: classe de cliente (residencial, comercial, industrial); tipo de ligação (monofásica, bifásica, trifásica); consumo médio mensal faturado (em kWh); subestação (PAM, JBR); localização geográfica (coordenadas UTM X e Y); tipo de tarifa (B1 Baixa Renda, B1 Residencial, B3); situação comercial (adimplente, inadimplente); indicativo de fraude (com/sem indicativo) e resultado da inspeção (tipo de irregularidade).

Estas variáveis, utilizadas como base para distinguir diferentes perfis de cliente, serão doravante denominadas *variáveis descritivas do cliente*.

B.2 Determinação do “DNA” do cliente

Para codificar as variáveis descritivas de cada cliente em um formato adequado para *clusterização* por meio de rede neural, foi utilizada uma codificação binária, na qual cada “gen” pode assumir exclusivamente os valores 0 (falso) ou 1 (verdadeiro).

Para cada critério de agrupamento, é definido um conjunto de “gens”, dentre os quais um único “gen” estará ativo (com valor 1) e os demais estarão inativos (valor 0). A especificação do conjunto de “gens” associado a cada variável descritiva depende do tipo de variável descritiva.

No modelo existem dois tipos básicos de variáveis descritivas (discretas e contínuas).

Variáveis discretas são aquelas que assumem um dentre um conjunto limitado de valores possíveis, como por exemplo a classe de cliente (residencial, comercial, industrial, etc.) e o tipo de ligação (monofásica, bifásica, trifásica).

Variáveis contínuas são aquelas que assumem valores numéricos contínuos, como por exemplo o consumo médio mensal e as coordenadas geográficas X e Y do cliente.

No caso de variáveis discretas, a codificação é feita da seguinte forma:

- Aloca-se um “gen” para cada possível valor de cada variável descritiva.
- No “DNA” de cada cliente, o “gen” associado ao valor que a variável descritiva assume para aquele cliente recebe o valor 1 (verdadeiro), enquanto que os gens associados aos demais valores possíveis daquela variável descritiva recebem o valor 0 (falso).

Já no caso de variáveis contínuas, a codificação é feita da seguinte forma:

- Os valores assumidos pela variável descritiva para cada cliente são ordenados em ordem crescente e segmentados em 5 faixas de valores, de modo que haja quantidade aproximadamente igual de ocorrências (20% do total de clientes) em cada faixa (isto é feito na etapa de segmentação das variáveis contínuas descrita anteriormente).

- Aloca-se um “gen” para cada faixa de cada variável descritiva.

- No “DNA” de cada cliente, o “gen” associado à faixa que a variável descritiva assume para aquele cliente recebe o valor 1 (verdadeiro), enquanto que os “gens” associados às demais faixas recebem o valor 0 (falso).

Desta forma, as 10 variáveis descritivas foram codificadas em 36 gens, como ilustrado na Tabela III.

B.3 Treinamento das Redes Neurais

Duas redes neurais do tipo *Mapa de Kohonen* foram treinadas com os dados cadastrais dos clientes que responderam à pesquisa social de posses e hábitos, tendo sido uma delas treinada com os dados dos clientes residenciais e outra com os dados dos clientes comerciais e industriais.

O conjunto de dados utilizado para treinamento da rede neural foi composto pelo “DNA” dos 932 clientes que responderam à pesquisa.

Foram executadas 10.000 iterações de treinamento, ou seja, o conjunto de dados dos 932 clientes foi apresentado à rede 10.000 vezes, em diferente ordem a cada apresentação.

Foram utilizadas redes neurais do tipo *Mapa de Kohonen* com 42 neurônios, iniciadas com uma estrutura em grade, composta por 6 linhas x 7 colunas, com estado inicial uniformemente distribuído ao longo do espaço de possibilidades.

O número de neurônios foi dimensionado de forma a ser obtida uma média da ordem de 25 clientes por cluster no conjunto de treinamento e pelo menos 1 neurônio independente para cada dimensão do espaço de possibilidades.

B.4 Classificação dos Clientes nos Clusters

Uma vez treinadas as redes neurais e formados os clusters, os demais clientes (aqueles que não foram entrevistados e ficaram de fora do conjunto selecionado aleatoriamente para treinamento da rede) foram classificados nos clusters por meio da simulação das redes neurais.

No processo de simulação da rede neural, cada novo cliente apresentado à rede ativa cada neurônio com maior ou menor intensidade, dependendo de seu “DNA”. O cliente é classificado no cluster correspondente ao neurônio que for ativado com maior intensidade (aquele que “vence a competição”).

TABELA III. SIGNIFICADO DOS “GENS” NO “DNA” DO CLIENTE.

GEN	DESCRIÇÃO
1	CLASSE DO CLIENTE – RESIDENCIAL
2	CLASSE DO CLIENTE – COMERCIAL
3	CLASSE DO CLIENTE – INDUSTRIAL
4	TIPO DE LIGAÇÃO – MONOFÁSICO
5	TIPO DE LIGAÇÃO – BIFÁSICO
6	TIPO DE LIGAÇÃO – TRIFÁSICO
7	CONSUMOMÉDIO MENSAL – FAIXA 1
8	CONSUMOMÉDIO MENSAL – FAIXA 2
9	CONSUMOMÉDIO MENSAL – FAIXA 3
10	CONSUMOMÉDIO MENSAL – FAIXA 4
11	CONSUMOMÉDIO MENSAL – FAIXA 5
12	SUBESTAÇÃO PAM
13	SUBESTAÇÃO JBR
14	COORDENADA UTM X – FAIXA 1
15	COORDENADA UTM X – FAIXA 2
16	COORDENADA UTM X – FAIXA 3
17	COORDENADA UTM X – FAIXA 4
18	COORDENADA UTM X – FAIXA 5
19	COORDENADA UTM Y – FAIXA 1
20	COORDENADA UTM Y – FAIXA 2
21	COORDENADA UTM Y – FAIXA 3
22	COORDENADA UTM Y – FAIXA 4
23	COORDENADA UTM Y – FAIXA 5
24	TIPO DE TARIFA – B1 BAIXA RENDA
25	TIPO DE TARIFA – B1 RESIDENCIAL
26	TIPO DE TARIFA – B3
27	SITUAÇÃO COMERCIAL - ADIMPLENTE
28	SITUAÇÃO COMERCIAL - INADIMPLENTE
29	COM INDICATIVO DE FRAUDE
30	SEM INDICATIVO DE FRAUDE
31	INDICATIVO DE FRAUDE – NÃO INSPECIONADO
32	RESULTADO DE INSPEÇÃO – IRREGULARIDADE CODIGO 100 A 199
33	RESULTADO DE INSPEÇÃO – IRREGULARIDADE CODIGO 200 A 299
34	RESULTADO DE INSPEÇÃO – IRREGULARIDADE CODIGO 300 A 399
35	RESULTADO DE INSPEÇÃO – SEM IRREGULARIDADE (CODIGO 400)
36	RESULTADO DE INSPEÇÃO – NÃO INSPECIONADO

C Estimativa Inicial das Curvas de Carga Fuzzy

O processo de cálculo das curvas de carga tem por objetivo produzir uma estimativa a mais realista possível da variação hora a hora da demanda de carga em cada transformador de distribuição, em dias úteis, sábados, domingos e feriados, visando à localização de fraudes e ligações irregulares.

Esta estimativa se baseia em informações cadastrais, informações coletadas em pesquisa de campo e em informações obtidas de campanha de medidas.

O cálculo da estimativa inicial das curvas de carga para os transformadores de distribuição envolve as etapas: cálculo das curvas presumidas para cada cluster; cálculo das curvas presumidas por tipologia; cálculo das curvas de carga estimadas para clientes e agregação das curvas de carga por transformador.

C.1 Cálculo das curvas presumidas PSPH

Para cada cluster, é construída uma curva de carga padrão presumida, que reflete as informações fornecidas por membros do cluster que responderam à pesquisa social de posses e hábitos.

As curvas presumidas para cada cluster são construídas

com base na informação coletada na pesquisa de campo com relação à posse e aos hábitos de uso de equipamentos elétricos, por meio do seguinte processo:

- Para cada cliente entrevistado, é construída uma curva estimada de consumo hora a hora, de acordo com os aparelhos elétricos informados na pesquisa de campo e seus respectivos horários de utilização.
- As curvas de carga dos clientes entrevistados são normalizadas em p.u., de modo que o valor médio da curva estimada corresponda a 1 unidade.
- Para cada cluster, é calculada uma curva de carga correspondente à média das curvas normalizadas de cada cliente entrevistado. Usando como exemplo o cluster 29 da classe residencial, a linha azul da Fig. 1 representa esta média. Também são calculadas as curvas correspondentes aos limites inferior e superior, linhas verde e vermelha respectivamente da Fig. 1, estimadas a partir do desvio padrão calculado hora a hora.

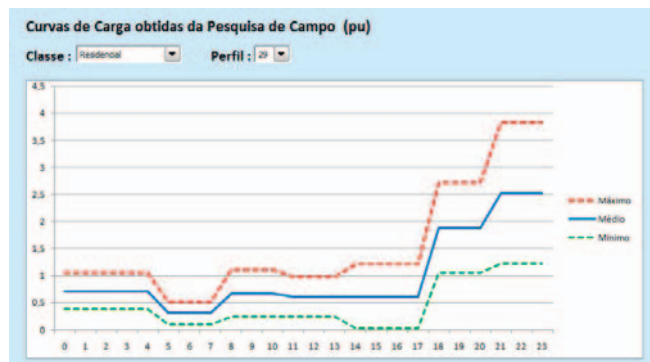


Figura 1. Curva de carga presumida para um cluster de clientes residenciais.

As curvas de carga presumidas para cada cluster são utilizadas para estimar a curva de carga a ser atribuída a cada cliente que seja classificado naqueles mesmos clusters.

C2. Cálculo das curvas presumidas por tipologia

As curvas presumidas por tipologia de cliente foram construídas com base nas informações coletadas na Terceira Campanha de Medidas, por meio do seguinte processo:

- Foram utilizadas 15 tipologias de cliente, correspondentes ao cruzamento de 3 classes de cliente (residencial, comercial e industrial) com 5 faixas de consumo (0 a 100 kWh, 101 a 300 kWh, 301 a 500 kWh, 501 a 1.000 kWh e acima de 1.000 kWh). Os dados coletados da Terceira Campanha de Medidas compreenderam as medições de uma amostra de 10 a 15 clientes de cada tipologia, efetuadas em dias úteis, sábados e domingos.
- Para cada tipologia de cliente, foram calculadas 9 curvas de carga padrão, cada uma com 24 medidas (1 medição por hora), correspondentes ao valor médio e aos limites inferior e superior estimados para sábados, domingos e dias úteis. As curvas correspondentes aos limites inferior e superior são estimadas com base no desvio padrão calculado hora a hora para a amostra de cada tipologia, para cada tipo de dia da semana (sábados,

domingos e dias úteis).

A título de ilustração, os valores mínimos, médios e máximos de uma curva de carga para clientes industriais com consumo médio mensal acima de 1.000 kWh em dias úteis são exibidas na Fig. 2.

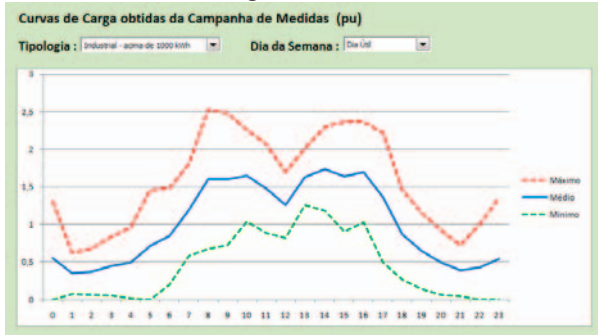


Figura 2. Curva de carga presumida para uma tipologia de cliente.

C3. Cálculo das curvas estimadas para cada cliente

As curvas de carga estimadas para cada cliente são calculadas a partir de uma composição da curva de carga presumida para o cluster em que o cliente foi classificado e da curva de carga presumida para a tipologia em que o cliente se enquadra.

Estas curvas de carga são construídas por meio do seguinte processo:

- O cluster a que cada cliente pertence é determinado com base no processo descrito na subseção B (formação dos clusters).
- A tipologia de cada cliente cadastrado é determinada com base na classe do cliente e no seu consumo médio dos últimos 12 meses.
- Uma curva de carga *fuzzy* é estimada para cada cliente. O valor da curva em cada hora é um valor *fuzzy* representado pelos quatro parâmetros a , b , c e d que delimitam o trapézio ilustrado na Fig. 3, correspondente ao grau de pertinência do consumo do cliente às faixas de valores em kWh naquele horário.

C4. Agregação das Curvas de Carga por Transformador

O cálculo das curvas de carga *fuzzy* agregadas por transformador de distribuição foi efetuado por meio do seguinte processo:

- Para cada cliente das subestações (JBR) e (PAM), foram determinados o alimentador e o transformador de distribuição ao qual o cliente está conectado.
- Para cada transformador de distribuição, as curvas de carga dos clientes conectados a esse transformador foram convertidas em kW, com base no consumo médio mensal de cada cliente, e somadas por agregação *fuzzy*, produzindo como resultado a curva de carga *fuzzy* agregada correspondente ao transformador.

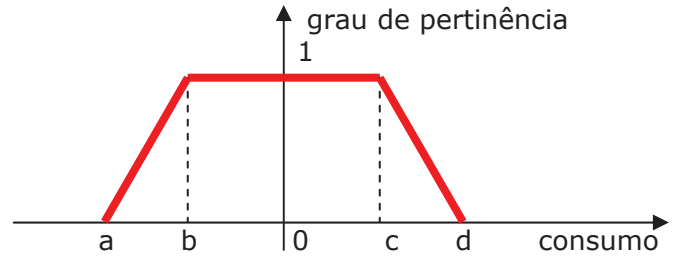


Figura 3. Representação trapezoidal dos valores *fuzzy* de consumo.

A título de ilustração, a curva de carga *fuzzy* prevista pelo software para o transformador T25365 do alimentador 01C3 da subestação PAM em um dia útil é exibida na Fig. 4.

As curvas a , b , c e d na Fig. 5 representam os parâmetros da representação trapezoidal dos valores *fuzzy* do consumo hora a hora.

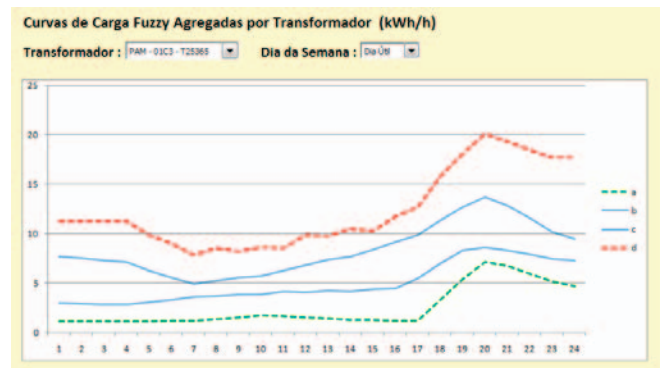


Figura 4. Curva de carga presumida para uma tipologia de cliente.

D. Correção das Curvas de Carga por Estimação Fuzzy

A demanda lida do medidor de saída do alimentador é modelada como uma medida *fuzzy*, capaz de representar uma perda ôhmica estimada. A diferença horária *fuzzy* entre o valor medido na subestação e o somatório dos valores estimados para cada transformador é a perda comercial *fuzzy*.

Essa perda precisa ser alocada entre os transformadores. O problema, ilustrado pela Fig. 5, consiste em determinar a demanda de cada transformador de distribuição alimentado por um circuito primário, baseando-se nas estimativas iniciais P_i^c , $i=1,2,\dots,n$, em função do consumo e dos limites inferior e superior C_i^m e C_i^M , $i=1,2,\dots,n$, determinados por curvas típicas e em uma leitura na saída do circuito.

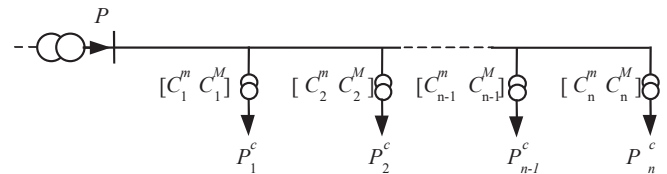
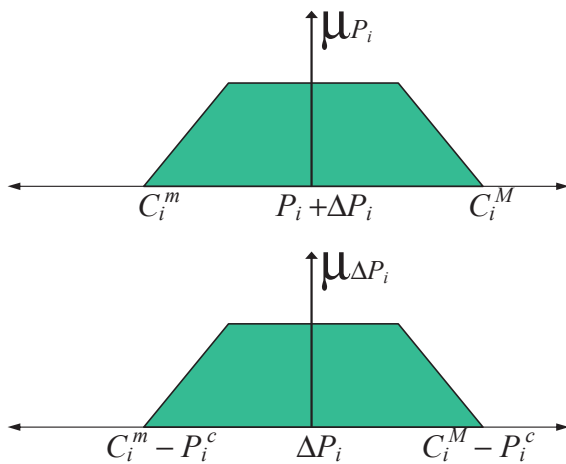


Figura 5. Circuito radial típico considerado na alocação das perdas.

O modelo calcula as correções ΔP_i , $i=1,2,\dots,n$, a serem aplicadas aos valores estimados P_i^c , $i=1,2,\dots,n$, de modo que a diferença ΔP entre o valor medido P e a soma das estimativas corrigidas $\sum_{i=1,2,\dots,n} (P_i^c + \Delta P_i)$ seja a menor possível e as correções ΔP_i , $i=1,2,\dots,n$, estejam dentro da região de pertinência ilustrada na Fig. 6.

Figura 6. Curvas de pertinência para P_i e ΔP_i .

As correções ΔP_i , $i=1,2,\dots,n$, são obtidas resolvendo-se o problema de otimização descrito nas equações (1) e (2).

$$\text{minimizar } \sum_{i=1,2,\dots,n} w_i \cdot \Delta P_i \quad (1)$$

$$\text{sujeito a } C_i^m - P_i^c \leq \Delta P_i \leq C_i^M - P_i^c, \quad i=1,2,\dots,n \quad (2)$$

onde: ΔP_i = correção a ser aplicada ao transformador i
(variável de decisão a ser otimizada)

w_i = ponderação aplicada ao transformador i

P_i^c = valor esperado para o transformador i

C_i^m = limite inferior para o transformador i

C_i^M = limite superior para o transformador i

O problema descrito nas equações (1) e (2) pode ser formulado como um problema de programação linear e solucionado pelo algoritmo Simplex.

E. Processo Iterativo de Localização de Medidores

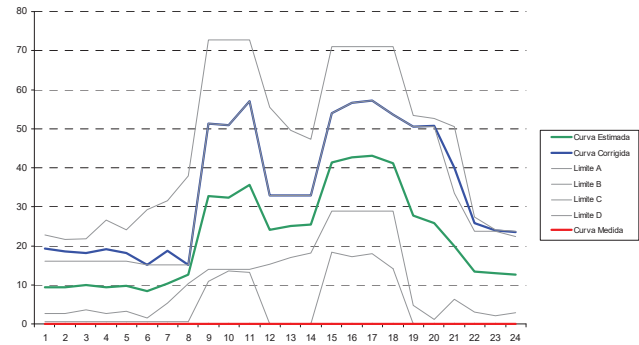
Após o cálculo das curvas de carga *fuzzy* dos consumidores e transformadores, o processo de rastreamento de perdas utiliza um numero reduzido de medidores para melhorar a estimação das correções de cada transformador de distribuição. O processo de escolha do local que deverá ter instalado um medidor de demanda consiste em resolver o sistema linear *fuzzy*, pelo método simplex, sem nenhum medidor além do instalado na saída do alimentador. O passo seguinte é analisar o conjunto de curvas de carga *fuzzy* e a corrigida, resultado do simplex e adota-se o seguinte critério:

- colocar medidor onde os valores corrigidos oscilam em relação ao valor estimado;
- colocar medidor onde o valor corrigido apresenta forma diferente do esperado.

O critério adotado é explicado pela premissa que se a ligação de clientes ao transformador estiver atualizada, se o alimentador não estiver manobrado e demais erros possíveis citados na seção 1 não estiverem presentes, o resultado esperado, uma vez que existe uma diferença positiva entre o valor de demanda lido à saída do alimentador e o somatório

das demandas de todos os transformadores do alimentador, é o da adição ao valor previsto inicialmente.

A Fig. 7 ilustra a curva de carga de um transformador cujo valor corrigido pelo simplex, em azul, está coerente com a previsão inicial, em verde. Ou seja, a adição ao valor estimado, linha verde, não ultrapassa o limite superior A e mantém certa similaridade com a curva média. Esse transformador não precisa ter medidor.

Figura 7. Curvas de carga *fuzzy* estimada e corrigida, no caso de um transformador que não precisa ter medidor.

A Fig. 8 mostra, num gráfico de barras, que a diferença entre os valores corrigidos e estimados, linha azul menos linha verde, se comporta de maneira uniforme, ou seja, em todos os horários a correção é positiva.

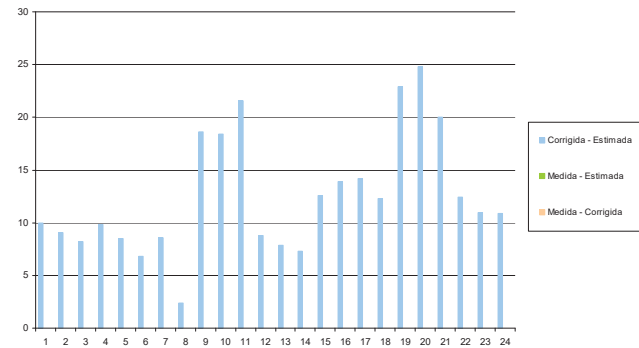


Figura 8. Diferença entre valor corrigido e estimado em um transformador sem necessidade de medidor.

A Fig. 9 mostra um exemplo de um transformador em que o valor corrigido oscila em relação ao estimado, isto é, hora a correção é positiva e hora negativa. Além disto, como a diferença entre a medida na subestação e o somatório das demandas dos transformadores é positiva, espera-se que a demanda corrigida seja maior que a estimada. Esse transformador é candidato a ter um medidor instalado.

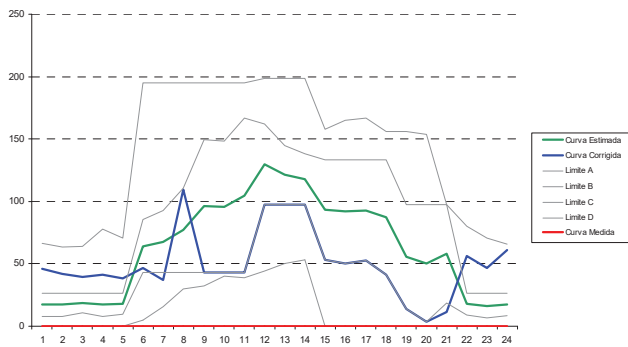


Figura 9. Curvas de carga *fuzzy* estimada e corrigida, no caso de um transformador que precisa ter medidor.

A Fig. 10 ilustra o gráfico de barras que a diferença entre os valores corrigidos e estimados, linha azul menos linha verde, apresenta valores negativos entre 6 e 21 horas.

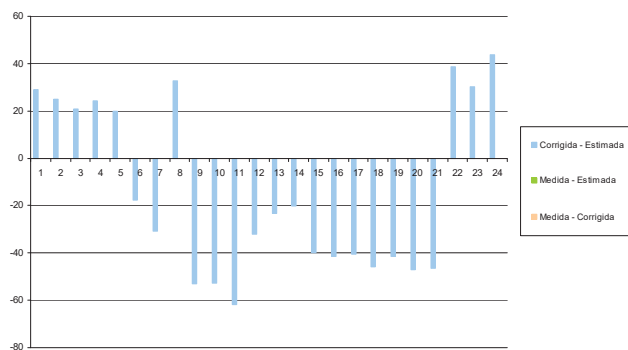


Figura 10. Diferença entre valor corrigido e estimado em um transformador com necessidade de medidor.

A Fig. 11 mostra um caso em que, embora as diferenças sejam todas positivas, como mostra a Fig. 12, a curva corrigida pelo método simplex apresenta uma forma distorcida da estimada. A linha azul tem uma descontinuidade anômala no horário entre 12 e 19 horas. Esse detalhe está marcado na Fig. 11 por um círculo vermelho e traduz, devido à faixa horária, a possibilidade de presença de erros na caracterização dos consumidores comerciais e/ou industriais ligados a esse transformador. Este transformador também é selecionado para ter um medidor instalado no lado de baixa tensão.

F. Operacionalização do Processo

O processo é implantado em duas etapas. A primeira consiste no tratamento de dados e solução do sistema linear *fuzzy* para áreas que possuem medidor. A segunda consiste na solução do sistema linear *fuzzy*, considerando as correções efetuadas na primeira etapa, para os demais transformadores sem medidores instalados.

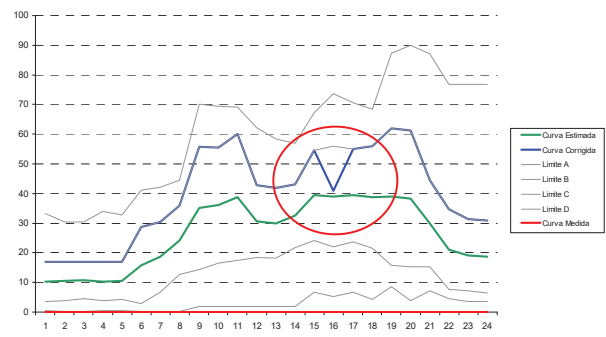


Figura 11. Resultado para transformador com forma diferente do esperado.

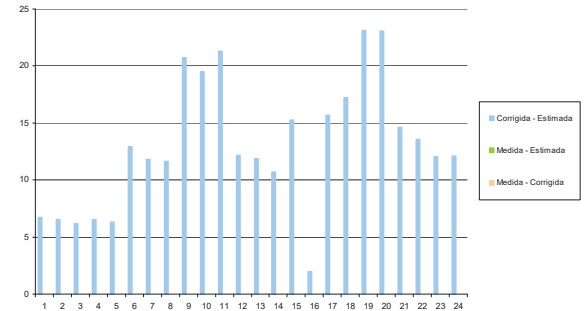


Figura 12. Diferença entre valores corrigidos e estimados das curvas da Fig. 11.

F.1. Primeira Etapa

A tarefa de tratamento de dados consiste em estimar inicialmente as curvas de carga dos clientes e dos transformadores, em estimar uma correção inicial pelo método simplex considerando apenas o medidor da subestação, em localizar e instalar os medidores nos transformadores e por fim em executar novamente o simplex, porém desta vez considerando os valores coletados dos medidores instalados nos transformadores.

Para os transformadores com medidor, calcular valores “corrigidos – estimados” e “medidos – estimados”. Verificar em campo a vinculação correta de consumidores para o caso em que “corrigidos – estimados” e “medidos – estimados” tiverem sinais contrários ou descontinuidades anômalas.

A Fig. 13 mostra a medição, em vermelho, coletada para o transformador da Fig. 9.

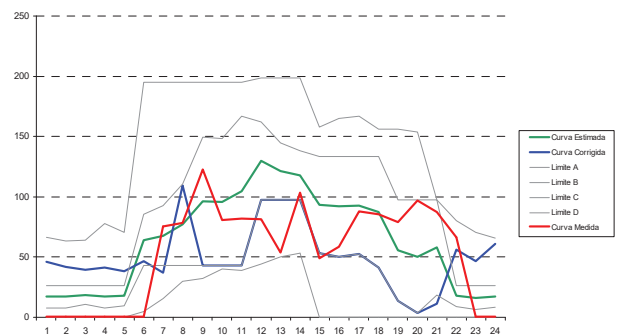


Figura 13. Gráfico da medição do transformador da Fig. 9.

As diferenças “corrigidos – estimados” e “medidos – estimados” estão representadas no gráfico de barras da Fig. 14.

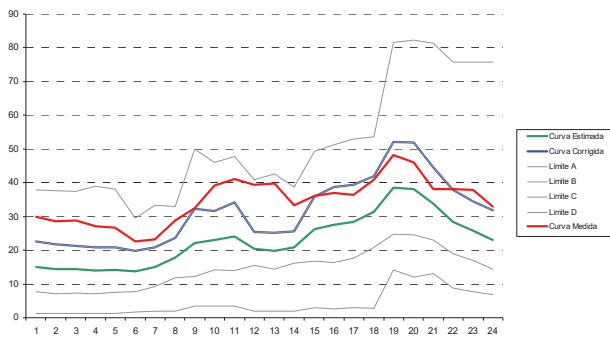


Figura 17. Transformador T – 24697 de PAM com pequenos erros de irregularidades no cadastro.

A tabela IV mostra que quando a rede está pouco desatualizada quanto à topologia de rede e vinculação do consumidor ao transformador, o estimador *fuzzy* melhora a estimativa do valor real, representado pela medição.

TABELA IV. PRECISÃO DO SOFTWARE PARA TRANSFORMADORES COM POUCOS PROBLEMAS DE ATUALIZAÇÃO DE DADOS.

Transformador	Consumo Estimado por Faturamento (kWh/dia)	Consumo Corrigido por Estimação Fuzzy (kWh/dia)	Consumo por Medição (kWh/dia)
C12213	1482,80	1291,83	1202,33
C42595	52,16	82,88	81,21
T21633	349,8	419,76	358,28
T23340	214,78	341,27	351,01
T24688	415,37	532,50	843,30
T24690	620,90	793,74	1290,03
T24694	667,25	869,85	741,69
T24697	550,82	755,00	833,02
T25250	623,46	748,15	736,75
T25377	278,67	409,40	740,15
T25383	153,39	184,42	171,72

V. RESULTADO DAS INSPEÇÕES

Foram indicados 170 clientes para inspeção, 15 eram da classe comercial e 155 da Residencial, dentre estes 123 da classe B1 e 32 de baixa renda.

Como resultados das inspeções, foram verificadas as seguintes situações constantes da Tabela V.

As Figs. 18, 19 e 20 ilustram os resultados para clientes comerciais, das inspeções da classe residencial B1 e o resultado das inspeções da classe residencial Baixa Renda, respectivamente.

As inspeções nos consumidores comerciais, sem a etapa de correção de vinculação, comprovaram que havia 60 % de clientes em condição irregular. 13 % apontaram vestígios de adulteração de medidor(clientes roubando energia) e 47% de desatualização de cadastro. Este resultado reforça a importância da pré-análise do estado do cadastro elétrico em relação ao comercial.

TABELA V. LEGENDA DOS RESULTADOS DA INSPEÇÃO.

LEGENDA	DESCRIÇÃO
NORMAL	UNIDADE CONSUMIDORA SEM FRAUDE
NÃO LOCALIZADO	UNIDADE CONSUMIDORA NÃO LOCALIZADA
FECHADO	3ª VISITA UNIDADE CONSUMIDORA FECHADA
SELO AUSENTE	EQUIPAMENTO DE MEDIÇÃO SEM SELO DE LABORATÓRIO HAVENDO ACESSO A PARTE INTERNA
SELO DESPAD.	O EQUIPAMENTO DE MEDIÇÃO ENCONTRA-SE COM OS SELOS DESPADRONIZADOS
SL. PRENSADOS	O EQUIPAMENTO DE MEDIÇÃO ENCONTRA-SE COM OS SELOS PRENSADOS
P BLO TERM	O EQUIPAMENTO DE MEDIÇÃO ENCONTRA-SE COM UMA PONTE DE BLOCO DE TERMINAIS
SELO ROMPIDO	O EQUIPAMENTO DE MEDIÇÃO ENCONTRA-SE COM OS SELOS DE LABORATÓRIO ROMPIDOS
DISP APAGADO	MEDIDOR COM DISPLAY APAGADO
C CORTADA	UNIDADE CONSUMIDORA CONTINUA CORTADA
IMPEDIDO	CLIENTE IMPEDIU A INSPEÇÃO
DESOCUPADO	IMÓVEL DESOCUPADO
E M IRREGULAR	EQUIPAMENTO DE MEDIÇÃO COM BOBINA IRREGULAR
LIGADO DIRETO	A UNIDADE CONSUMIDORA ENCONTRA-SE LIGADO DIRETO SEM PASSAR PELA MEDIÇÃO
DERI CLANDEST	UNIDADE CONSUMIDORA COM DERIVAÇÃO CLANDESTINA ANTES DA MEDIÇÃO
E M FURADO	EQUIPAMENTO DE MEDIÇÃO ENCONTRA-SE COM SUA TAMPA FURADA HAVENDO ACESSO A SUA PARTE INTERNA

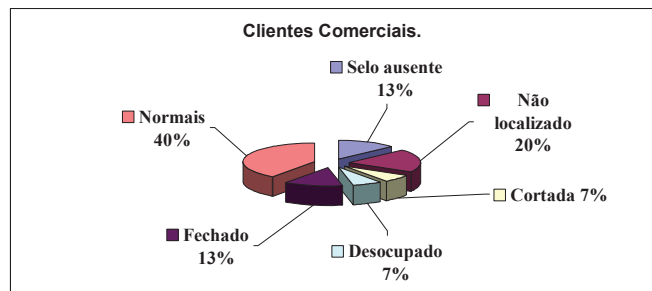


Figura 18 Ocorrências verificadas nas inspeções de clientes comerciais.

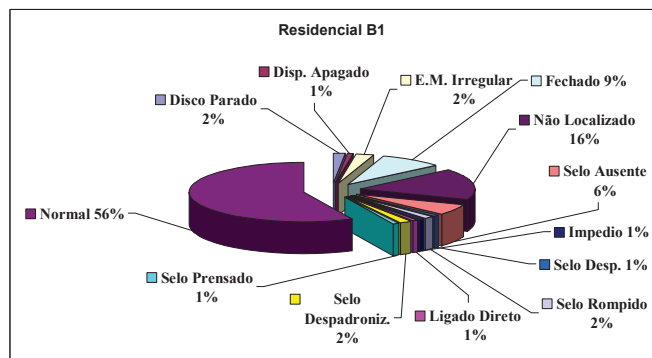


Figura 19. Ocorrências verificadas em clientes residenciais B1.

No caso de clientes residenciais da classe B1, conforme ilustrado na Fig. 19, as inspeções apontaram 44% de irregularidades, sendo 19% relacionados com furto de energia e 25% relacionados à desatualização cadastral.

No caso de clientes de baixa renda, conforme ilustrado na Fig. 20, as inspeções indicaram 43% de irregularidades, sendo que 16% relativas à desatualização de cadastro e 27% relativas a furto de energia.

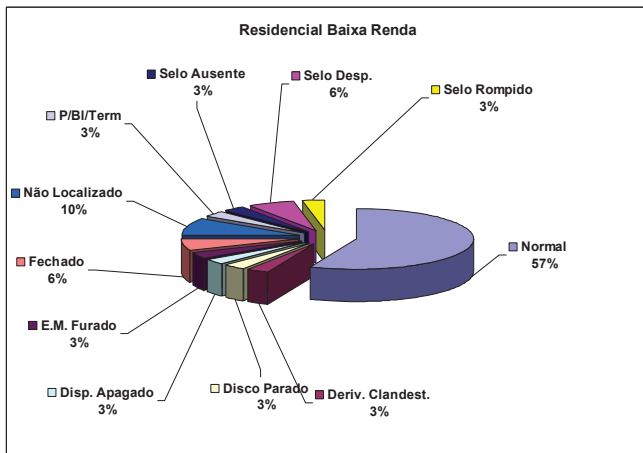


Figura 20. Ocorrências verificadas em clientes residenciais baixa renda.

O resultado apresentado pode ser considerado muito bom, uma vez que não houve atualização de vinculação de cliente a transformador e nem a aplicação por ramal de serviço, pois essas etapas foram desenvolvidas a partir do resultado do teste de campo. Foram verificados 33 casos de irregularidades ligadas a furto de energia envolvendo 30 casos de problemas ligados ao medidor, 2 ligações clandestinas e 1 cliente que impediu a inspeção.

Após a introdução do software desenvolvido na empresa de energia, tornou-se possível planejar inspeções mais direcionadas e pontuais, o que antes era feito de forma massiva, por bairro, rua ou área alimentada por determinada subestação. A assertividade de irregularidades encontradas por número de consumidores inspecionados também apresentou melhores resultados.

VI. CONCLUSÕES E COMENTÁRIOS

Os testes de campo serviram para organizar a operacionalização do processo, com a ajuda do software desenvolvido, nos demais alimentadores. O método proposto avalia quais os transformadores precisam ser verificados, reduzindo-se o custo de verificar todos. Uma boa assertividade depende também da atualização da topologia da rede, ou seja, devem estar atualizados quais os transformadores que estão ligados ao alimentador em estudo e quais os clientes que estão ligados a cada transformador. A primeira etapa do processo proposto aponta os seguimentos de rede que necessitam de atualização.

No caso da concessionária em estudo, as quase totalidades de furtos de energia foram feitas através da adulteração dos medidores.

Espera-se que com o incremento do sistema de detecção de fraudes ocorra uma redução considerável de furtos de energia na concessionária. Deve-se considerar que quanto maior o número de linhas monitoradas, reduz-se o custo de combate a ações de fraudes devido à sistematização dos processos e treinamento das equipes de análise.

REFERÊNCIAS

- [1] D. Bassaikhan, D. Sodnomdori, Kh Enklargal, M. Narantuya and L. Dolgorsuren, "Commercial Losses of UB Electrical Distribution Network Shareholding Company (CO.,SH) and their Reducing", In: *Proceedings of 7th Korea-Russia International Symposium*, p.38-44, 2003.
- [2] S.S. Mustafa, M.H. Yaseen, H.H. Abdullah and H.K. Hazaa, "Evaluation of Electric Energy Losses in Kirkuk Distribution Electric System Area", *Iraq Journal Electrical and Electronic Engineering*, Vol.7, n.2, p.144-150, 2011.
- [3] R.K. Subramanian, G. Wacker and R. Billinton, "Understanding Commercial Losses Resulting from Electric Service Interruptions", *IEEE Transactions on Industry Applications*, Vol.2, p.1711-1715, 1990. doi: 10.1109/IAS.1990.152416.
- [4] H.O.Henriques, A.P.L. Barbero, R.M.Ribeiro, M.Z.Fortes, W. Zanco, O.S.Xavier and R.M.Amorim, "Development of adapted ammeter for fraud detection in low-voltage installations", *Measurement*, Vol.56, p.1-7, 2014. doi: 10.1016/j.measurement.2014.06.015.
- [5] H.O.Henriques, M.Z.Fortes, L.Hudson, N.S.V.Silva and F.O.Teixeira, "Use of Radar for illegal connections prospecting in buried or embedded cables", *Measurement*, Vol.47, p.221-227, 2014. doi: 10.1016/j.measurement.2013.08.061.
- [6] V.Ford, A.Siraj and W.Eberle, "Smart grid energy fraud detection using artificial neural networks", In: *IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*, pp.1-6, 2014. doi: 10.1109/CIASG.2014.7011557.
- [7] S.Salinas, L.Ming and L.Pan, "Privacy-preserving energy theft detection in smart grids", *IEEE Journal of Selected Areas in Communications*, Vol.31, No.9, pp.257-267, 2013. doi: 10.1109/JSAC.2013.SUP.0513023.
- [8] R. Jiang, R.Lu, Y. Wang, J. Luo, C. Shen and X. Shen, "Energy-Theft Detection Issues for Advanced Metering Infrastructure in Smart Grid", *Tsinghua Science and Technology*, Vol.19, No.2, pp.105-120, 2014.
- [9] P.Jokar, N.Arianpoo and V.C.M.Leung, "Electricity Theft Detection in AMI Using Customers' Consumption Patterns", *IEEE Transactions on Smart Grid*, Vol. PP, No.99, pp.1-10. doi: 10.1109/TSG.2015.2425222.
- [10] A.V.Mogilenko and D.A.;Pavlyuchenko, "The Estimation of Electric Power Losses in Electrical Networks by Fuzzy Regression Model using Genetic Algorithm", *Mathware & Soft Computing*, Vol.11, p. 13-30, 2004.
- [11] G. Vieira, R.Barbosa, L. Varela and G. Putnik, "Electrical Energy Losses Determination in Low Voltage – A case study", *Revista Eletrônica Sistema & Gestão*, Vol.6, p.91-116, 2011.
- [12] A.A.Penteado Jr., E.L. Ramalho and M.T.O. Andrade, "Regulation of the Commercial Loss Process in the Brazilian Electrical Sector", In: *19th International Conference on Electricity Distribution*, Paper no 0269, Session 6, 2007.
- [13] M. Stojkov, D. Zagar and K. Trupinic, "Measurement Procedures for Commercial Loss Reduction in a Distribution Power System", *Strojarsvo: Journal for Theory and Application in Mechanical Engineering*, Vol.51, no.4, p.371-389, 2009.
- [14] H.J. Zimmerman, *Fuzzy set Theory and its Applications*, Kluwer Boston. 3rd Edition, 1996..
- [15] H.O.Henriques, "Aplicação de Sistemas Inteligentes na Previsão e Localização de Perdas em Sistemas de Distribuição", Tese de Doutorado, Universidade Federal do Rio de Janeiro, 2003.
- [16] C.M.P. dos Santos, "Determination of Electric Power Losses in Distribution Systems", In: *IEEE /PES Transmission & Distribution Conference and Exposition: Latin America*, p.1-5, 2006. doi: 10.1109/TDCLA.2006.311585.
- [17] ELETROBRAS. Documentos diversos. <http://www.eletrobras.com.br/elb/procel/main.asp>. Acesso em: 15 dez 2014
- [18] T. Kohonen, *Self-Organization and Associate Memory*. Springer-Verlag, Berlin, 1987.



Henrique de Oliveira Henriques, é graduado em Engenharia Elétrica pela Universidade Federal do Rio de Janeiro (1978), mestrado em Computação Aplicada e Automação pela Universidade Federal Fluminense (1998) e doutorado em Engenharia Elétrica pela Universidade Federal do Rio de Janeiro (2003). Atualmente é professor associado da Universidade Federal Fluminense, coordenador do LETD – Laboratório de Estudos de Transmissão e Distribuição de Energia Elétrica.



Heraldo Luis Silveira de Almeida, é graduado em Engenharia Eletrônica pela Universidade Federal do Rio de Janeiro (1988). Possui Mestrado e Doutorado em Engenharia Elétrica pela COPPE/UFRJ (1991/2003), respectivamente. Atualmente é Professor Associado no Departamento de Engenharia Eletrônica e de Computação da Escola Politécnica da UFRJ.



Márcio Zamboti Fortes recebeu o título de Mestre em Engenharia da energia pela Universidade Federal de Itajuba (UNIFEI) em 2000 e de Doutor em Engenharia Elétrica pela Universidade de São Paulo (USP) em 2007. Atualmente é professor da Universidade Federal Fluminense (UFF) .



Felipe Borda Carregosa graduou-se em Engenharia Eletrônica e Computação pela Universidade Federal do Rio de Janeiro (UFRJ) em 2014. Possui interesse nos temas: combate a fraudes e inteligência artificial.



Wlademir Lacerda de Moura graduou-se em Engenharia Elétrica pela Universidade de Pernambuco (UPE) em 2001. Possui Mestrado em Engenharia Elétrica pela Universidade Federal de Pernambuco (UFPE) finalizado em 2010. Atualmente é Gerente do Departamento de Gestão de Mercado - CGM da Companhia Energética de Pernambuco - CELPE.