

SENDI 2004
XVI SEMINÁRIO NACIONAL DE DISTRIBUIÇÃO DE ENERGIA ELÉTRICA

Sistema Computacional para Construção de Tipologias de Curvas de Carga

A.L.A.Araújo
COSERN

V.L.O.Castellani
FPLF

P.M.Ribeiro
CEPEL

J.F.M.Pessanha
CEPEL

andre.araujo@cosern.com.br e francisc@cepel.br

Palavras-chaves : análise de agrupamentos, curvas de carga, sistema computacional, tipologias

Resumo : Os perfis típicos da demanda horária de eletricidade são informações fundamentais no cálculo das tarifas que remuneram o serviço de distribuição de energia elétrica. Estes perfis são obtidos por meio de um estudo de caracterização da carga, onde são utilizados métodos de análise de agrupamentos para identificar os perfis típicos ou tipologias, a partir de uma amostra de medições de curvas de carga de clientes e pontos da rede de distribuição. Visando agilizar a caracterização da carga, de forma a beneficiar particularmente os estudos tarifários, o CEPEL, em conjunto com a COSERN, desenvolveu o ANATIPO, um sistema computacional dotado de diversas funcionalidades para auxiliar a construção das tipologias. Dentre estas funcionalidades destacam-se as rotinas para leitura de medições provenientes de medidores de diferentes fabricantes, rotinas gráficas para identificação das curvas características de cada medição e a disponibilização de três técnicas de análise de agrupamentos. Destaca-se também a rotina para identificação automática das curvas características, capaz de acelerar a análise das medições. As tipologias obtidas pelo programa são automaticamente ajustadas ao mercado e gravadas em planilhas eletrônicas, prontas para serem incluídas nas propostas tarifárias a serem enviadas à ANEEL. Enfim, este artigo tem como objetivo apresentar o sistema ANATIPO e as suas principais funcionalidades.

1. INTRODUÇÃO

Os perfis típicos da demanda horária de eletricidade, também conhecidos como tipologias, são informações fundamentais no cálculo das tarifas [1,2] que remuneram o serviço de distribuição de energia elétrica e úteis na indicação de alternativas para o gerenciamento pelo lado da demanda [3,4].

As tipologias de clientes são definidas para cada classe de consumo e representam a diversidade do comportamento dos consumidores da classe. Da mesma forma são definidas as tipologias de redes que representam os perfis típicos das solicitações de potência em pontos selecionados da rede de distribuição, alimentadores e transformadores. Estas informações, em conjunto com os custos marginais por nível de tensão, permitem elaborar a estrutura de custos marginais de fornecimento, aberta nos postos ponta e fora de ponta, que serve de base para o cálculo das tarifas de uso dos sistemas de distribuição [5].

Basicamente, a construção das tipologias de um determinado segmento (nível de tensão, classe de consumo ou tipo de rede) inicia com a coleta de uma amostra de medições de curvas de carga de clientes ou redes. Em seguida, aplica-se alguma técnica de análise de agrupamentos (*cluster analysis*) [6-9] na amostra de medições para particioná-la em agrupamentos (*clusters*) mutuamente exclusivos, de tal forma que as medições em um mesmo *cluster* sejam semelhantes entre si, porém diferentes das curvas pertencentes aos outros *clusters*. Os perfis horários típicos existentes na amostra são os perfis médios dos *clusters*.

No setor elétrico brasileiro, o SNACC¹ foi o primeiro sistema computacional para construção de tipologias de curvas de carga. Neste sistema, desenvolvido em meados da década de 1980, os *clusters* são obtidos pelo uso combinado de dois métodos estatísticos para análise de agrupamentos: nuvens dinâmicas – NUDYC [10] e método de descrição e classificação ascendente hierárquico – DESCR2 [10].

Os dois métodos são executados em sequência: primeiro o método das nuvens dinâmicas identifica os *clusters* iniciais (formas fortes) e em seguida estes *clusters* são agrupados pelo DESCR2, um método hierárquico semelhante ao de Ward [8]. Apesar da maneira sofisticada de formar os *clusters*, o SNACC não é um sistema amigável e possui limitações que dificultam o processo de construção das tipologias, entre as quais destaca-se o fato de não disponibilizar saídas gráficas para visualização dos agrupamentos.

Nos sistemas desenvolvidos mais recentemente, a disponibilidade de várias saídas gráficas representou um grande avanço. Entretanto, mesmo nestes sistemas, a construção das tipologias ainda é pouco confortável e pode ser melhorada por meio da incorporação de novas rotinas que facilitem e tornem mais produtivo o processo de construção das tipologias.

Visando agilizar o estudo de caracterização da carga, de forma a beneficiar particularmente os estudos tarifários, o Centro de Pesquisas de Energia Elétrica - CEPTEL, em conjunto com a Companhia Energética do Rio Grande do Norte - COSERN, desenvolveu o ANATIPO, um sistema computacional dotado de diversas funcionalidades. Dentre estas funcionalidades destacam-se as rotinas para leitura de medições provenientes de medidores de diferentes fabricantes (arquivos em formato público ou em planilhas eletrônicas), rotinas gráficas que facilitam a identificação visual das curvas características dos dias úteis, sábado e domingo de cada medição e uma variedade de técnicas de agrupamento disponibilizadas pelo programa: *K-Means* [6,7], *Ward* [8] e *FCM* [9]. Destaca-se também a rotina para identificação automática das curvas características, um item inovador capaz de acelerar substancialmente a análise das medições. As tipologias, principais resultados obtidos pelo programa, são automaticamente ajustadas ao mercado e gravadas em planilhas eletrônicas, prontas para serem incluídas nas propostas tarifárias a serem enviadas à Agência Nacional de Energia Elétrica – ANEEL. Enfim, este artigo tem como objetivo apresentar o sistema computacional ANATIPO e as suas principais funcionalidades.

¹ Sistema Nacional de Análise de Curvas de Carga - SNACC, desenvolvido pelo Departamento Nacional de Águas e Energia Elétrica (DNAEE)

A seguir, a seção 2 apresenta o processo de construção das tipologias. A seção 3 apresenta a heurística adotada para fazer a seleção automática das curvas características. Os algoritmos dos métodos de análise de agrupamentos implementados no ANATIPO são apresentados na seção 4. A seção 5 apresenta algumas saídas gráficas do sistema ANATIPO e, finalmente, as principais conclusões do trabalho são resumidas na seção 6.

2. CONSTRUÇÃO DAS TIPOLOGIAS

A construção das tipologias inicia com a obtenção das medições de curvas de carga de clientes e redes. Na baixa tensão o número de consumidores é bastante elevado, os medidores instalados são do tipo eletromecânico e não possuem memória de massa para armazenar os registros de demanda horária. Neste caso, deve-se realizar uma amostragem, ou melhor, uma campanha de medidas, para coletar as medições de curva de carga de clientes e transformadores de distribuição.

Por outro lado, na alta tensão o número de pontos a serem medidos é menor e os medidores instalados nos clientes são eletrônicos, com capacidade para armazenar os registros de demanda horária. Neste caso, a coleta das medições é direta, bastando apenas recuperar as medições armazenadas na memória de massa dos medidores. Nas duas situações, recomenda-se que cada medição tenha uma abrangência de no mínimo cinco dias úteis, um sábado e um domingo completos.

Após a coleta das medições deve-se fazer uma inspeção visual de cada medição, com o objetivo de identificar três curvas diárias características do ponto de medição: uma curva típica para o dia útil, uma para o sábado e uma para o domingo, conforme apresentado na Figura 1.

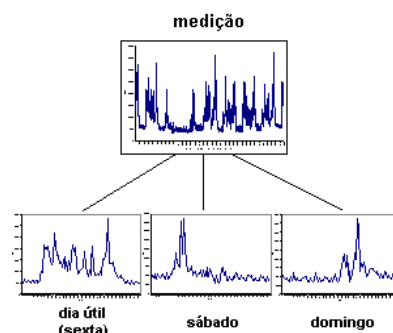


FIGURA 1 - SELEÇÃO DAS CURVAS CARACTERÍSTICAS

A identificação das tipologias de um determinado segmento (nível de tensão, classe de consumo ou tipo de rede) envolve a aplicação de alguma técnica de análise de agrupamentos ou *cluster analysis*, com a finalidade de dividir o conjunto de curvas características do dia útil em agrupamentos internamente homogêneos (*clusters*), a partir dos quais são extraídas as tipologias, por exemplo, a média das curvas em cada *cluster*. Este processo é resumido na Figura 2.

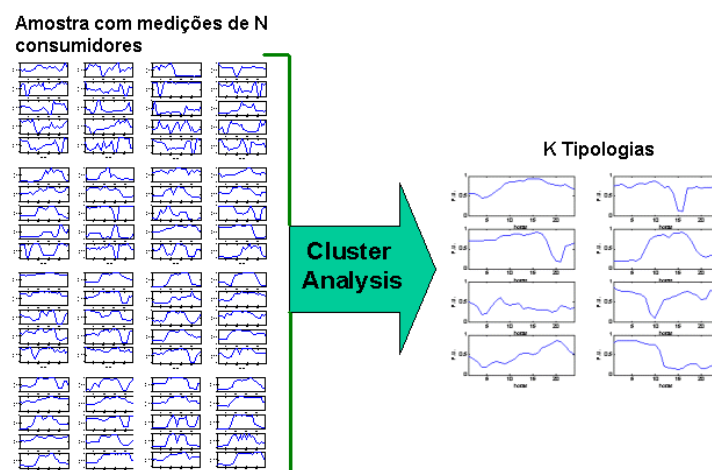


FIGURA 2 – K TIPOLOGIAS OBTIDAS A PARTIR DE N MEDIÇÕES ($K < N$)

As curvas características do dia útil formam o conjunto de curvas a serem classificadas, portanto, são elas que definem os perfis das tipologias. As curvas de sábado e domingo servem apenas para calcular fatores de ponderação usados no ajuste [11] dos perfis típicos ao montante de energia anual (MWh) do segmento que eles representam. Este ajuste é conduzido de acordo com o procedimento a seguir :

- 1) Em cada *cluster* i , $i=1,k$, some separadamente as curvas do dia útil (U), do sábado (S) e do domingo (D), obtendo três curvas agregadas (perfis típicos) para cada *cluster* i : U_h^i , S_h^i , D_h^i , $\forall h = 1, 24$ horas.

- 2) Totalize as curvas horárias do dia útil, sábado e domingo em cada *cluster* i :

$$E_u^i = \sum_{h=1}^{24} U_h^i \quad E_s^i = \sum_{h=1}^{24} S_h^i \quad E_d^i = \sum_{h=1}^{24} D_h^i \quad (1)$$

- 3) Calcule os fatores de ponderação do sábado (P_s^i) e do domingo (P_d^i) de cada *cluster* i :

$$P_s^i = \frac{E_s^i}{E_u^i} \quad P_d^i = \frac{E_d^i}{E_u^i} \quad (2)$$

- 4) Calcule a energia anual da curva agregada do *cluster* i :

$$C_i^{anual} = E_u^i \cdot 24(N_u + N_s \cdot P_s^i + N_d \cdot P_d^i) \quad (3)$$

onde N_u , N_s e N_d são, respectivamente, o número de dias úteis, sábados e domingos mais feriados do ano.

- 5) Calcule a participação do *cluster* i no mercado total do segmento que ele representa:

$$\alpha_i = \frac{C_i^{anual}}{\sum_{i=1}^k C_i^{anual}} \quad (4)$$

- 6) Obtenha a parcela do mercado correspondente ao *cluster* i :

$$ET_i = \alpha_i \cdot M \quad (5)$$

onde M é o mercado anual em MWh do segmento que ele representa.

- 7) Calcule a demanda média (MW) da curva agregada do *cluster* i :

$$D_i^{média} = \frac{ET_i}{24 \cdot (N_u + N_s P_s^i + N_d P_d^i)} \quad (6)$$

- 8) Finalmente, ajuste a tipologia i ao mercado, multiplicando a demanda média pela demanda horária em p.u. da média:

$$T_h^i = D_i^{média} \cdot \frac{U_h^i}{\left(\frac{\sum_{h=1}^{24} U_h^i}{24} \right)} \quad (7)$$

3. IDENTIFICAÇÃO AUTOMÁTICA DAS CURVAS CARACTERÍSTICAS

A identificação das curvas características pode ser realizada por meio de uma inspeção visual de cada arquivo de medição. Entretanto, este procedimento consome muito tempo do analista conhecedor do mercado. Com o intuito de acelerar esta identificação, introduziu-se no ANATIPO uma rotina que seleciona automaticamente as curvas características do dia útil, do sábado e domingo. Esta rotina não elimina o trabalho do analista, mas visa auxiliá-lo. A idéia é que no momento da inspeção visual o

programa forneça a indicação de três curvas ao analista, deixando para ele a decisão final de aceitar a sugestão do programa ou a selecionar manualmente outras curvas.

O critério implementado na rotina de seleção automática é bastante simples: dado um arquivo de medição, calculam-se as três curvas diárias médias, uma para sábado, uma para domingo e outra para o dia útil. Em seguida identificam-se para cada um destes dias quais as medições diárias registradas mais próximas de suas curvas médias.

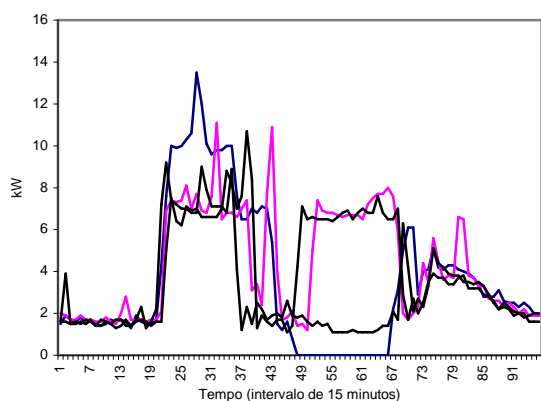
Como a média não é uma medida robusta [12], as medições que possuam valores nulos e as realizadas em dias de instalação e retirada do medidor não participarão do cálculo das curvas médias. Para o caso das curvas do dia útil adota-se um critério adicional, com o objetivo de identificar curvas com energias atípicas. Neste critério calculam-se inicialmente os quartis da distribuição de energias das medições do dia útil e, a partir deles, definem-se os seguintes limites:

$$CS = Q3 + \alpha(Q3-Q1) \quad (8)$$

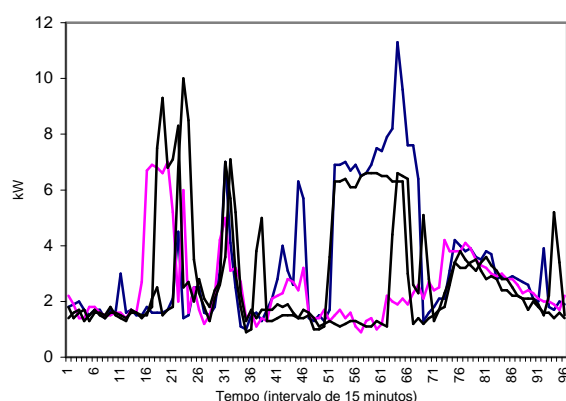
$$CI = Q1 - \alpha(Q3-Q1) \quad (9)$$

onde Q1 e Q3 representam, respectivamente, o primeiro e terceiro quartis da distribuição de energias das medições do dia útil e α representa uma constante arbitrária, usualmente igual a 1,5 [12].

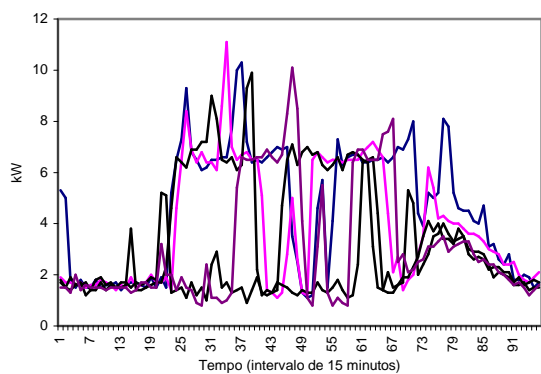
Todas as curvas que possuam energias abaixo do limite inferior ou acima do limite superior são eliminadas do cálculo da média. Para ilustrar a metodologia de seleção automática, considere um arquivo de medição, cujas medições diárias são apresentadas nas Figuras 3 e 4.



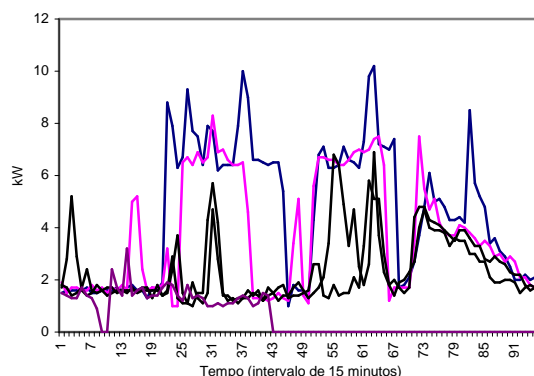
Sábado



Domingo

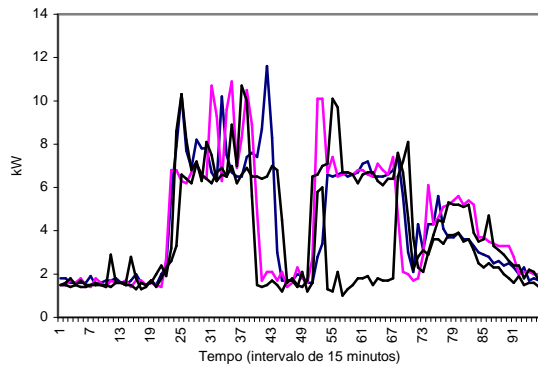


Segunda-feira

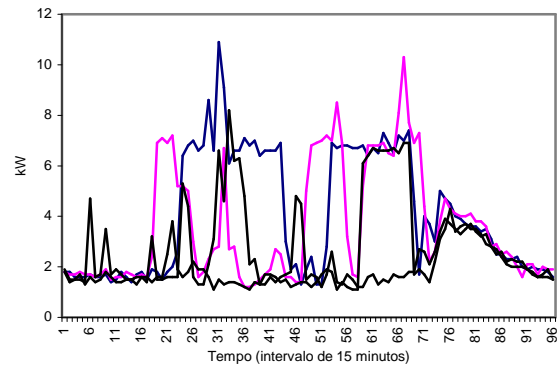


Terça-feira

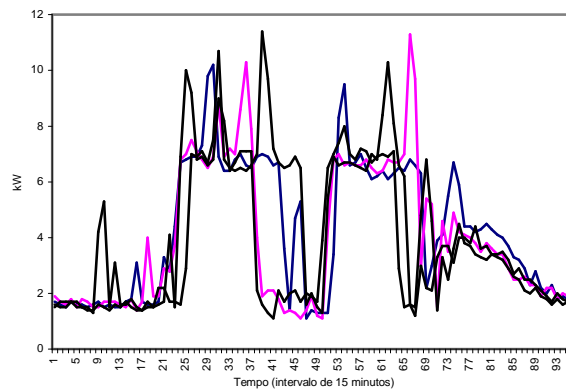
FIGURA 3 – MEDIÇÕES DE CURVA DE CARGA



Quarta-feira



Quinta-feira



Sexta-feira

FIGURA 4 – MEDIÇÕES DE CURVA DE CARGA

Aplicando o critério de seleção automática nas curvas apresentadas nas Figuras 3 e 4, obtêm-se três curvas características, ilustradas na Figura 5:

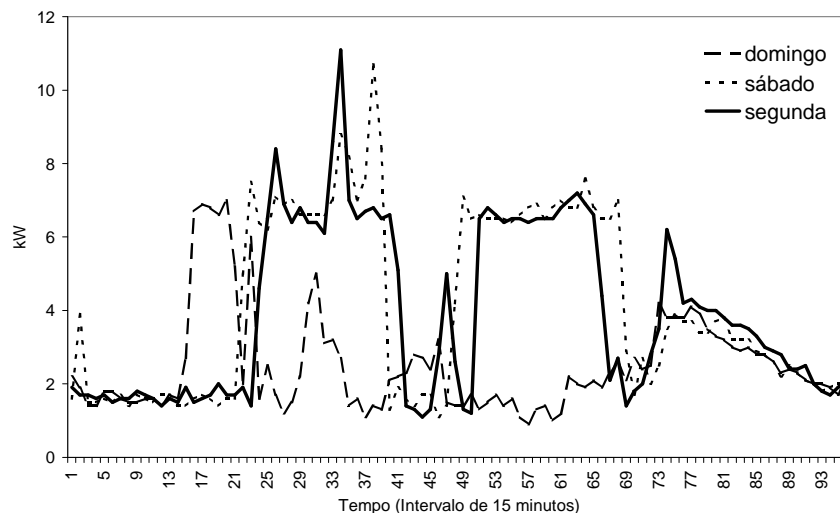


FIGURA 5 – CURVAS CARACTERÍSTICAS SELECIONADAS AUTOMATICAMENTE

4. MÉTODOS DE ANÁLISE DE AGRUPAMENTOS IMPLANTADOS NO ANATIPO

A análise de agrupamentos denota uma ampla variedade de métodos de classificação não supervisionados [13], usados para encontrar uma estrutura natural de agrupamentos em um conjunto de objetos multidimensionais, baseados em uma medida de similaridade ou distância entre os objetos. Os vários métodos de análise de agrupamentos visam resolver o problema de como dividir um

conjunto com n objetos em k *clusters* mutuamente exclusivos, de tal forma, que os objetos em um mesmo *cluster* sejam semelhantes entre si, mas diferentes dos objetos pertencentes aos outros *clusters*.

Admitindo que os n objetos sejam caracterizados por p atributos de natureza quantitativa, cada objeto pode ser representado por um vetor x_i , $i=1,2,\dots,n \in \mathfrak{R}^p$. Neste caso, a distância euclidiana é usualmente a métrica adotada para avaliar o grau de semelhança entre os objetos.

O sistema ANATIPO disponibiliza ao analista três métodos de análise de agrupamentos: *K-Means*, *Ward* e *FCM*. Qualquer um dos métodos pode ser escolhido e uma comparação dos desempenhos destes métodos, quando aplicados na classificação de curvas de carga, pode ser encontrada na referência [14].

4.1. K-Means

Um dos mais conhecidos métodos de *cluster analysis* é o *K-Means*. Este método divide um conjunto de n vetores em k *clusters*, $C_j, j=1,k$, de tal forma que os centros de cada *cluster*, $c_j, j=1,k$, minimizem a soma dos quadrados intra-*cluster* (SQIntra):

$$SQIntra = \sum_{j=1}^k \left(\sum_{i, x_i \in C_j} \|x_i - c_j\|^2 \right) \quad (10)$$

Na equação 1, $\|x_i - c_j\|^2$ é o quadrado da distância euclidiana entre o objeto x_i e o centro de gravidade c_j , do *cluster* C_j , onde o objeto foi alocado. Cada objeto pode pertencer a apenas um dos k *clusters*, portanto, a classificação dos n objetos pode ser representada por uma matriz binária U de dimensões $n \times k$, onde $u_{ij} = 1$ se o i -ésimo objeto pertence ao j -ésimo *cluster* e $u_{ij} = 0$ caso contrário. Se os centros dos k *clusters* são fixos, a partição ótima consiste em alocar cada objeto no *cluster* com o centro de gravidade mais próximo, ou seja, os valores de u_{ij} são definidos pela seguinte regra:

$$u_{ij} \begin{cases} 1 & \text{se } \|x_i - c_j\| \leq \|x_i - c_k\| \forall k \neq j \\ 0 & \text{caso contrário} \end{cases} \quad (11)$$

A matriz U tem as seguintes propriedades:

$$\sum_{j=1}^k u_{ij} = 1 \quad \text{A soma dos elementos da } i\text{-ésima linha é } 1$$

$$\sum_{i=1}^n u_{ij} = n_j \quad \text{A soma dos elementos da } j\text{-ésima coluna é o total de objetos no } j\text{-ésimo } cluster$$

$$\sum_{j=1}^k \sum_{i=1}^n u_{ij} = n_1 + \dots + n_k = n \quad \text{A soma de todos os elementos de } U \text{ é igual ao total de objetos do conjunto analisado}$$

Por outro lado, para uma dada partição dos n objetos a matriz U é fixa e o centro do j -ésimo *cluster*, $j=1,k$, que minimiza a função 1 é a média dos objetos do *cluster*:

$$c_j = \frac{1}{n_j} \sum_{i, x_i \in C_j} x_i \quad (12)$$

Computacionalmente o método *K-Means* pode ser implementado segundo o algoritmo a seguir, onde a matriz U e os centros dos *clusters* são obtidos de forma iterativa:

Passo 1 - Inicialize os centros dos *clusters*. Usualmente a inicialização se dá pelo sorteio de k objetos entre os n a serem classificados.

Passo 2 - Determine a matriz U de acordo com o critério do centro mais próximo, conforme em 11.

Passo 3 - Calcule o valor da função SQIntra. Pare se o valor da função estiver abaixo de uma tolerância pré-especificada ou se a melhoria em relação à iteração anterior for desprezível.

Passo 4 - Atualize os centros dos *clusters* de acordo com a equação 11 e volte para o passo 2.

Este algoritmo é computacionalmente eficiente e produz bons resultados se os *clusters* são compactos, esféricos e bem separados no espaço [13]. Entretanto, o algoritmo não garante a convergência para uma solução ótima e o seu desempenho depende dos centros iniciais escolhidos no passo 1.

4.2. Fuzzy Clustering Method

Para melhorar o desempenho do algoritmo *K-Means*, Bezdek em 1973 [9] propôs a incorporação do conceito *fuzzy* na função SQIntra, resultando no *Fuzzy C-Means Clustering* (FCM). Neste algoritmo, inicialmente cada objeto pertence a todos os *clusters*, porém com distintos graus de pertinência. Após a convergência do algoritmo, cada objeto é alocado no *cluster* onde apresenta maior grau de pertinência, ou seja, a *defuzzificação* é pelo máximo. O FCM é uma generalização do método *K-Means*, com partição binária ou rígida ($u_{ij} = 1$ ou $u_{ij} = 0$), para um método onde a partição é *fuzzy* ($0 \leq u_{ij} \leq 1$).

No FCM a partição ótima resulta da solução do seguinte problema de otimização não linear com restrições de soma unitária para os graus de pertinência de cada objeto:

$$\begin{aligned} \text{Min } J &= \sum_{j=1}^k \sum_{i=1}^n u_{ij}^m \|x_i - c_j\|^2 \\ \text{s.a.} & \\ & \sum_{j=1}^k u_{1j} = 1 \\ & \dots \\ & \sum_{j=1}^k u_{nj} = 1 \end{aligned} \quad (13)$$

No problema de otimização em 13, m é uma constante maior que 1 e reflete a natureza *fuzzy* da partição. Para $m \rightarrow 1$ a partição tende a ser rígida ($u_{ij} \rightarrow 0$ ou $u_{ij} \rightarrow 1$) e para $m \rightarrow \infty$ a partição é totalmente *fuzzy* ($u_{ij} \rightarrow 1/k$). Usualmente m é 1,25 ou 2 [15]. Com o auxílio da função Lagrangeana o problema de otimização em 13 pode ser escrito como:

$$\text{Min } J = \sum_{j=1}^k \sum_{i=1}^n u_{ij}^m \|x_i - c_j\|^2 + \sum_{i=1}^n \lambda_i \left(\sum_{j=1}^k u_{ij} - 1 \right) \quad (14)$$

Na equação 14, λ_i , $i=1, n$ são os multiplicadores de Lagrange para as n restrições de igualdade.

Derivando a equação 14 em relação a u_{ij} , c_j e λ_j obtém-se as seguintes condições necessárias:

$$c_j = \frac{\sum_{i=1}^n u_{ij}^m x_i}{\sum_{i=1}^n u_{ij}^m} \quad (15)$$

$$u_{ij} = \frac{1}{\sum_{t=1}^k \left(\frac{\|x_i - c_j\|}{\|x_i - c_t\|} \right)^{\frac{2}{m-1}}} \quad (16)$$

Usando as equações 15 e 16 pode-se implementar o algoritmo do FCM, onde a matriz U e os centros dos k clusters são obtidos iterativamente:

Passo 1 - Inicialize a matriz U com valores entre 0 e 1, observando que em cada linha da matriz a soma dos valores deve ser unitária. Esta etapa pode ser vista como a *fuzzyficação* em um sistema de inferência *fuzzy*.

Passo 2 - Use a equação 15 para calcular as coordenadas dos k centros

Passo 3 - Calcule $J = \sum_{j=1}^k \sum_{i=1}^n u_{ij}^m \|x_i - c_j\|^2$. Pare se o valor de J estiver abaixo de uma tolerância ou se a melhoria em relação à iteração anterior for desprezível.

Passo 4 - Use a equação 16 para atualizar a matriz U e volte para o passo 2

Após a convergência do algoritmo, cada objeto é alocado no *cluster* onde apresenta o maior grau de pertinência. Da mesma forma que no *K-Means*, o algoritmo FCM não garante a convergência para a solução ótima e a performance também depende das coordenadas dos centros iniciais.

4.3. Método de Ward

O método de *Ward* é um método hierárquico e aglomerativo, significando que ele particiona um conjunto de n objetos sucessivamente em $n-1, n-2, \dots, 2$ clusters. Por exemplo, para um conjunto com 5 objetos a, b, c, d, e , as possíveis partições geradas por um método hierárquico aglomerativo são:

$$\begin{aligned} 5 \text{ clusters} &- a / b / c / d / e \\ 4 \text{ clusters} &- ab / c / d / e \\ 3 \text{ clusters} &- ab / cd / e \\ 2 \text{ clusters} &- ab / cde \end{aligned}$$

A SQIntra indica o grau de homogeneidade interna dos *clusters*: quanto menor for esta soma, mais homogêneas são as classes e melhor a qualidade da partição. A SQIntra cresce monotonicamente a medida que o número de classes diminui, indicando que as classes ficam maiores e cada vez mais heterogêneas. O exemplo anterior mostra que nos métodos hierárquicos, a cada iteração é feita a fusão de dois *clusters* e, por isso, são inevitáveis o acréscimo da SQIntra e a perda de qualidade da partição a cada iteração.

Para minimizar o acréscimo da SQIntra na passagem de k para $k-1$ clusters e não degenerar a qualidade da partição, o método de *Ward* agrupa os dois clusters, cuja agregação resulte no menor acréscimo da SQIntra. Este critério torna o método de *Ward* o mais eficiente dos métodos hierárquicos [6].

Para fazer a partição sucessiva dos n objetos em $n-1, n-2, \dots, 2$ clusters, os métodos hierárquicos necessitam de uma matriz simétrica de ordem n , onde cada elemento representa a distância entre dois objetos. À medida que os n objetos vão sendo agrupados, a ordem da matriz diminui e cada elemento passa a fornecer a distância entre os clusters. No exemplo dos objetos a, b, c, d, e a ordem da matriz de distância decai na sequência 5,4,3,2. A distância entre dois clusters a ser colocada na matriz de distâncias é dada por:

$$d_{ij} = \frac{p_i p_j}{p_i + p_j} \|c_i - c_j\|^2 \quad (17)$$

onde p_i e p_j são os pesos, ou melhor, o número de objetos nos *clusters* i e j respectivamente, e $\|c_i - c_j\|^2$ é o quadrado da distância euclidiana entre os centros dos *clusters* i e j .

5. SISTEMA COMPUTACIONAL ANATIPO

O sistema computacional ANATIPO possui uma interface gráfica amigável que permite ao usuário manipular com total facilidade os dados provenientes da importação de arquivos com medições de curva de carga, em formato público ou em planilha eletrônica. A seguir a Figura 6 ilustra uma tela com do ANATIPO com a lista das medições de consumidores residenciais consideradas no estudo.

The screenshot shows the 'Dados do Estudo de Clientes' window. It contains several input fields for search criteria: Código (005), ESTUDO SOBRE CLIENTES COMPLETO, Mercado (MWh) (1000000), Nível de Tensão (7 BT), Classe de Consumo (1 Residencial), and Faixa de Consumo (2 0 à 500 KW). Below these fields is a table with the following columns: Código, Nome, Ativo, and Arquivo.

Código	Nome	Ativo	Arquivo
005-00001-001	Consumidor Número 005-00001-001	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\527338hb.jod
005-00001-002	Consumidor Número 005-00001-002	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\861508ei.ams
005-00001-003	Consumidor Número 005-00001-003	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\861788gn.gde
005-00001-004	Consumidor Número 005-00001-004	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\861938gn.gtr
005-00001-005	Consumidor Número 005-00001-005	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\862478gn.gle
005-00001-006	Consumidor Número 005-00001-006	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\863138ei.fob
005-00001-007	Consumidor Número 005-00001-007	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\863238SC.AJL
005-00001-008	Consumidor Número 005-00001-008	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\863488GS.AGT
005-00001-009	Consumidor Número 005-00001-009	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\865188MF.AEP
005-00001-010	Consumidor Número 005-00001-010	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\865778PT.AAK
005-00001-011	Consumidor Número 005-00001-011	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\868108NH.ABJ
005-00001-012	Consumidor Número 005-00001-012	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_Publico\868878JC.AKT
005-00001-013	Consumidor Número 005-00001-013	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_EXCEL\CG_TRB1_034
005-00001-014	Consumidor Número 005-00001-014	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_EXCEL\CG_TRB1_040
005-00001-015	Consumidor Número 005-00001-015	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_EXCEL\CG_TRB4_009
005-00001-016	Consumidor Número 005-00001-016	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_EXCEL\CG_TRB4_010
005-00001-017	Consumidor Número 005-00001-017	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_EXCEL\CG_TRB4_011
005-00001-018	Consumidor Número 005-00001-018	Sim	C:\CEPEL\AnaTpo\INPUT\Formato_EXCEL\CG_TRB4_014

FIGURA 6 – LISTA DE ARQUIVOS COM MEDIÇÕES DE CURVAS DE CARGA

No sistema ANATIPO, o processo de seleção das curvas características ocorre em um ambiente gráfico amigável, onde é possível visualizar simultaneamente todas as curvas, classificadas por dia da semana, registradas em um arquivo de medição (Figura 7). Nesta saída gráfica pode-se comparar facilmente o perfil de carga em diferentes dias da semana e selecionar diretamente no gráfico as curvas do dia útil, sábado e domingo que melhor caracterizem a medição. Apenas as curvas características são armazenadas na base de dados e podem ser visualizadas a qualquer momento (Figura 8).

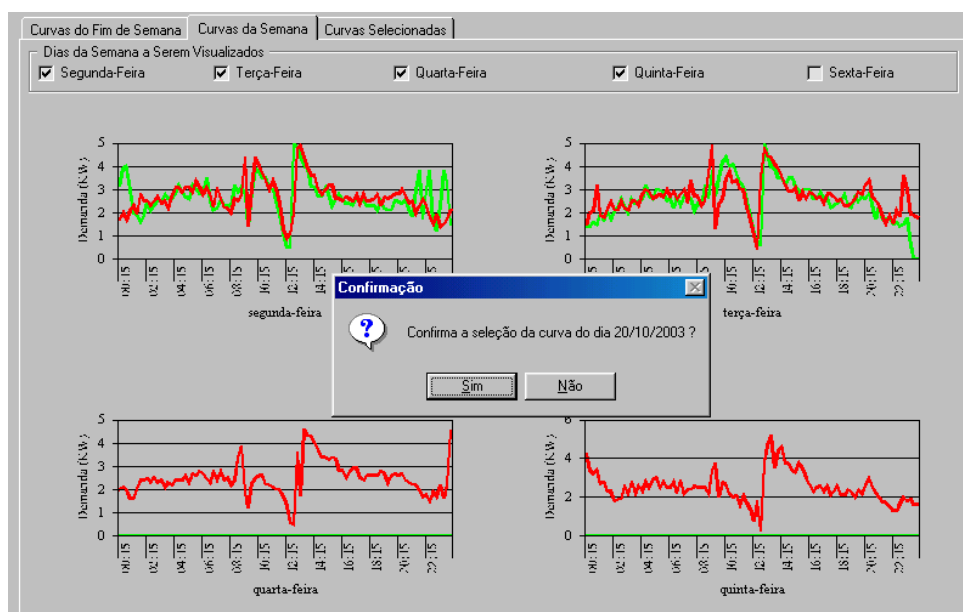


FIGURA 7 – TELA PARA VISUALIZAÇÃO DE UMA MEDIÇÃO E SELEÇÃO MANUAL DAS RESPECTIVAS CURVAS CARACTERÍSTICAS

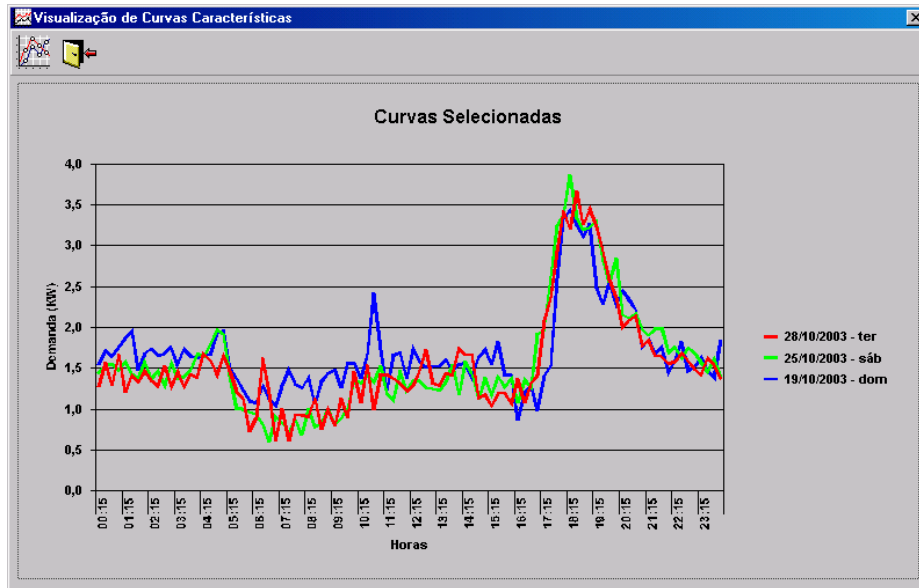


FIGURA 8 – CURVAS CARACTERÍSTICAS SELECIONADAS

Os resultados, obtidos por qualquer um dos três métodos de *cluster analysis*, são apresentados em uma saída gráfica que permite visualizar até quatro tipologias simultaneamente, conforme mostra a Figura 9.

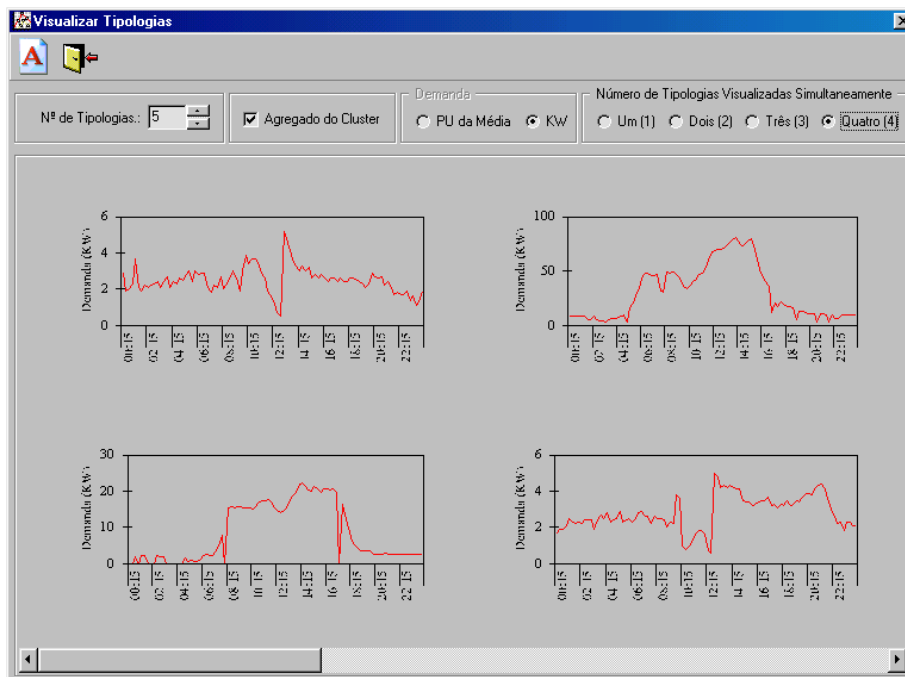


FIGURA 9 – TIPOLOGIAS

O ANATIPO também faz o ajuste das tipologias ao mercado. Este processo é realizado de forma automática, cabendo ao usuário apenas fornecer o mercado (MWh) do segmento a ser caracterizado pelas tipologias. Como resultado o ANATIPO gera uma planilha eletrônica (Figura 10) com as tipologias para os dias útil, sábado e domingo, ajustadas ao mercado e prontas para serem anexadas às propostas tarifárias da distribuidora.

INTERVALO	HORA	TIPO 1			TIPO 2			AGREGADO		
		útil	sábado	domingo	útil	sábado	domingo	útil	sábado	domingo
00:00-01:00	1	15.253	18.113	15.600	0.953	0.867	0.780	16.207	18.980	16.38
01:00-02:00	2	14.907	18.113	16.120	0.867	0.953	0.693	15.773	19.067	16.813
02:00-03:00	3	15.167	17.940	15.947	0.780	0.780	0.780	15.947	18.720	16.727
03:00-04:00	4	15.167	17.680	16.033	0.867	0.867	0.693	16.033	18.547	16.727
04:00-05:00	5	15.253	17.767	15.600	0.693	0.953	0.780	15.947	18.720	16.38
05:00-06:00	6	18.287	19.674	16.033	0.867	0.867	0.693	19.154	20.540	16.727
06:00-07:00	7	32.760	35.707	16.380	1.040	1.127	0.953	33.800	36.834	17.333
07:00-08:00	8	44.460	63.787	16.293	1.300	1.733	0.953	45.760	65.521	17.247
08:00-09:00	9	142.568	123.588	16.207	1.820	1.387	0.693	144.388	124.974	16.9
09:00-10:00	10	306.976	227.935	15.773	6.067	1.387	1.127	313.043	229.322	16.9
10:00-11:00	11	317.896	244.055	16.293	6.413	1.213	0.693	324.309	245.269	16.987
11:00-12:00	12	287.042	234.522	18.200	1.560	1.127	0.780	288.603	235.649	18.98
12:00-13:00	13	296.749	225.855	17.680	4.680	1.040	0.607	301.429	226.895	18.287
13:00-14:00	14	219.962	184.428	17.073	0.780	2.253	0.347	220.742	186.682	17.42
14:00-15:00	15	245.442	204.275	16.293	6.153	1.560	0.867	251.596	205.835	17.16
15:00-16:00	16	250.989	202.455	17.247	5.893	1.040	0.520	256.882	203.495	17.767
16:00-17:00	17	232.875	209.475	16.467	5.807	0.693	0.520	238.682	210.169	16.987
17:00-18:00	18	228.109	215.282	16.900	0.780	1.560	1.300	228.889	216.842	18.2
18:00-19:00	19	178.622	131.474	16.380	1.127	0.867	0.780	179.748	132.341	17.16
19:00-20:00	20	117.868	94.901	17.247	1.907	0.953	0.867	119.774	95.854	18.113
20:00-21:00	21	111.974	47.927	17.593	1.387	1.647	0.953	113.361	49.574	18.547
21:00-22:00	22	79.041	32.414	17.333	1.127	1.387	0.347	80.167	33.800	17.68
22:00-23:00	23	44.114	25.827	16.900	0.780	1.213	0.433	44.894	27.040	17.333
23:00-24:00	24	23.054	17.680	16.033	1.473	1.040	0.173	24.527	18.720	16.207
% mercado		98.347			1.653			100		
mercado MWh		983472.4			16527.56			1000000		
demanda média MW		135.606			2.297			137.902		
poderação sábado		0.808			0.517			0.804		
poderação domingo		0.122			0.314			0.125		
fator de carga %		42.657			35.811			42.522		
demanda máx MW Ponta		178.622			1.907			179.748		
demanda máx MW Fora_Ponta		317.896			6.413			324.309		

FIGURA 10 – PLANILHA COM TIPOLOGIAS AJUSTADAS AO MERCADO

6. CONCLUSÕES

Neste artigo foi apresentado o ANATIPO, um sistema computacional desenvolvido pelo CEPEL em conjunto com a COSERN, por meio de um projeto de P&D ANEEL, com o objetivo de agilizar os estudos de caracterização da carga, beneficiando particularmente os estudos tarifários. O sistema é dotado de interface gráfica amigável e conta com uma série de funcionalidades que tornam mais confortável a construção das tipologias de curvas de carga, desde a aquisição das medições, passando pela seleção das curvas características até a agregação e ajuste das tipologias ao mercado.

Entre as diversas funcionalidades do programa destacam-se as rotinas para leitura de arquivos de medição em formato público e planilha eletrônica, o que confere uma grande flexibilidade no tratamento dos dados provenientes das campanhas de medidas. Outras funcionalidades importantes, apresentadas resumidamente neste trabalho, são as rotinas gráficas para identificação manual e automática das curvas características de cada medição, as rotinas gráficas para visualização dos *clusters* de curvas de carga, e os três métodos de análise de agrupamentos (*K-Means*, *Ward* e *FCM*) implementados no programa.

Enfim, estas funcionalidades proporcionam maior agilidade na construção das tipologias de curvas, e com certeza contribuem também para a elaboração das propostas tarifárias.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] BEBER, S. DANIEL *Cronograma de tareas para realizar el estudio de un cuadro tarifario*. II Reunión Internacional de Distribución y Comercialización de Energía Eléctrica, Lima, Peru, agosto de 2001
- [2] CHEN, C.S., HWANG, J.C., HUANG, C.W. *Application of Load Survey Systems to Proper Tariff Design*. IEEE Transactions on Power Systems, Vol. 12, nº 4, November 1997
- [3] JANNUZI, GILBERTO M., SWISHER, JOEL N.P. *Planejamento Integrado de Recursos Energéticos*. Campinas, São Paulo, Editora Autores Associados, 1997

- [4] HWANG, J.C. *Assesment of Air Condition Load management by load Survey in Taipower*. IEEE Transactions on Power Systems, Vol, 16 no 4, November 2001
- [5] PESSANHA, J.F.M., VELASQUEZ, R.M.G., PASSOS JR., R., MELO, A C.G., CALDAS, R.P., TANURE, J.E.P.S., SANTOS, P.E.S. *Metodologia e Aplicação do Cálculo dos Custos Marginais de Fornecimento e das Tarifas de Uso dos Sistemas de Distribuição*, Latin America Power & Gas, Rio de Janeiro, 2001
- [6] JOHNSON, RICHARD A., WINCHERN, DEAN W. *Applied Multivariate Statistical Analysis*, 4th edition, Prentice-Hall, 1998
- [7] BOUROCHE, JEAN-MARIE, SAPORTA, GILBERT *Análise de Dados*, Zahar Editores, Rio de Janeiro, 1980
- [8] ALDENDERFER, M.S., BLASHFIELD, R.K. *Cluster Analysis*, Series: Quantitative Applications in the Social Sciences, SAGE publications
- [9] JANG, J.S.R., SUN, C.T., MIZUTANI, E. *Neuro-Fuzzy and Soft Computing: a computational approach to learning and machine intelligence*, Prentice Hall Inc, 1997
- [10] BRASIL, Ministério das Minas e Energia, DNAEE, Eletrobrás, Empresas Concessionárias de Energia Elétrica; *Nova Tarifa de Energia Elétrica : metodologia e aplicação* , DNAEE, Brasília, 1985
- [11] BRASIL, Ministério das Minas e Energia, Secretaria de Energia, DNAEE; *Programa de Revisão Tarifária - PRT, Projeto 1 - Caracterização da Carga*, 1994.
- [12] MURTEIRA, BENTO J.F. *Análise Exploratória de Dados – Estatística Descritiva*, McGraw Hill, Lisboa, 1993
- [13] JAIN, A.K., DUIN, R.P.W., MAO, J. *Statistical Pattern Recognition: A Review*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, nº 1, January 2000.
- [14] PESSANHA, J.F.M., VELASQUEZ, R.M.G., MELO, A.C.G., CALDAS, R.P. *Técnicas de Cluster Analysis na Construção de Tipologias de Curvas de Carga*, XV Seminário Nacional de Distribuição de Energia Elétrica - SENDI, Salvador, 2002
- [15]** FRANCO, CLÁUDIA R. *Categorização Nebulosa e Validação de Categorias Aplicadas a Problemas de Reconhecimento de Padrões*, Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, IM/NCE, 2002.