

Previsão de Carga a Curto Prazo Usando Método de Agrupamento Nebuloso

Ricardo Menezes Salgado, Takaaki Ohishi e Rosângela Ballini

Resumo—O presente trabalho aplica um modelo baseado em agrupamento nebuloso no problema de previsão de carga de curto prazo para as 24 horas do próximo dia. Neste modelo, é utilizado o algoritmo de agrupamento *fuzzy c-means* para explorar a estrutura dos dados históricos, e reconhecimento de padrões para capturar similaridades na tendência das séries de consumo de carga discretizadas em base horária. A metodologia é testada em um ponto de medição tipicamente residencial localizado na região nordeste do Brasil. Os resultados mostram erros percentuais absolutos médios na ordem de 1,99%.

Index Terms—Previsão de Carga, Séries Temporais, Algoritmos de Agrupamento, Fuzzy C-Means, Reconhecimento de Padrão.

I. INTRODUÇÃO

Atualmente, em um sistema de potência o controle de carga é exigência básica nos processos de planejamento e operação. Para fornecer energia elétrica com boa qualidade, de maneira segura e econômica, a empresa precisa dispor de mecanismos que possibilitem a resolução de vários problemas de nível técnico e operacional. Especificamente, no contexto do planejamento de curto prazo, a previsão da carga é importante na elaboração do programa de operação dos próximos dias em análises de segurança e estabilidade, pois erros na previsão podem afetar a eficiência e segurança do sistema. Em concordância com este objetivo, o conhecimento do comportamento da carga futura é o primeiro pré-requisito para um planejamento seguro e confiável do sistema de energia elétrica. Aumentar a segurança e a economia da operação de sistemas são as motivações para a realização de previsões de cargas mais confiáveis nos sistemas elétricos de potência.

O problema da previsão de carga é classificado como previsão de longo prazo, médio prazo e curto prazo, dependendo da área de atuação. A previsão de longo prazo é necessária para o planejamento da expansão da capacidade do sistema; a previsão de médio prazo é necessária para a programação do suprimento de combustível, operações de manutenção e planejamento do intercâmbio; e a previsão de curto prazo é utilizada na operação diária do sistema elétrico para a programação da operação, transferência de energia e gerenciamento da demanda [1].

O presente trabalho enfoca a previsão da demanda de carga de curto prazo, sendo que o grande desafio é manter o equilíbrio entre a energia gerada e a consumida. Um desequilíbrio, tanto de escassez como de excesso de geração de carga, pode causar problemas de subtensão ou de sobretensão no sistema. As

soluções possíveis para equilibrar essa escassez ou excesso de energia são as reprogramações das geradoras, ou as transferências de energias através dos intercâmbios entre os sistemas. Entretanto, tanto a reprogramação, quanto o intercâmbio, demandam um certo intervalo de tempo, daí a importância de se realizar uma boa previsão de carga com antecedência suficiente.

As redes neurais artificiais (RNAs) e variantes têm sido largamente adotadas para previsão de séries temporais que apresentam um comportamento não-linear [2]. A ideia de utilizar RNAs em previsão não é nova, sendo que a primeira aplicação data de 1964, realizada por [3]. A utilização de redes neurais para a previsão de séries temporais foi proposta por vários autores, mostrando a viabilidade de utilizar estes modelos [4], [5]. Diversas estruturas de RNAs vêm sendo propostas baseadas principalmente nos modelos perceptron multi-camadas, mapas de Hopfield e redes auto-organizáveis de Kohonen. Em problemas de previsão, os modelos mais utilizados são redes estáticas multi-camadas, treinadas através do algoritmo de retropropagação, ou *backpropagation* [6], cuja base é o método do gradiente descendente [7].

Recentemente, sistemas nebulosos vêm sendo combinados às RNAs com o objetivo de explorar a capacidade de representação e o poder de aprendizagem dessas abordagens, obtendo bons resultados [8].

Com o objetivo de explorar a estrutura de dados históricos da série temporal, o agrupamento de dados, reconhecido como um dos paradigmas mais importantes para o reconhecimento e classificação de padrões, é empregado em problemas de previsão.

Este artigo utiliza um modelo baseado em agrupamento nebuloso o qual utiliza procedimentos de reconhecimento de padrões para capturar informações relevantes na série histórica e obter a previsão. Este modelo foi aplicado para previsão de vazões afluentes em [9]. Este tipo de abordagem proporciona menor complexidade e menor tempo de execução do que os métodos baseados em redes neurais e redes neurais nebulosas.

Neste artigo, este modelo é aplicado à previsão de séries temporais de carga ativa com discretização horária (MW/h), medidas em um ponto de medição localizado no nordeste do Brasil. Os resultados obtidos são comparados com o modelo de rede neural multi-camadas.

II. AGRUPAMENTO DE DADOS EM PREVISÃO

Técnicas de agrupamento são empregadas para análise exploratória de dados, auxiliando na resolução de problemas de agrupamento. O objetivo do processo de agrupamento é separar um conjunto de dados em diferentes grupos de modo que os elementos de um dado grupo apresentem um certo grau de similaridade entre eles. O uso dessa técnica é apropriado quando se conhece pouco ou nada sobre a estrutura de um conjunto de

R. M. Salgado - Departamento de Engenharia de Sistemas, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, e-mail: ricardo@cose.fee.unicamp.br

T. Ohishi - Departamento de Engenharia de Sistemas, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, e-mail: taka@densis.fee.unicamp.br

R. Ballini - Departamento de Teoria Econômica - DTE, Instituto de Economia - IE, Universidade Estadual de Campinas, Campinas, e-mail: ballini@eco.unicamp.br

dados [10].

A idéia de aplicar algoritmos de agrupamento em previsão consiste em organizar os elementos da série em grupos com características similares. Apesar do agrupamento de dados ser um conceito comum em modelos estatísticos, existem também algoritmos de agrupamento baseados na teoria de conjuntos nebulosos. Esses algoritmos constituem generalizações dos algoritmos clássicos e são utilizados, nesse trabalho, como base para o desenvolvimento de modelos destinados à previsão.

Em previsão de séries temporais, os algoritmos de agrupamento são utilizados como componentes de metodologias híbridas. Geralmente, essas metodologias híbridas são desenvolvidas em duas fases. Na primeira, através dos algoritmos de agrupamento, o objetivo é organizar os dados em grupos com características similares, estabelecendo padrões e capturando comportamentos semelhantes ao longo da série. Supondo que os dados estejam classificados, a segunda fase consiste em aplicar métodos de previsão a cada uma das diferentes classes pré-estabelecidas. A principal diferença entre os modelos existentes reside na escolha do método de agrupamento [11], [12].

Neste trabalho, é utilizado o algoritmo de agrupamento *Fuzzy C-Means - FCM*, proposto por [13], para realizar o agrupamento dos dados.

A. Algoritmo Fuzzy C-Means

O algoritmo FCM, apresentado no Quadro 1, utiliza os seguintes conceitos e notações. Seja $X = \{x_1, x_2, \dots, x_p\}$ um conjunto finito de dados, $X \in \mathbb{R}^p$, onde \mathbb{R}^p é um espaço Euclidiano p -dimensional; c o número de grupos, $2 \leq c \leq n$; U a matriz de pertinência $c \times n$, sendo u_{ik} com $1 \leq i \leq c$ e $1 \leq k \leq n$, o grau de pertinência do ponto x_k ao grupo i . Uma c -partição de X é definida por:

$$U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \dots & \dots & \dots & \dots \\ u_{c1} & u_{c2} & \dots & u_{cn} \end{bmatrix} \quad (1)$$

Sendo $u_{ik} \in [0, 1], \forall i, k$ e

$$\sum_{i=1}^c u_{ik} = 1; 0 < \sum_{k=1}^n u_{ik} < n, \quad (2)$$

O algoritmo FCM agrupa os dados através de um procedimento que tem como objetivo minimizar a seguinte função objetivo:

$$J_m(U, V) = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m d_{ik}^2, U \in M_{fc}, 1 < m < \infty \quad (3)$$

a qual $d = \|x_k - v_i\|_A$ é a distância entre x_k e v_i , $\|\cdot\|$ é uma norma induzida por um produto interno, como por exemplo $\|x\|_A^2 = x^T A x$, sendo A uma matriz $p \times p$ definida positiva; $V = \{v_1, v_2, \dots, v_c\}$ é o conjunto de centros do grupo, representado por uma matriz $c \times p$, $v_i \in \mathbb{R}^p, 1 \leq i \leq c$, o ponto v_i é chamado de centro do i -ésimo grupo; e m é o fator que define o grau de pertinência da partição *fuzzy* do sistema.

Quadro 1: Algoritmo FCM.

Dado X , escolher o número de grupos $2 \leq c \leq n$, o parâmetro $m > 1$, os critérios de parada $\epsilon > 0$ e o número máximo de iterações $lmax$.

1. \rightarrow Inicializar $U^{(0)}$ e o contador de iterações $l = 1$.
2. \rightarrow Calcular os c centros dos grupos $\{v_1^l, v_2^l, \dots, v_n^l\}$ usando $U^{(l)}$, com a seguinte equação:

$$v_i^l = \frac{\left(\sum_{k=1}^n (u_{ik}^{l-1})^m x_k \right)}{\left(\sum_{k=1}^n (u_{ik}^{l-1})^m \right)}$$

3. \rightarrow Atualizar $U^{(l)}$ com o seguinte procedimento:
Para $1 \leq k \leq n$:

Se $\|x_k - v_i^l\|^2 > 0$

$$u_{ik} = \left[\sum_{j=1}^c \left(\frac{\|x_k - v_i^l\|^2}{\|x_k - v_j^l\|^2} \right)^{\frac{1}{m-1}} \right]^{-1}, i = 1, 2, \dots, c.$$

Se $\|x_k - v_i^l\|^2 = 0$

$u_{ik}^l = 1, 1 < i < c$.

4. \rightarrow Calcular $\Delta = \|U^l - U^{l-1}\| = \max_{i,j} |u_{ij}^l - u_{ij}^{l-1}|$
Se $\Delta > \epsilon$ ou $l < lmax$,
 $l = l + 1$ e voltar ao passo 2.
Caso contrário parar.

III. MODELO DE PREVISÃO BASEADO EM AGRUPAMENTO NEBULOSO

O modelo de previsão baseado em agrupamento nebuloso (PAN) é composto por duas fases. A primeira é referente ao agrupamento de padrões e a segunda a classificação de novos padrões e previsão [9] e [14].

Neste trabalho, considera-se a previsão um passo à frente, isto é, dadas $p_{t-1} \in \mathbb{R}, 1 < t < k$ amostras de uma série temporal, o objetivo é estimar o valor de p_t utilizando informação de um conjunto de valores passados.

Seja um padrão genérico p_i , $(k + 1)$ dimensional construído pela união de p valores amostrais de uma série temporal da seguinte maneira:

$$p_i = [c_{t-k}^i \quad c_{t-(k+1)}^i \quad \dots \quad c_{t-2}^i \quad c_{t-1}^i \quad c_t^i] \quad (4)$$

sendo $1 < i < N$ o número de padrões.

Os elementos c_{t-k}^i são padrões de carga escolhidos para proporcionar uma previsão de melhor qualidade. No caso de previsão de carga horária à curto prazo analisou-se o desempenho de $k = 2$ e $k = 4$ para formar os padrões de previsão. Desta maneira, os padrões seguem o seguinte modelo:

$$p_i = [c_{t-4}^i \quad c_{t-3}^i \quad c_{t-2}^i \quad c_{t-1}^i \quad c_t^i] \quad (5)$$

O algoritmo FCM, descrito anteriormente, é aplicado na primeira fase do modelo aos padrões de dados definidos pela Equação (5). Como resultado, obtém-se a matriz de centros dos grupos dada por:

$$Mc = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1(k+1)} \\ c_{21} & c_{22} & \dots & c_{2(k+1)} \\ \vdots & \vdots & \ddots & \vdots \\ c_{j1} & c_{j2} & \dots & c_{j(k+1)} \end{pmatrix} = \begin{pmatrix} mc_1 \\ mc_2 \\ \vdots \\ mc_j \end{pmatrix}$$

sendo $c_i \in \mathfrak{R}^{k+1}$. O número apropriado de centros j é escolhido empiricamente. Esses centros serão utilizados no próximo estágio, para a classificação de novos padrões de entrada.

A segunda fase classifica novos padrões p_i , denominados padrões de previsão, de acordo com os grupos nebulosos obtidos na primeira fase. O padrão de previsão tem como último componente, o valor a ser previsto. Os demais p elementos compõem o padrão de previsão de forma similar à composição dos padrões de dados definidos na Equação (4).

A classificação dos padrões de previsão é feita utilizando a seguinte equação do algoritmo FCM:

$$u_{wp} = \left[\sum_{j=1}^c \left(\frac{\|p_p - mc_w\|^2}{\|p_p - mc_j\|^2} \right)^{\frac{1}{m-1}} \right]^{-1} \quad (6)$$

sendo u_{wp} o grau de pertinência do padrão de previsão p_p ao grupo w . Assim, o valor da previsão é computado como segue:

$$\hat{c}_t^p = \sum_{j=1}^c u_{wp} \times mc_{w(k+1)} \quad (7)$$

Note que como o último componente do padrão de previsão p_p é o valor a ser previsto. Assim para a etapa de previsão deve-se utilizar um mecanismo que substitua este componente para que se possa realizar a classificação dos padrões. Isso significa que o valor de c_t^p em p_p deve ser substituído por uma aproximação adequada. Neste trabalho, utiliza-se o procedimento de reconhecimento de padrões descrito a seguir.

A. Técnica de Reconhecimento de Padrão - PANrp

A idéia principal deste procedimento é incluir informações sobre a tendência do padrão de previsão para selecionar o valor que melhor aproxima o valor real c_t^p .

A aproximação desejada é obtida pelo valor de c_j^i do padrão de dados p_i que mais se aproxima do padrão de previsão p_p . O procedimento de reconhecimento de padrões (PANrp) compara os N valores dos padrões de dados com um padrão de previsão, utilizando a norma Euclidiana como medida de distância. Logo, o valor c_t^i do padrão de dados, cuja norma é menor, fornece a melhor aproximação para \hat{c}_t^p . Matematicamente, o padrão de dados mais próximo de p_p , é encontrado de acordo com a seguinte expressão:

$$\|p_p - p_d\| = \min_{1 < i < N} \|p_p - p_i\| \quad (8)$$

Desta maneira, o valor \hat{c}_t^p é o valor c_t^d pertencente ao padrão de dados p_d . O modelo PANrp segue os passos descritos no Quadro 2.

Quadro 2: Algoritmo - PANrp.

1. Escolher o número de componentes k referentes aos padrões de dados.
2. Construir os padrões de dados p_i e o padrão de previsão p_p .
3. Determinar o número de centros j e aplicar o algoritmo FCM para obter a matriz de centros

$$Mc = [c_i \quad | \quad c_j]$$

4. Determinar o padrão p_d de acordo com a seguinte expressão:

$$\|p_p - p_d\| = \min_{1 < i < N} \|p_p - p_i\|$$

5. Tomar o valor \hat{c}_t^p como sendo igual ao valor c_t^d pertencente ao padrão de dados p_d .
6. Através da matriz Mc classificar o padrão de previsão de acordo com a equação:

$$u_{wd} = \left[\sum_{j=1}^c \left(\frac{\|p_d - mc_w\|^2}{\|p_d - mc_j\|^2} \right)^{\frac{1}{m-1}} \right]^{-1}$$

7. Encontrar o valor previsto através da seguinte equação:

$$\hat{c}_t^p = \sum_{j=1}^c u_{wd} \times mc_{w(k+1)}$$

IV. ANÁLISE DOS DADOS DE CARGA

Análise de séries temporais tem como objetivo básico à avaliação do fenômeno sob consideração, para se obter conclusões em termos estatísticos, bem como, sumarizar propriedades estatísticas e caracterizar seu comportamento [15].

Conhecer o conjunto de dados que será utilizado na fase de ajuste e validação dos modelos de previsão é fundamental em qualquer etapa do desenvolvimento de uma pesquisa. A análise prévia dos dados pode fornecer informações relevantes que serão levadas em consideração em etapas futuras no ajuste de um modelo de previsão.

Neste trabalho utilizou-se uma série diária de consumo de energia pertencente ao sistema elétrico da região nordeste do Brasil. As medições que foram feitas no período de 01/06/2001 a 04/10/2001 (período do racionamento de energia).

A Tabela I mostra a tensão, a carga máxima e mínima, o valor médio, o desvio padrão e a variância da série considerada. A comparação do desvio padrão com o valor médio da carga mostra o comportamento com pouca variabilidade da série analisada.

A Figura 1 mostra os perfis de consumo da série analisada. Pode-se observar que o consumo, para este ponto de medição, apresenta grande regularidade com demanda de ponta ocorrendo no período das 17 as 22h, indicando um comportamento de consumo tipicamente residencial.

O fato das curvas possuírem níveis distintos de consumo alimenta a necessidade da realização de previsores que sejam capazes de perceber tais distinções, principalmente entre dias úteis e fins de semana, e proporcionar resultados mais acurados.

TABELA I
DADOS ESTATÍSTICOS.

Tensão	Carga Máxima	Carga Mínima	Carga Média	Desvio Padrão	Variância
69KV	64,5 MW	21,2 MW	35,2 MW	2,36 MW	6,10 MW

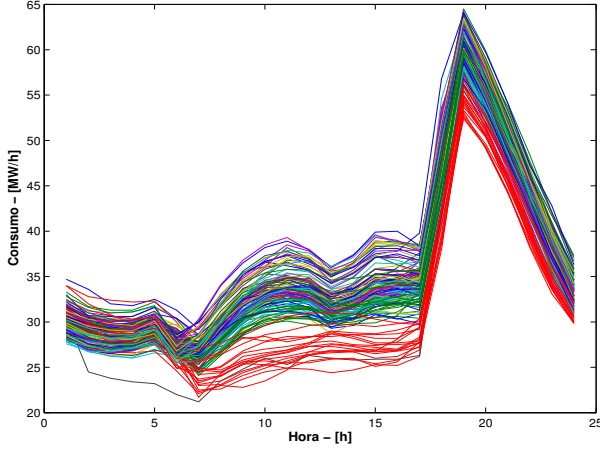


Figura 1. Dados de Consumo.

V. RESULTADOS

Neste trabalho, o desempenho do modelo de agrupamento nebuloso foi comparado com o desempenho do modelo de rede neural multi-camadas com algoritmo de retropropagação.

Para o ajuste dos modelos utilizou-se o histórico de dados do período de 01/06/20001 até 25/09/2001 e para a fase de validação utilizou-se o dia de 02/10/2001. Para avaliar e comparar o desempenho dos modelos, as medidas adotadas foram o erro relativo percentual médio ($ERM(\%)$), o erro absoluto médio (EAM) e o erro absoluto máximo (EA_{max}), entre as cargas observadas y_j e as cargas previstas \hat{y}_j , definidos como:

$$ERM(\%) = \frac{100}{24} \sum_{j=1}^{24} \frac{|y_j - \hat{y}_j|}{y_j} \quad (9)$$

$$EAM = \frac{1}{24} \sum_{j=1}^{24} |y_j - \hat{y}_j| \quad (10)$$

$$EA_{max} = \max |y_j - \hat{y}_j| \quad (11)$$

A estrutura do modelo de RNA é estática com uma camada intermediária, composta de 10 neurônios. A taxa fixa de aprendizagem foi de 0,001 e o termo *momentum* igual a 0,9.

O desempenho dos modelos para a previsão hora-a-hora, foi analisado considerando dois tipos de padrão de entrada: um primeiro padrão constituído por 4 componentes, ou seja, 4 dias anteriores ao dia a ser previsto para uma determinada hora h ; o segundo padrão é composto por 2 dias anteriores ao dia a ser previsto, na hora h . Por exemplo, para se prever a primeira hora da terça-feira dia 02/10/2001 os modelos foram considerados

utilizando as seguintes padrões como entradas:

$$p_i = \left[\underbrace{c_{t-4}^i}_{Sex} \quad \underbrace{c_{t-3}^i}_{Sab} \quad \underbrace{c_{t-2}^i}_{Dom} \quad \underbrace{c_{t-1}^i}_{Seg} \quad \vdots \quad \underbrace{\tilde{c}_t^i}_{Ter} \right]$$

Padrão com 4 componentes

$$p_i = \left[\underbrace{c_{t-2}^i}_{Dom} \quad \underbrace{c_{t-1}^i}_{Seg} \quad \vdots \quad \underbrace{\tilde{c}_t^i}_{Ter} \right]$$

Padrão com 2 componentes

A seguir são analisados os resultados para estes dois casos.

A. Padrão com 4 componentes

A Tabela II mostra os erros $ERM(\%)$, EAM e EA_{max} para os dois modelos de previsão. Como pode ser observado o modelo PANrp apresentou melhor desempenho mostrando que a metodologia foi eficaz na obtenção das previsões.

TABELA II
ERROS TERÇA-FEIRA 02/10/2001 - PADRÃO COM 4 COMPONENTES.

	$ERM(\%)$	EAM (MW)	EA_{max} (MW)
PANrp	1,99	0,75	1,91
RNA	2,42	0,89	2,14

A Figura 2 mostra as curvas estimadas pelos modelos bem como a curva de consumo verificada na terça-feira (02/10/2001). Nota-se que modelo de RNA apresentou a curva menos similar a curva verificada.

A Figura 3 apresenta os erros percentuais ao longo das 24h do dia. Nota-se que os erros para ambos os modelos são mais altos durante o período das 7 as 15h, Figura 3, sendo este período o de maior variabilidade nos perfis de consumo, como pode ser observado na Figura 1. Também pode-se notar que o modelo PANrp apresentou menor erro percentual para todo o período.

B. Padrão com 2 componentes

A Tabela III mostra o $ERM(\%)$, EAM e EA_{max} para a previsão feita com 2 componentes, ou seja, dois dados anteriores ao dia de previsão hora-a-hora. Neste caso, a RNA apresentou um melhor desempenho que o modelo PANrp. Isto deve-se ao fato de que o número de entradas foi insuficiente para o modelo PANrp generalizar o comportamento dos padrões ocasionando perda de qualidade nas previsões.

A Figura 4 mostra as curvas estimadas pelos modelos bem como a curva de consumo verificada na terça-feira 02/10/2001.

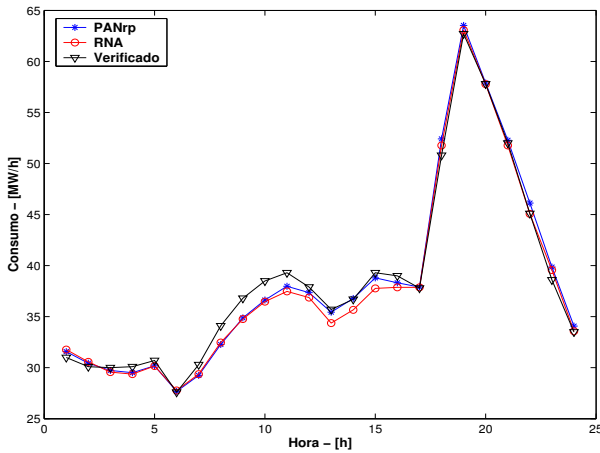


Figura 2. Curvas Estimadas 02/10/2001 - Padrão com 4 Componentes.

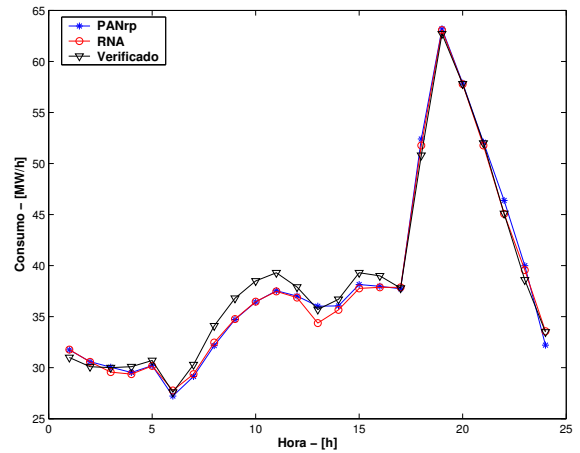


Figura 4. Curvas Estimadas 02/10/2001 - Padrão com 2 Componentes.

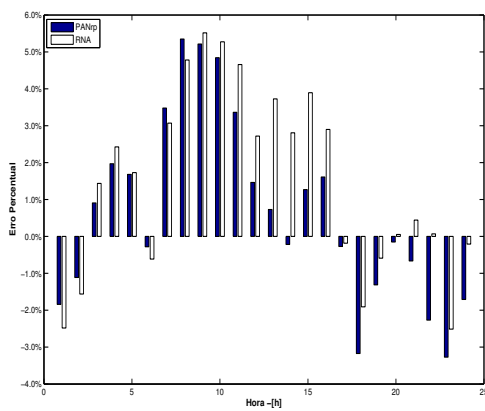


Figura 3. Erro Percentual 02/10/2001 (Base Horária) - Padrão com 4 Componentes.

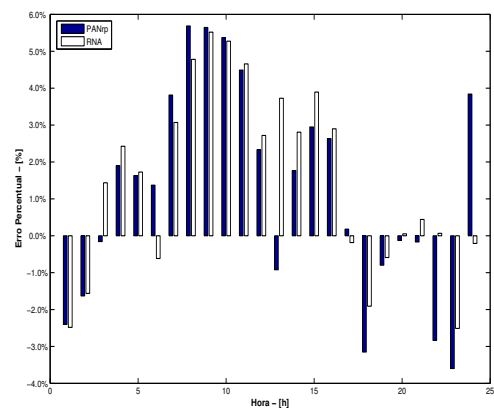


Figura 5. Erro Percentual 02/10/2001 (Base Horária) - Padrão com 2 Componentes.

Como descrito acima, o modelo RNA apresentou a curva mais similar a curva desejada. Nota-se também que o período das 9hs às 13h foi o de maior variabilidade nos erros, para ambos os modelos.

A Figura 5 apresenta o erro percentual calculado hora-a-hora para a terça-feira (02/10/2001). Como pode-se notar, ambos os modelos apresentaram dificuldades em prever o valor de carga às 12h, que é um ponto crítico na curva de consumo.

Vale ressaltar que o esforço computacional e o tempo de execução do modelo PANrp é menor em comparação ao modelo RNA. Isso acontece porque os procedimentos utilizados no modelo PANrp, não exigem um grande esforço computacional e são executados rapidamente. Contrariamente, uma RNA

com um número grande de neurônios na camada intermediária demanda um tempo de execução e um esforço computacional relativamente altos devido às formulações matemáticas envolvidas no processo de treinamento.

VI. CONCLUSÕES

Neste estudo foram comparados o desempenho de dois modelos de previsão de carga a curto prazo: modelo baseado em agrupamento nebuloso e reconhecimento de padrões (PANrp) e uma rede neural multi-camadas (RNA).

O modelo PANrp apresentou um bom desempenho, com erros relativo percentual médio e absoluto médio menores que o modelo RNA, quando um número maior de informações foram dadas ao modelo PANrp.

Uma ferramenta de previsão deve apresentar erros compatíveis com a reserva de carga do sistema. O erro máximo é um aspecto importante na operação em tempo real, o sistema elétrico pode não estar preparado para suprir uma dada demanda de carga muito maior do que a prevista. Um aspecto a destacar é que tanto o modelo PANrp como a RNA obtiveram erros máximos menores que 3,0% da carga da ponta. Este valor está de acordo com a faixa de segurança que usualmente é adotado

TABELA III
ERROS TERÇA-FEIRA 02/10/2001 - PADRÃO COM 2 COMPONENTES.

	ERM (%)	EAM (MW)	EAm_{max} (MW)
PANrp	2,47	0,91	2,07
RNA	2,31	0,84	2,03

nas empresas elétricas.

Em termos de custo computacional o modelo PANrp mostrou-se mais eficiente que o modelo neural. Isto pode ser verificado pelo fato do PANrp ter como base o algoritmo *fuzzy c-means* que apresenta baixo custo computacional; por outro lado, o modelo RNA tem um custo computacional bastante elevado devido às conexões existentes entre os neurônios.

O processo de agrupamento mostrou-se eficiente no tratamento de previsão de carga à curto prazo. Desta forma, modelos de agrupamento, como PANrp, se despontam como uma alternativa com grande capacidade de representação de dados e previsão de séries temporais de consumo elétrico.

AGRADECIMENTOS

Esta pesquisa teve o suporte da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) processos #04/07879 – 9 e #03/10019 – 9.

REFERÊNCIAS

- [1] S. Rahman, "Formulation and analysis of a rule-based short-term load forecasting algorithm," *Proceedings of IEEE*, vol. 78, no. 5, pp. 805–816, 1990.
- [2] N. Gershenfeld and A. Weigend, "The future of time series: learning and understanding," in *Time Series Prediction: Forecasting the Future and Understanding the Past*. Addison-Wesley Publishing Company, 1996, pp. 1–70.
- [3] Q. Zhang, B. Patuwo, and M. Hu, "Forecasting with artificial neural networks: the state of the art," *International Journal of Forecasting*, vol. 14, pp. 35–62, 1998.
- [4] A. Lapedes and R. Farber, "Nonlinear signal processing using neural networks: Prediction and system modelling," Los Alamos National Laboratory, Tech. Rep., 1987.
- [5] A. S. Weigend, D. Rumelhart, and B. A. Huberman, "Predicting the future: A connectionist approach," *International Journal Neural System*, vol. 1, no. 3, 1990.
- [6] D. Rumelhart, G. Hinton, and R. Williams, *Learning internal representation by error propagation, Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge: MIT Press, 1986, vol. 1.
- [7] H. Maier and G. Dandy, "Understanding the behaviour and optimising the performance of back-propagation neural networks: an empirical study," *Environmental Modelling and Software*, vol. 13, pp. 179–191, 1998.
- [8] W. Pedrycz and F. Gomide, *An Introduction to Fuzzy sets: Analysis and Design*. Cambridge, M.A: MIT Press, 1998.
- [9] M. Magalhães, R. Ballini, R. Gonçalves, and F. Gomide, "Predictive fuzzy clustering model for natural streamflow forecasting," *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 390–394, 2004.
- [10] E. Backer, *Computer Assisted Reasoning in Cluster Analysis*. New York: Prentice Hall, 1995.
- [11] A. Geva, "Non-stationary time series prediction using fuzzy clustering," *18th International Conference of the North American Fuzzy Information Processing Society - NAFIPS*, pp. 413–417, 1999.
- [12] C. Shalizi, K. Shalizi, and J. Crutchfield, "Pattern discovery in time series, part i: theory, algorithm, analysis, and convergence," *Journal of Machine Learning Research*, October 2002.
- [13] C. J. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition*. New York - USA: IEEE Press, 1992.
- [14] M. Magalhães, R. Ballini, P. Molck, and F. Gomide, "Combining forecasts for natural streamflow prediction," *Proceedings of the International Conference of the North American Fuzzy Information Processing Society*, 2004.
- [15] R. Ballini, "Análise e previsão de vazões utilizando modelos de séries temporais, redes neurais e redes neurais nebulosas," Ph.D. dissertation, Unicamp, Campinas - SP, 2000.