

# Prospecção de Fraudes e Identificação de Ações de Combate a Perdas Comerciais

Henrique de O. Henriques, Heraldo L. S. de Almeida, Felipe B. Carregosa, Otávio M. Junior, José A. Lopes, Fábio B. Ribeiro e Wladimir L. de Moura

**Resumo** – Os processos de prospecção das perdas comerciais tradicionais não têm sido efetivos. O presente projeto baseia-se numa metodologia de inspeção de redes direcionadas por medições em alguns pontos estratégicos da rede. As cargas das unidades consumidoras e transformadores de distribuição são previstas *a priori*, baseadas na tipologia da carga existente na CELPE e em dados de pesquisa social de posses e hábitos de consumo. Essas previsões são representadas como números *fuzzy* trapezoidais.

A perda comercial é detectada pela diferença entre a leitura da demanda fornecida na saída do alimentador e o somatório das cargas dos transformadores de distribuição somadas às perdas ôhmicas. A origem dessa perda comercial é estimada resolvendo-se um sistema linear *fuzzy*, que calcula a correção de demanda horária dos transformadores e UCs de distribuição, inicialmente previstos de forma imprecisa pela tipologia de carga. Os transformadores e UCs com valores maiores de correção são selecionados para inspeção local.

**Palavras-chave** – Curvas de carga, fraudes, inspeções inteligentes e lógica *fuzzy*.

## I. INTRODUÇÃO

Esse artigo descreve a metodologia desenvolvida e os principais resultados do Projeto “Prospecção de Fraudes e Identificação de Ações de Combate a Perdas Comerciais” do programa de P&D da Companhia Energética de Pernambuco, CELPE, ciclo 2005/2006, com suporte financeiro do Programa de Pesquisa e Desenvolvimento Tecnológico do Setor de Energia Elétrica da Agência Nacional de Energia Elétrica, ANEEL. O projeto foi desenvolvido pela Fundação

---

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 VI Congresso de Inovação Tecnológica em Energia Elétrica (VI CITENEL), realizado em Fortaleza/CE, no período de 17 a 19 de agosto de 2011.

Dr. Henrique de O. Henriques e Eng<sup>o</sup> Otávio M. Junior trabalham no Laboratório de Estudos de Transmissão e Distribuição do Departamento de Engenharia Elétrica, LETD da UFF, como coordenador e colaborador, respectivamente (rique@vm.uff.br, omagalhaesjr@gmail.com).

Dr. Heraldo L. S. de Almeida e Felipe B. Carregosa pertencem ao Departamento de Engenharia Eletrônica e de Computação da Escola Politécnica da UFRJ e são pesquisadores associados ao Laboratório de Estudos de Transmissão e Distribuição do Departamento de Engenharia Elétrica, LETD da UFF (heraldo@ufrj.br, felipecarregosa@gmail.com).

Eng<sup>o</sup> José Aderaldo Lopes, Eng<sup>o</sup> Fábio Barros Ribeiro e Eng<sup>o</sup> Wladimir Lacerda de Moura, trabalham, respectivamente, nos Departamentos de Planejamento de Investimentos, Serviços de Rede e Gestão de Mercado, da Companhia Energética de Pernambuco, CELPE (jose.lopes@celpe.com.br, fabiobarros@celpe.com.br, wladimir@celpe.com.br).

Euclides da Cunha de Apoio Institucional a Universidade Federal Fluminense, FEC / UFF, através do Laboratório de Estudos de Transmissão e Distribuição, LETD, do Departamento de Engenharia Elétrica da UFF.

O principal produto desse trabalho é um software que será utilizado na CELPE para orientar e planejar inspeções mais efetivas. No desenvolvimento desse software foram utilizadas técnicas de inteligência computacional e medições feitas em um percentual de até 30% dentre os transformadores de distribuição de um alimentador.

Para se detectar e calcular as perdas comerciais é necessário obter o valor da energia fornecida às cargas alimentadas pela distribuidora e retirar as perdas ôhmicas das redes e equipamentos, e a energia faturada. Basicamente, a perda comercial pode existir pelos seguintes motivos:

- erros no processo de faturamento;
- furto de energia;
- ligações clandestinas;
- ligações irregulares.

As diferenças 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.

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.

O processo proposto neste projeto para avaliação e localização de perdas comerciais procura calcular a energia consumida pelos consumidores e, através das diferenças entre esta e a energia faturada, verificar onde a perda acontece. Esse cálculo depende da demanda ao longo do período faturado. Então, em primeiro lugar, deve-se estimar a curva de carga de cada unidade consumidora. No presente trabalho, essa estimativa é feita de duas maneiras:

- baseado na energia faturada e na tipologia de cargas;
- baseado na tipologia de cargas e na Pesquisa Social de Posses e Hábitos de consumo, PSPH.

A demanda estimada pelo método baseado na energia faturada pode conter erros grosseiros devido à possibilidade do consumo estar contaminado por irregularidades na unidade consumidora. Nesse método, portanto, o cálculo da curva de carga é feito após a obtenção da energia faturada para o cliente em um determinado período. Calcula-se a demanda média e multiplica-se pelos números, em p.u., da curva de carga de sua tipologia correspondente.

A demanda estimada através da PSPH utiliza dados declarados de uso de aparelhos elétricos por uma amostra de consumidores típicos. Desse modo, a metodologia se desvia dos possíveis erros de irregularidades no faturamento, mas, em geral, possui um desvio padrão bem elevado.

Outro problema que causa erros em ambos os processos é a defasagem entre o cadastro comercial e o elétrico das empresas. Essa defasagem, na maioria dos casos, é devido ao tempo requerido para atualização dos dados operacionais e de expansão da rede elétrica, bem como dificuldades para manutenção do cadastro comercial. Conhecendo os principais motivos que levam ao erro nos processos de previsão das curvas de carga das unidades consumidoras (UCs), os objetivos do presente projeto são:

- desenvolver software para localização de fraudes e ligações clandestinas em UCs de baixa tensão;
- desenvolver modelo de representação das incertezas da demanda das cargas utilizando lógica *fuzzy*;
- desenvolver modelo de agregação de cargas *fuzzy* para compor um alimentador primário;
- corrigir as incertezas das cargas, baseado em um modelo de estimação linear *fuzzy*;
- desenvolver modelo de inspeção supervisionada.

A solução encontrada para a representação horária da carga de cada UC, nesse trabalho, devido a todos os problemas encontrados na maioria das empresas de energia elétrica, é feita por números *fuzzy* trapezoidais [1], cujas incertezas são traduzidas pelos valores limites da base inferior do trapézio. Esses limites são determinados pelo ajuste da tipologia de carga e da pesquisa de campo. Nessa última, constam perguntas sobre, identificação do cliente, caracterização do domicílio, informações sobre energia elétrica, posses e hábitos de consumo de aparelhos e iluminação, dados sócio-econômicos, 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, conforme detalhado na seção II deste trabalho.

As curvas de carga de todas as UCs 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 grande desafio da pesquisa é descobrir um meio de redistribuir essa perda o mais próximo à realidade possível, pois dessa maneira os transformadores de maior perda são identificados.

O processo de alocação sugerido nesse projeto é o de es-

timação linear *fuzzy* [2], 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 ordenados em função do índice de perdas apresentados. 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.

O software desenvolvido nesse trabalho torna a escolha de UCs a inspecionar direcionada, mais efetiva e com custos bem menores do que as campanhas de inspeções massivas, realizadas em bairros, ruas ou comunidades.

A seção II desse trabalho descreve com detalhes todos os passos do processo. A seção III descreve os testes de campo realizados. A seção IV apresenta os resultados da aplicação do método no alimentador número 3 da subestação de Pau Amarelo, na Região Metropolitana do Recife, Pernambuco. A seção V apresenta as conclusões e próximos trabalhos.

## II. 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;
- operacionalização do processo.

A figura 1 a seguir exibe uma visão geral da metodologia proposta, ilustrando, passo a passo, a seqüência de processos automáticos (executados por algoritmos de inteligência computacional) e manuais (executados por análise humana) utilizada para obter as curvas de carga e as listas de UCs a inspecionar, com base nas informações obtidas do cadastro de clientes, da rede elétrica, da campanha de medidas e da pesquisa social de posses e hábitos.

Estas etapas são descritas detalhadamente nas subseções a seguir.

### 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 os seguintes:

- dados cadastrais dos clientes;
- informações de conexão à rede elétrica;
- resultados da pesquisa social de posses e hábitos;
- tabelas de consumo típico de aparelhos elétricos;
- curvas de carga obtidas na campanha de medidas.

### AI. Dados cadastrais e de conexão à rede elétrica

Os dados cadastrais e as informações de conexão das UCs à rede elétrica foram obtidos de arquivos extraídos das bases de dados CELPE.

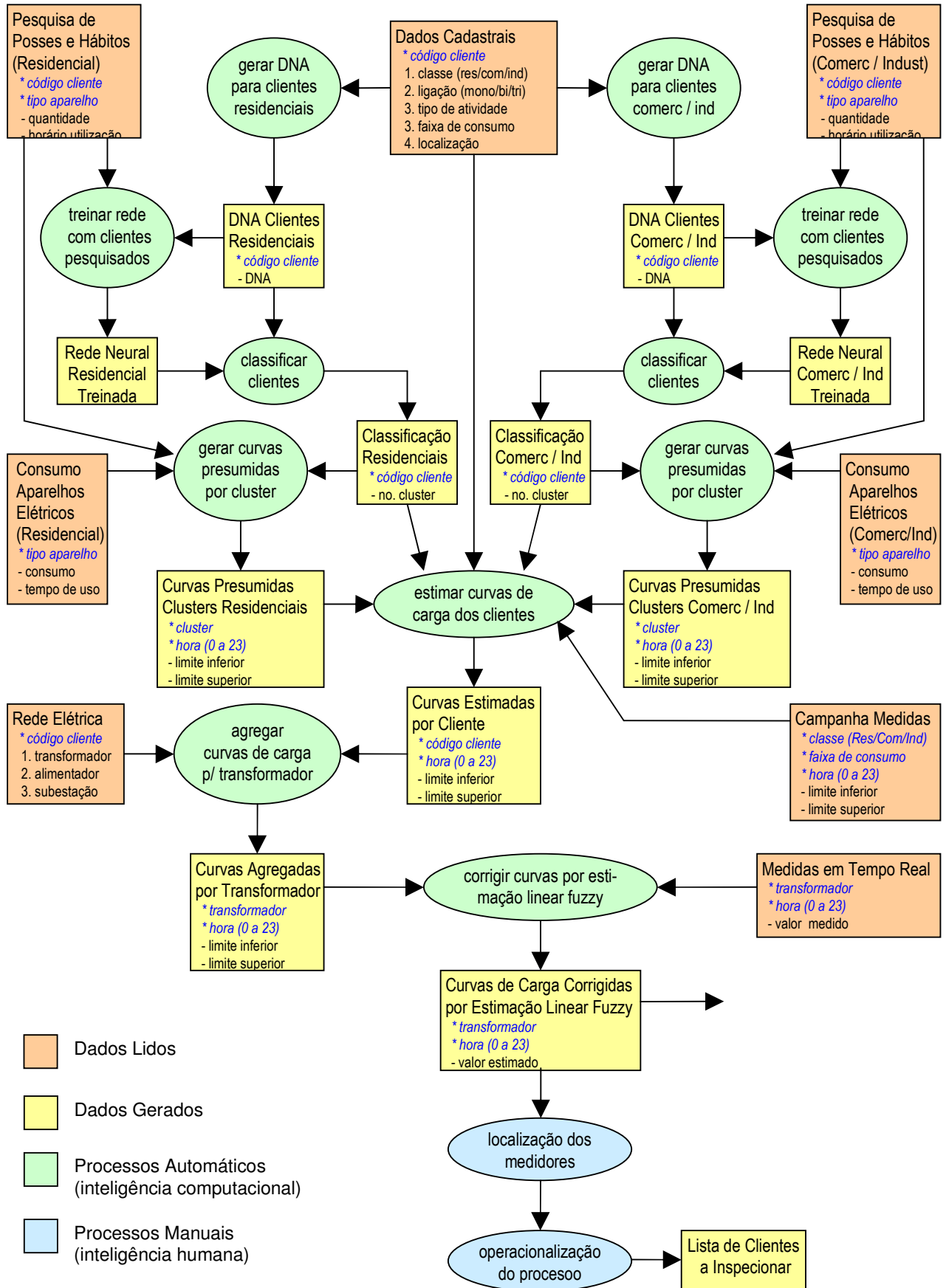


Figura 1. Visão geral do processo.

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 das correspondentes UCs 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 da CELPE 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
- coordenada X
- coordenada Y
- classe (residencial, comercial, industrial, etc.)
- tipo de tarifa
- tensão (monofásica, bifásica ou trifásica)
- consumo no mês de dezembro
- 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)
- indicativo de fraude (quando houver)

#### A2. Resultados da pesquisa social de posses e hábitos

A Pesquisa Social de Posses e Hábitos foi realizada no ano de 2008 com 932 consumidores residenciais, comerciais e industriais.

Além de informações sócio-econômicas 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 a seguir enumera as faixas de horário considerados na pesquisa de campo.

Tabela I. Faixas de horário consideradas na pesquisa de campo

Horário	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

#### A3. Tabelas de consumo típico 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 [3].

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.

#### A4. Curvas de carga 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 CELPE, realizada no ano de 2004.

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

Tabela II. Estrutura de Dados da Terceira Campanha de Medidas

Campo	Descrição
1	Classe de cliente (residencial, comercial, industrial, etc.)
2	Faixa de consumo médio mensal (0-100, 101-300, 301-500, 501-1000 e acima de 1000 kWh)
3	Identificação do cliente medido
4	Dia da semana (útil, sábado, domingo)
5	Hora (0 a 23)
6	Consumo medido no horário (kWh)

#### B. Formação dos Clusters por Rede Neural

O processo de formação dos clusters tem por objetivo a definição de grupos de consumidores que compartilham características comuns, visando identificar conjuntos de consumidores com tendências comportamentais semelhantes, tanto no que se refere aos hábitos de consumo quanto à propensão à fraude.

Inicialmente, as informações cadastrais dos clientes residenciais, comerciais e industriais 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* [4].

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 seguintes etapas:

1. Definição dos critérios de agrupamento
2. Determinação do “DNA” dos clientes
3. Treinamento das redes neurais
4. Classificação dos clientes nos clusters

### B1. Critérios de agrupamento

Os clientes das subestações Pau Amarelo (PAM) e João de Barros (JBR) foram agrupados por semelhança de perfil, de acordo com os seguintes 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)
- resultado da inspeção (tipo de irregularidade)

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

### B2. Determinação do “DNA” dos clientes

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, que são:

- variáveis discretas
- variáveis 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 da unidade consumidora.

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

1. Aloca-se um “gen” para cada possível valor de cada variável descritiva.
2. 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).

Por exemplo, para um cliente da classe industrial, o gen associado ao atributo “industrial” recebe o valor 1 (verdadeiro), enquanto que os gens associados às demais classes de clientes recebem o valor 0 (falso).

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

1. 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 quanti-

dade aproximadamente igual de ocorrências (20% do total de clientes) em cada faixa (isto é feito na etapa de segmentação das variáveis contínuas descrita anteriormente).

2. Aloca-se um “gen” para cada faixa de cada variável descritiva.
3. 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).

Por exemplo, para um cliente que cujo consumo médio nos últimos 12 meses esteja situado na faixa 3, o gen associado à faixa 3 da variável descritiva referente ao consumo médio recebe o valor 1 (verdadeiro), enquanto que os gens associados às faixas 1, 2, 4 e 5 desta mesma variável descritiva recebem o valor 0 (falso).

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

Tabela III. Significado dos “gens” no “DNA” do cliente

Gen	Descrição
01	Classe de Cliente = residencial
02	Classe de Cliente = comercial
03	Classe de Cliente = industrial
04	Tipo de Ligação = monofásica
05	Tipo de Ligação = bifásica
06	Tipo de Ligação = trifásica
07	Consumo Médio Mensal = faixa 1 ( até 37 kWh )
08	Consumo Médio Mensal = faixa 2 ( 37 a 75 kWh )
09	Consumo Médio Mensal = faixa 3 ( 75 a 107 kWh )
10	Consumo Médio Mensal = faixa 4 ( 107 a 161 kWh )
11	Consumo Médio Mensal = faixa 5 ( acima de 161 kWh )
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	Indicativo de Fraude = com indicativo
30	Indicativo de Fraude = sem indicativo
31	Indicativo de Fraude = não inspecionado
32	Resultado Inspeção = irregularidade código de 100 a 199
33	Resultado Inspeção = irregularidade código de 200 a 299
34	Resultado Inspeção = irregularidade código de 300 a 399
35	Resultado Inspeção = sem irregularidade (código 400)
36	Resultado Inspeção = não inspecionado

### B3. Treinamento das Redes Neurais

Dois 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.

#### B4. 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”).

#### 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 a 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 seguintes etapas:

1. Cálculo das curvas presumidas para cada cluster.
2. Cálculo das curvas presumidas por tipologia.
3. Cálculo das curvas de carga estimadas para clientes.
4. Agregação das curvas de carga por transformador.

##### C1. 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:

1. 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.

2. As curvas de carga correspondentes aos clientes entrevistados são normalizadas em p.u., de modo que o valor médio da curva estimada corresponda a 1 unidade.
3. Para cada cluster, é calculada uma curva de carga correspondente à média das curvas normalizadas de cada cliente entrevistado, que pertença àquele cluster, bem como curvas correspondentes a limites inferior e superior, estimadas a partir do desvio padrão calculado hora a hora.

A título de ilustração, a curva de carga presumida, calculada pelo software, para clientes pertencentes ao cluster 29 da classe residencial é exibida na Figura 2 a seguir.

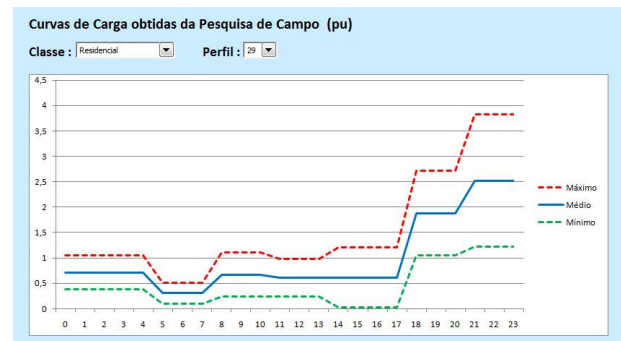


Figura 2. Curva de carga presumida para um cluster de clientes residenciais

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:

1. 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 1000 kWh e acima de 1000 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.
2. 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 1000 kWh em dias úteis são exibidas na figura 3 a seguir.

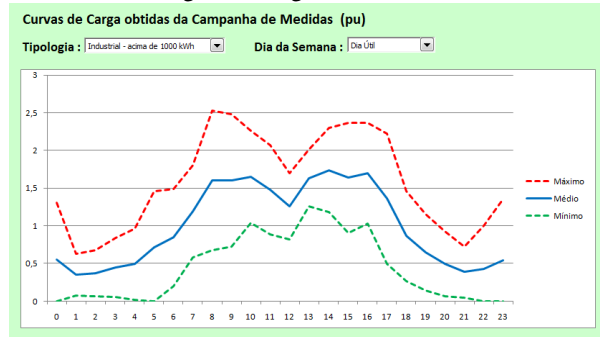


Figura 3. Curva de carga presumida para uma tipologia de cliente

### 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:

1. O cluster a que cada cliente pertence é determinado com base no processo descrito na subseção B (formação dos clusters).
2. A tipologia de cada cliente cadastrado é determinada com base na classe do cliente e no seu consumo médio dos últimos 12 meses.
3. 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 figura 4 abaixo, correspondente ao grau de pertinência do consumo do cliente às faixas de valores em kWh naquele horário.

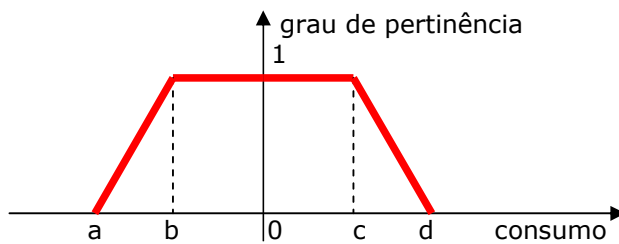


Figura 4. Representação trapezoidal dos valores *fuzzy* de consumo

### 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:

1. Para cada cliente das subestações JBR e PAM, foram determinados o alimentador e o transformador de distribuição ao qual o cliente está conectado.

2. 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.

A título de ilustração, a curva de carga *fuzzy* prevista pelo software para o trafo T25365 do alimentador 01C3 da subestação PAM em um dia útil é exibida na figura 5 a seguir.

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

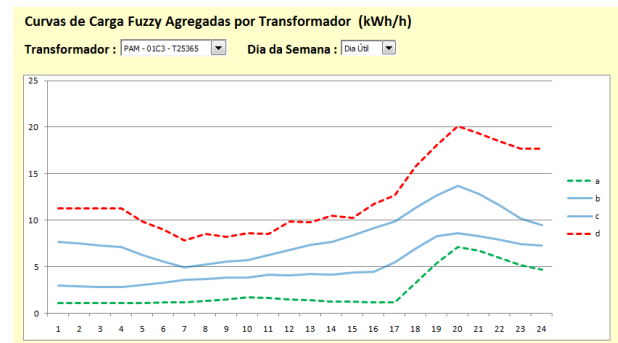


Figura 5. Curva de carga presumida para uma tipologia de cliente

### 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 figura 6 a seguir, 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.

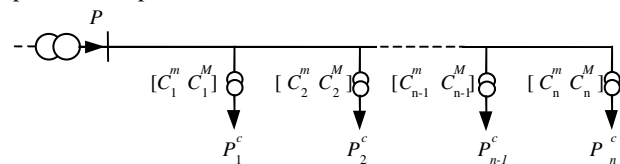


Figura 6. Circuito radial típico considerado na alocação das perdas

O modelo calcula as correções  $\Delta P_i$ ,  $i=1,2,\dots,n$ , a serem aplicadas aos valores estimados  $P_i^c$ ,  $i=1,2,\dots,n$ , de modo que a diferença  $\Delta P$  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  $\Delta P_i$ ,  $i=1,2,\dots,n$ , estejam dentro da região de pertinência ilustrada na figura 7 a seguir.

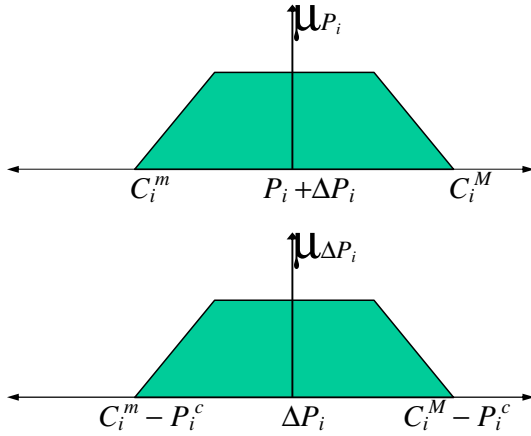


Figura 7. Curvas de pertinência para  $P_i$  e  $\Delta P_i$

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

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

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

- onde  $\Delta P_i$  = 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 número 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 I 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 figura 8 a seguir mostra um transformador cujo valor corrigido pelo simplex, em azul, está coerente com a previsão inicial, em verde. Esse transformador não precisa ter medidor.

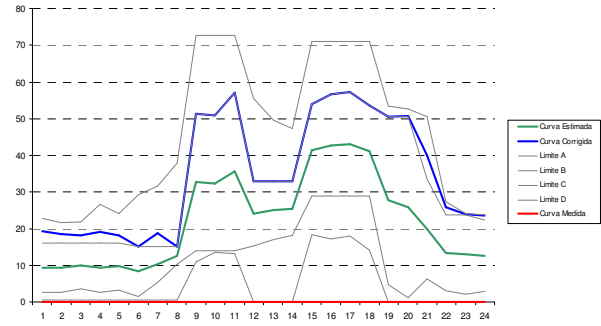


Figura 8. Curvas de carga *fuzzy* estimada e corrigida, no caso de um transformador que não precisa ter medidor

A figura 9 a seguir mostra, num gráfico de barras, que a diferença entre os valores corrigidos e estimados se comporta de maneira uniforme, ou seja, em todos os horários a correção é positiva.

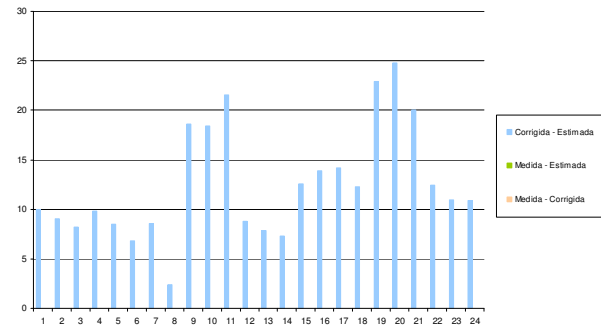


Figura 9. Diferença entre valor corrigido e estimado em um transformador sem necessidade de medidor.

A figura 10 a seguir mostra um exemplo em que o valor corrigido oscila em relação ao estimado, isto é, hora a correção é positiva e hora negativa. Esse transformador é candidato a ter um medidor instalado. A figura 11 mostra o correspondente gráfico de barras.

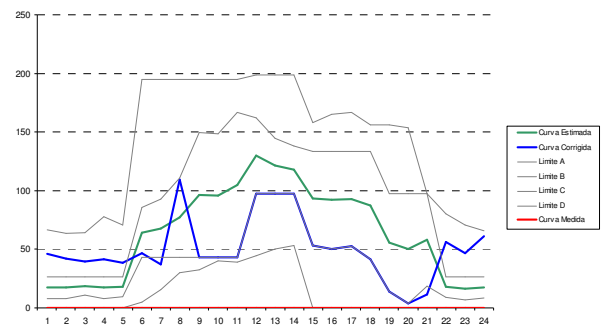


Figura 10. Curvas de carga *fuzzy* estimada e corrigida, no caso de um transformador que precisa ter medidor



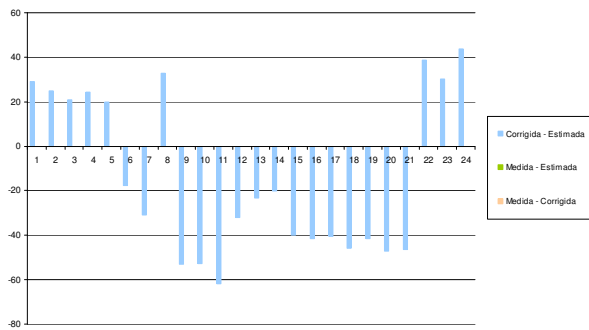


Figura 11. Diferença entre valor corrigido e estimado em um transformador com necessidade de medidor.

A figura 12 a seguir mostra um caso em que, embora as diferenças sejam todas positivas, como mostra a figura 13, a curva corrigida pelo método simplex apresenta uma forma distorcida da estimada. Esse detalhe está marcado na figura 12 por um círculo vermelho e traduz a presença de erros na caracterização dos consumidores ligados a esse transformador.

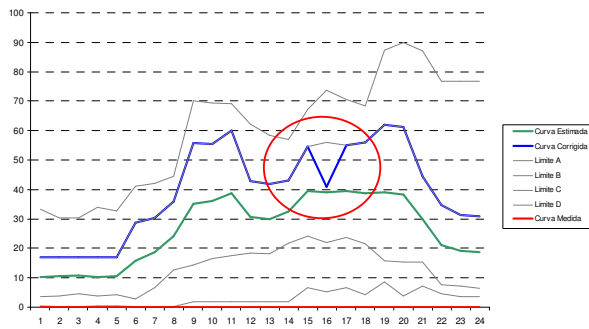


Figura 12. Resultado para transformador com forma diferente do esperado

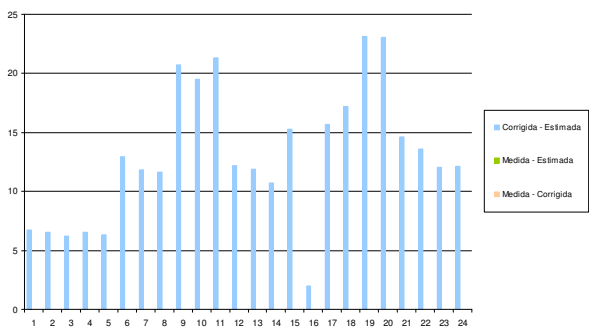


Figura 13. Diferença entre valores corrigidos e estimados das curvas da figura 12.

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.

#### G1. 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.

A figura 14 a seguir mostra a medição, em vermelho, coletada para o transformador da figura 10.

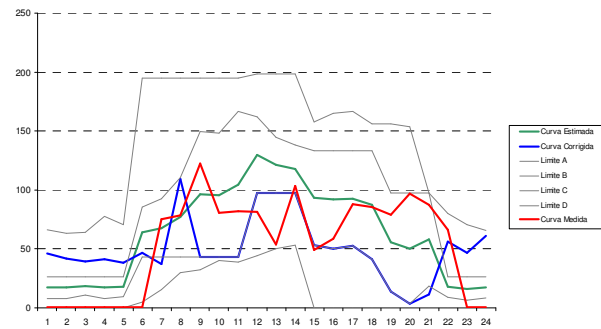


Figura 14. Gráfico da medição do transformador da figura 10

As diferenças “corrigidos – estimados” e “medidos – estimados” estão representadas no gráfico de barras da figura 15 a seguir.

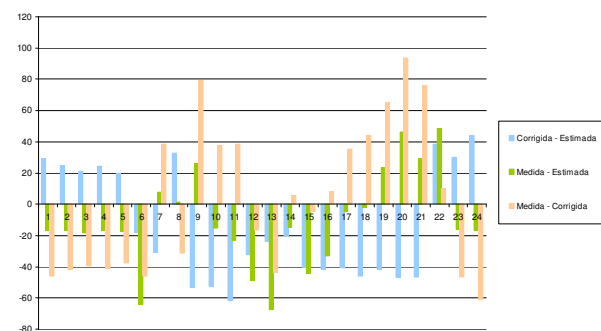


Figura 15. Diferenças do transformado da figura 14

Supondo:

$X$  = “corrigidos – estimados”;

$Y$  = “medidos – estimados”.

Para saber se existe no transformador algum erro de vinculação, avaliam-se os sinais das diferenças de  $X$  e  $Y$ . Se  $X$  e  $Y$  tiverem sinais contrários, existe problema de vinculação. O gráfico da Figura 16 a seguir ilustra essa situação.

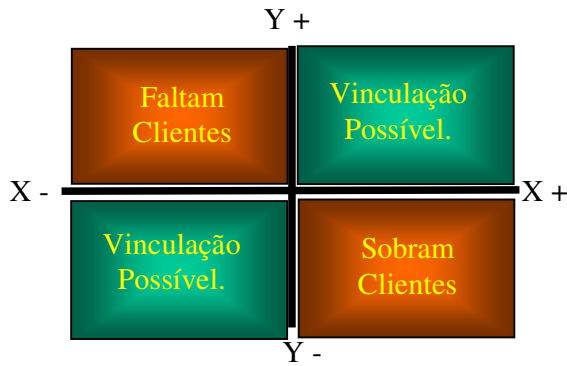


Figura 16. Erros de vinculação possíveis em um transformador

No caso de X ser positivo e Y negativo, isto quer dizer que o transformador tem menos clientes do que foi considerado no cálculo, isto é, sobram clientes. Caso X seja negativo e Y positivo, isto quer dizer que estão faltando alguns clientes que não foram considerados no cálculo. Em ambos os casos deve-se ir ao campo efetuar a atualização do cadastro elétrico.

Após as devidas correções, o próximo passo é resolver o sistema linear *fuzzy* para cada transformador com medidor e gerar a primeira lista de inspeção contendo clientes cujo ramal de ligação tenha resultado em um valor corrigido maior que 30 % do valor estimado. A figura 17 a seguir ilustra o processo.

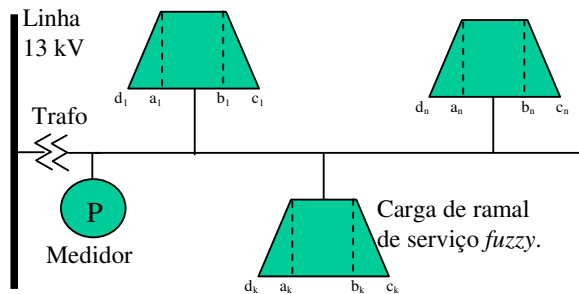


Figura 17. Diagrama unifilar da modelagem de um transformador de distribuição

Alguns medidores podem ser instalados nos ramais de serviço dos clientes de baixa tensão. O critério de onde instalá-los é o mesmo que foi utilizado em transformadores.

### G2. Segunda Etapa

Após acertar erros de vinculação e regularizar os clientes dos transformadores da primeira etapa, deve ser resolvido novamente o sistema linear *fuzzy* para todos os transformadores, considerando agora os medidores já instalados. Aqueles que estiverem valores corrigidos mais de 30 % acima do valor estimado devem ter clientes selecionados para inspeção.

Em seguida, deve ser calculado o simplex para os transformadores anteriores e listados os clientes individuais que possuam valores corrigidos maiores que 30 % dos valores estimados.

## III. TESTES DE CAMPO

Os testes de campo foram essenciais para o desenvolvimento do processo de inspeção por medições. Foi através dele que se verificou a necessidade de se construir um indicador que apontasse a necessidade de verificação da vinculação do cliente ao transformador e da necessidade de se verificar a atualização de alguns dados cadastrais, tais como a classe do cliente.

Os testes de campo foram feitos medindo-se todos os transformadores do alimentador 3 da subestação de Pau Amarelo. Porém, para a correção pelo método simplex, escolheu-se apenas 9, ou 30% dos transformadores. As demais medidas serviram para comparar o valor estimado pelo processo e o valor real medido.

A figura 18 a seguir mostra o diagrama unifilar do alimentador. Os transformadores em vermelho foram os selecionados para terem seus medidores considerados no cálculo de estimação dos valores corrigidos. Os transformadores marcados com uma seta foram escolhidos para inspeção, levando-se em consideração apenas a diferença maior ou igual a 30% entre os valores “corrigidos – estimados”, sem fazer correção de vinculação.

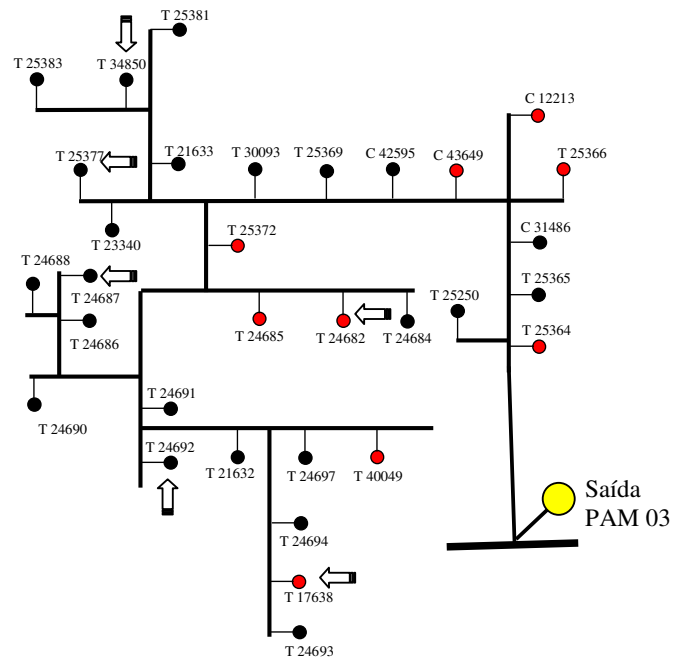


Figura 18. Diagrama unifilar da subrede Pau Amarelo 3 com os medidores localizados

## IV. RESULTADOS 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 IV, a seguir.

Tabela IV. Legenda dos resultados da inspeção

Legenda	Descrição
NORMAL	Unidade consumidora sem fraude.
N LOCALIZADO	Unidade consumidora não localizada.
FECHADO	3ª visita unidade consumidora fechada.
SELO AUSENTE	Equipamento de medição sem selo de laboratório havendo acesso a sua parte interna.
SELO DESPAD	O equipamento de medição encontra-se com os selos despadrionados.
SL PRENSADOS	O equipamento de medição encontra-se com os selos prensados.
P BLO TERM	O equipamento de medição encontra-se com uma ponte no bloco de terminais.
SELO ROMPIDO	O equipamento de medição encontra-se com selos de laboratório rompidos.
DISP APAGADO	Medidor com display apagado.
C CORTADO	Unidade consumidora continua cortada.
IMPEDIDO	Cliente impediu a inspeção.
DESOCUPADO	Imóvel desocupado.
E M IRREGULAR	O equipamento de medição encontra-se com bobina irregular.
LIGADO DIRETO	A unidade consumidora encontra-se ligado direto sem passa pela medição.
DERI CLANDEST	A unidade consumidora encontra-se com derivação clandestina antes da medição.
E M FURADO	O equipamento de medição encontra-se com sua tampa furada havendo acesso a sua parte interna.

A figura 19 a seguir ilustra o resultado para clientes comerciais.

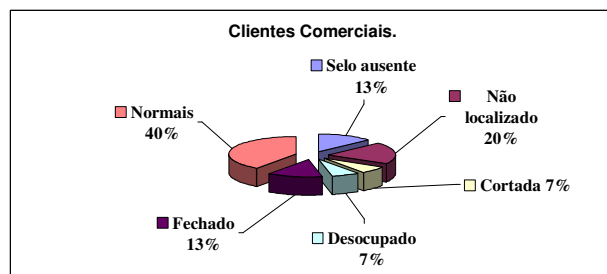


Figura 19. Ocorrências verificadas nas inspeções de clientes comerciais

A figura 20 a seguir ilustra o resultado das inspeções da classe residencial B1.

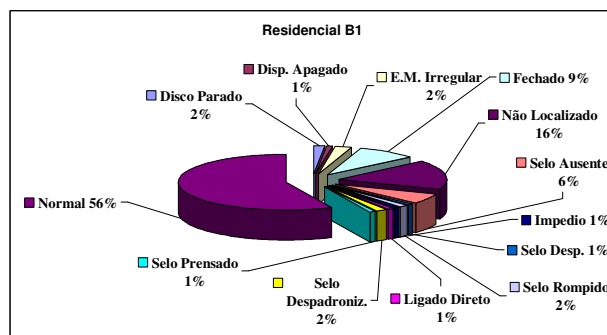


Figura 20. Ocorrências verificadas em clientes residenciais B1

A figura 21 a seguir ilustra o resultado das inspeções da classe residencial Baixa Renda.

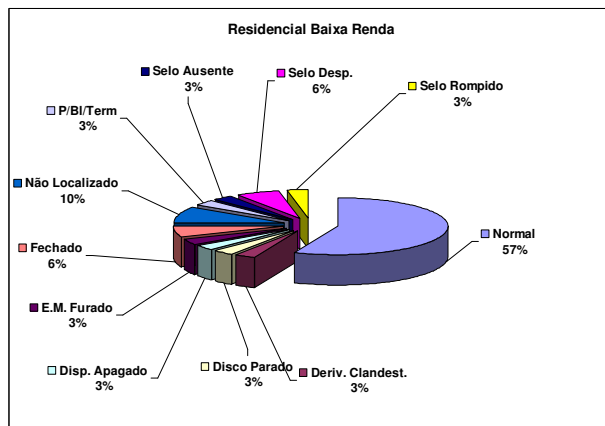


Figura 21. Ocorrências verificadas em clientes residenciais baixa renda

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 32 casos de irregularidades ligadas a furto de energia envolvendo 29 casos de problemas ligados ao medidor, 2 ligações clandestinas e 1 cliente que impediu a inspeção.

## V. CONCLUSÕES E TRABALHOS FUTUROS

Os testes de campo serviram para organizar a operacionalização do processo, com a ajuda do software desenvolvido, nos demais alimentadores. Salientamos que para a melhor assertividade do sistema, é necessário a permanente atualização dos cadastros comercial e da rede elétrica. O método proposto nessa pesquisa 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 os dados de quais os transformadores que estão ligados ao alimentador em estudo.

Como trabalho futuro, fica a recomendação para CELPE aplicar o método completo para as demais linhas das subestações de Pau Amarelo e João de Barros. Ao aplicar a segunda etapa, o número de clientes a inspecionar deve diminuir e o percentual de irregularidades apresentadas deve aumentar, reduzindo os custos das ações de combate aos furtos de energia.

## VI. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] H. J. Zimmerman, "Fuzzy Set Theory and its Applications", 3rd edition, Kluwer, Boston, 1996.
- [2] 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.
- [3] Eletrobrás, "http://www.eletrobras.com/elb/procel/main.asp", acesso em 13/09/2009.
- [4] T. Kohonen, "Self-Organization and Associative Memory", Berlin: Springer-Verlag, 1987.