Monthly Electric Energy Demand Forecasting by Fuzzy Inference System

Ivette Luna and Rosangela Ballini
Departament of Economic Theory, DTE
Institute of Economics, IE
University of Campinas, UNICAMP
{ivette,ballini}@eco.unicamp.br

Resumo – Este artigo apresenta os resultados de prospecção de um e doze passos à frente da demanda mensal de energia elétrica de uma concessionária de energia pertencente a região sudeste do Brasil. Neste trabalho a demanda de energia elétrica total é subdividida em três grupos de consumo: residencial, industrial e comercial. O modelo de previsão adotado é baseado em regras nebulosas do tipo Takagi-Sugeno (TS), sendo o número de regras obtido via algoritmo de agrupamento não supervisionado *Subtractive Clustering*. Uma base de regras nebulosa é determinada para cada classe de consumo e os parâmetros do sistema de inferência são ajustados usando o algoritmo de otimização de maximização da verossimilhança. Como variáveis de entrada são consideradas as observações de demanda em instantes anteriores além de variáveis explicativas de natureza macroeconômica. O desempenho do modelo é verificado por meio de medidas de erros calculadas dentro e fora da amostra e os resultados indicam que o sistema de inferência nebuloso atinge índices de desempenho na ordem anual de 3% para as classes de consumo.

Palavras-chave – Demanda de energia elétrica, sistema de inferência nebuloso, previsão, séries temporais.

Abstract – This paper presents the use of a fuzzy rule-based system for mid-term electric energy demand forecasting. The results are achieved for an specific region at the Southeastern part of Brazil. The total demand is divided in three groups of consumption: residential, industrial and commercial. The forecasting model adopted is based on Takagi-Sugeno fuzzy rules, where the number of fuzzy rules is defined by the Subtractive Clustering algorithm, an unsupervised approach applied over an in-sample data set. A fuzzy rule base is determined by each group of consumption and the model parameters are adjusted using the Expectation Maximization optimization algorithm. As input variables are considered the observations of demand in previous moments as well as macroeconomic explanatory variables. Forecasting tests over an in-sample and out-of-sample data sets are developed. The results show the adequacy of the models adjusted, achieving annual absolute percentage errors of 3% in average.

Keywords – Electric energy demand, fuzzy inference system, forecasting, time series.

1. INTRODUCTION

Econometric models have been developed to forecast electric energy demand in several countries. A survey of statistical methods to evaluate urban energy needs is presented by [1]. Determinants of energy demand in the literature include economics variables, such as, population growth, prices, income and GDP, and climatic variables, such as degree days temperature [2], [3].

In Brazil, the energy crisis in mid-2001 unchained a process of rationing of electric energy in the Southeastern region of the country and a called phenomenon "rationalization of consumption". This phenomenon was responsible for modifying habits and behaviors with respect to consumption and demand of electric energy. The consumption platforms were modified as much as the schedules of occurrence of the maximum demand of electric energy so that studies of demand forecasting in this sector have become even more necessary.

Following this line of research, different computational intelligence methods have emerged in the last two decades as alternatives for building effective predictors, due to the inherent non-linear nature that maps the relation among independent and dependent variables. Usually, these models are based on neural networks, fuzzy systems and hybrid approaches [4]. These techniques involve an iterative process of model adjustment during a sequence of offline parameters updates, considering at each iteration all data available for training the model.

As an alternative to existing models, this paper suggests a fuzzy rule-based structure and a learning algorithm for building time series models. This approach is based on the Takagi-Sugeno fuzzy systems (FIS) and the Expectation Maximization (EM) optimization technique [5]. Here, FIS structure is defined in two phases. In the first phase, an initial rule based system composed by a set of fuzzy rules is generated using a Subtractive Clustering algorithm (SC), originally proposed in [6]. In a second phase, the model is re-adjusted using the Expectation Maximization algorithm, where all the model parameters are adjusted considering as a start point results obtained with the SC algorithm. This model is applied for explaining the regional demand of electric energy because of its capability of dealing with non-linear relations between variables, which is more appropriate for the problem analyzed.

After this introduction, this paper proceeds as follows. Section II presents a brief description of the general FIS structure. Section III details the methodology and the case study. Finally, some conclusions and further research are presented in Section V.

2. FUZZY INFERENCE SYSTEM

2.1 Structure

Let $\mathbf{x}^k = [x_1^k,\ x_2^k,\ \ldots,\ x_p^k] \in \mathbb{R}^p$ denote the input vector at instant k, $k \in \mathbb{Z}_0^+$; $\hat{y}^k \in \mathbb{R}$ is the output model, for the correspondent input $\mathbf{x}^k$. The input space represented by $\mathbf{x}^k \in \mathbb{R}^p$, is partitioned into M sub-regions, each represented by a fuzzy rule; $k = 0, 1, 2, \ldots$ is the time index (Figure 1). The antecedents of each fuzzy **If-Then** rule (R_i) are represented by their respective centers $\mathbf{c}_i \in \mathbb{R}^p$ and covariance matrices $\mathbf{V}_i|_{p\times p}$. The consequents are represented by local linear models, with output y_i, $i = 1, \ldots, M$ defined by:

$$y_i^k = \phi^k \times {\theta_i}^T \tag{1}$$

where $\phi^k = [1\ x_1^k\ x_2^k\ \ldots x_p^k]$; $\theta_\mathbf{i} = [\theta_{i0}\ \theta_{i1}\ \ldots\ \theta_{ip}]$ is the coefficient vector of the local linear model for the i^{th} rule.

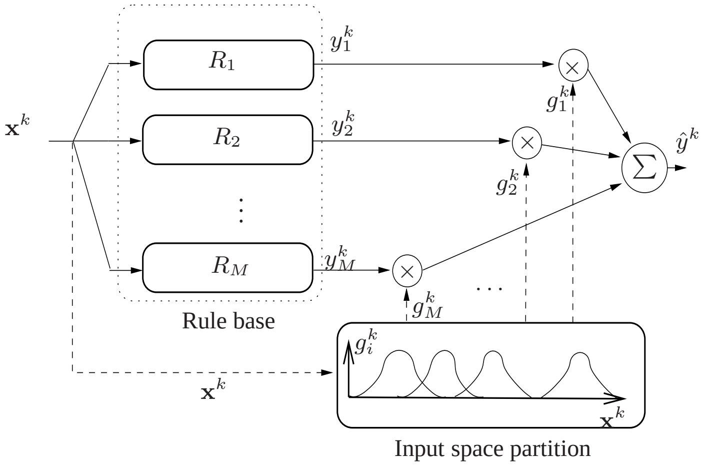

Figura 1: General FIS formulation.

Each input pattern has a membership degree associated with each region of the input space partition. This is calculated through membership functions $g_i(\mathbf{x}^k)$ that vary according to centers and covariance matrices related to the fuzzy partition, and are computed by:

$$g_i(\mathbf{x}^k) = g_i^k = \frac{\alpha_i \cdot P[\ i \mid \mathbf{x}^k\ ]}{\sum_{q=1}^{M} \alpha_q \cdot P[\ q \mid \mathbf{x}^k\ ]} \tag{2}$$

where α_i are positive coefficients satisfying $\sum_{i=1}^{M} \alpha_i = 1$ and $P[\ i \mid \mathbf{x}^k]$ is defined according to

$$P[\ i \mid \mathbf{x}^k\ ] = \frac{1}{(2\pi)^{p/2} \det(\mathbf{V}_i)^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}^k - \mathbf{c}_i)\mathbf{V}_i^{-1}(\mathbf{x}^k - \mathbf{c}_i)^T \right\} \tag{3}$$

where $\det(\cdot)$ is the determinant function. The model output $y(k) = \hat{y}^k$, which represents the predicted value for future time instant k, is calculated by means of a non-linear weighted averaging of local outputs y_i^k and its respective membership degrees g_i^k, i. e.,

$$\hat{y}(\mathbf{x}^k) = \hat{y}^k = \sum_{i=1}^{M} g_i^k\ y_i^k \tag{4}$$

2.2 Optimization

First, an initial structure composed by fuzzy rules is defined, and its parameters are adjusted via the traditional Expectation Maximization algorithm, originally proposed in [7] for mixture of experts models.

Model structure is initialized using the unsupervised clustering algorithm called the Subtractive Clustering Algorithm (SC), proposed in [6]. This algorithm provides a set of M clusters from a specific training data set presented to the algorithm. Patterns processed by the SC algorithm are composed by the input-output patterns used in a second stage for model optimization.

These groups are associated to a set of fuzzy rules codified in the FIS structure. Therefore, after the number of fuzzy rules is defined, we proceed to initialize the model parameters, for $i = 1, \ldots, M$, according to the following criteria:

- $\mathbf{c}_i^0 = \psi_i^0|_{1\ldots p}$, where $\psi_i^0|_{1\ldots p}$ is composed by the first p components of the $i-$th center found by the SC algorithm;

- $\sigma_i^0 = 1.0$;

- $\theta_i^0 = [\psi_i^0|_{p+1}\ 0\ \ldots\ 0]_{1\times p+1}$, where $\psi_i^0|_{p+1}$ is the $p+1-$th component of the $i-$th center found by the SC algorithm;
- $\mathbf{V}_i^0 = 10^{-4}\mathbf{I}$, where $\mathbf{I}$ is a $p \times p$ identity matrix;
- $\alpha_i^0 = 1/M$.

After this initialization, model parameters are re-adjusted based on the traditional offline EM algorithm, following an iterative sequence of EM steps, given incomplete data y^k. It means that, a complete data is composed by the output variable y^k and a missing data. The goal of the EM algorithm is to find a set of model parameters, which will maximize the log-likelihood $\mathcal{L}$, of the observed values of y^k at each M step of the learning process. This objective function is defined by

$$\mathcal{L}(D,\ \mathbf{\Omega}) = \sum_{k=1}^{N} \ln\left(\sum_{i=1}^{M} g_i(\mathbf{x}^k,\ \mathbf{C}) \times P(y^k \mid \mathbf{x}^k,\ \theta_i)\right) \tag{5}$$

where $D = \{\mathbf{x}^k, y^k | k = 1, \ldots, N\}$, $\mathbf{\Omega}$ contains all model parameters and $\mathbf{C}$ contains just the antecedents parameters (centers and covariance matrices). However, for maximizing $\mathcal{L}(D,\ \mathbf{\Omega})$, it is necessary to estimate the missing data h_i^k during the E step. This missing data, according to mixture of experts theory, is known as the posterior probability of $\mathbf{x}^k$ belong to the active region of the $i-$th local model.

When the EM algorithm is adapted for adjusting fuzzy systems, h_i^k may also be interpreted as a posterior estimate of membership functions defined by Eq. (2). So, h_i^k is calculated as

$$h_i^k = \frac{\alpha_i P(i \mid \mathbf{x}^k) P(y^k \mid \mathbf{x}^k,\ \theta_i)}{\sum_{q=1}^{M} \alpha_q P(q \mid \mathbf{x}^k) P(y^k \mid \mathbf{x}^k,\ \theta_q)} \tag{6}$$

for $i = 1, \ldots, M$. These estimates are called as "posterior", because these are calculated assuming $y^k,\ k = 1, \ldots, N$ as known. Moreover, conditional probability $P(y^k|\mathbf{x}^k, \theta_i)$ is defined by:

$$P(y^k \mid \mathbf{x}^k,\ \theta_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{[y^k - y_i^k]^2}{2\sigma_i^2}\right) \tag{7}$$

with σ_i^2 estimated as:

$$\sigma_i^2 = \left(\sum_{k=1}^{N} h_i^k [y^k - y_i^k]^2\right) / \sum_{k=1}^{N} h_i^k \tag{8}$$

Hence, the EM algorithm for determining FIS parameters can be summarizes as:

1. E step: Estimate h_i^k via Eq. (6);
2. M step: Maximize Eq. (5) and update model parameters, with optimal values calculated as:

$$\alpha_i = \frac{1}{N} \sum_{k=1}^{N} h_i^k \tag{9}$$

$$\mathbf{c}_i = \left(\sum_{k=1}^{N} h_i^k \mathbf{x}^k\right) / \sum_{k=1}^{N} h_i^k \tag{10}$$

$$\mathbf{V}_i = \left(\sum_{k=1}^{N} h_i^k (\mathbf{x}^k - \mathbf{c}_i)'(\mathbf{x}^k - \mathbf{c}_i)\right) / \sum_{k=1}^{N} h_i^k \tag{11}$$

   for $i = 1, \ldots, M$, where M is the size of the fuzzy rule base, N is the number of input-output patterns at the training set. For all these equations, $\mathbf{V}_i$ was considered as a positive diagonal matrix, as an alternative to simplify the problem and avoid infeasible solutions. An optimal solution for θ_i is derived solving the following equation:

$$\sum_{k=1}^{N} \frac{h_i^k}{\sigma_i^2} \left(y^k - \phi^k \times \theta_i\right) \cdot \phi^k = 0 \tag{12}$$

   where σ_i is the standard deviation for each local output $y_i,\ i = 1, \ldots, M$, with σ_i^2 defined by Eq.(8). After parameters adjustment, calculate the new value for $\mathcal{L}(D,\ \mathbf{\Omega})$.

3. If convergence is achieved, then stop the process, else return to step 1.

3. METODOLOGY AND CASE STUDY

3.1 Data analysis and preprocessing

Realized monthly energy consumption belongs to an energy enterprise in the Southeastern region of Brazil. Figure 2 presents historical monthly consumption data used for this case study (in MWh), going from January 2003 to December 2008. These historical series of energy consumption will be used for estimating the expected demand for future months.

The total electric energy consumption registered is depicted in Figure 2-(d). In this paper, this series is composed by the aggregation of three types of consumption: residential, industrial and commercial. Since the region to which these observations belong has a strong industrial activity, the industrial consumption depicted in Figure 2-(b) represents in average 69% of the total consumption, whereas the residential and the commercial ones represent 24% and 7%, respectively (Figure 2-(a) and (c)).

As observed, the residential, industrial and commercial consumption present a strong trend and a seasonal component. The last one is easier to observe in the commercial class. Since the FIS model works with stationary data, all these time series need to be integrated or transformed for removing the trend component.

As commonly used in a production planning framework [8], the forecasting of the total electric energy demand was estimated by the aggregation of individual forecasts for the three types of demand, following a hierarchical bottom-up approach, which estimates the parts for obtaining an aggregate total estimate [9].

The actual version of the forecasting models consider macroeconomic variables as inputs to the fuzzy inference systems.

In the case of the residential consumption, it was assumed that the observed trend is highly related to the population growth - which is also related to the number of consumers registered by the distribution company - as well as the number of billed days in a month. We remove the effect of the population trend by dividing the monthly data by the number of regional consumers. This procedure was adopted in order to make use of the information present in the series of monthly regional consumers. Indeed, in order to estimate the residential demand free of the effect of the billing date, the residential consumption per capita was divided by the number of monthly billing days that varies for each month and for each year. In this way, we obtain the residential consumption per capita and per billed day.

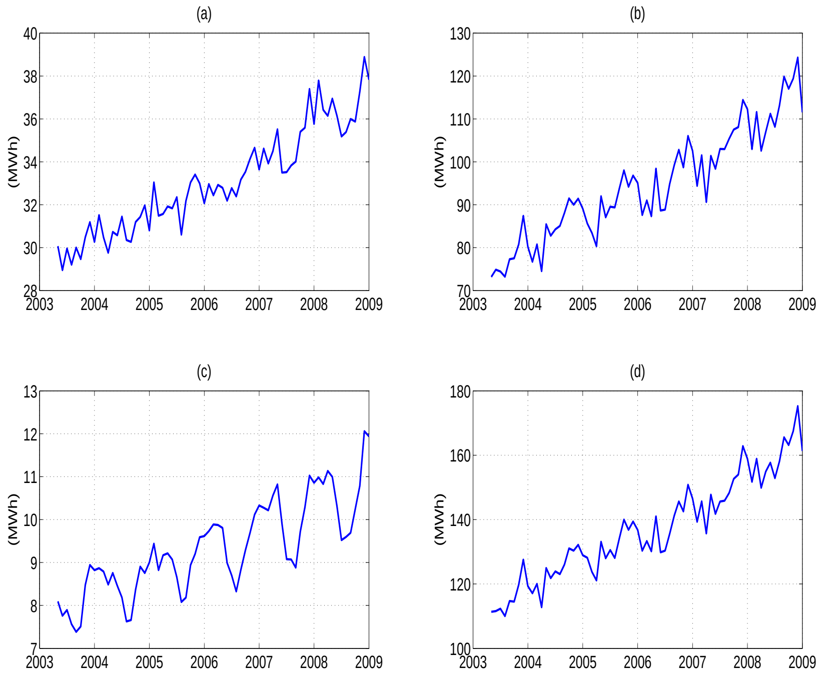

Figura 2: Monthly electric energy consumption from 2003 to 2008: (a) residential; (b) industrial; (c) commercial; (d) total.

Different works in the literature (see [3]) show evidence that climate conditions affect the residential consumption, effect that may be represented by the heating degree days (HDD) and cooling degree days (CDD) index which are calculated using daily maximum and minimum temperatures recorded. Unfortunately, daily records of regional temperature were not available, so that there was no possible to consider climate variations as an exogenous factor. Hence, we limited our study to models including economic factors.

In order to consider macro and socioeconomic variables for reflecting the regional economical development, the model adjusted for the residential consumption per capita and per billed day used as exogenous variable the total monthly regional employment rate. This variable was available at the CAGED portal (*Cadastro Geral de Empregados e Desempregados*) [10].

The effect of the billing date was also observed in the commercial consumption. Therefore, the commercial data was also divided by the number of monthly billing days.

The total employment rate was also available at the CAGED portal for each activity sector. Hence, the specific monthly employment rate related to regional commerce and service was considered as an explicative variable by the commercial demand model.

In the case of the industrial consumption, the final model considered as exogenous variable the regional monthly industrial production index, which estimates are based on the regional participation in all the industrial activities considered for calculating the state industrial production index (IPI), available at IBGE (*Instituto Brasileiro de Geografia e Estatística*) ([11]).

Since all the series presented a trend component, they were integrated once, and then normalized to the unit interval.

Data from 2003 to 2007 were used for calibration purposes, while the data from 2008 were considered for testing. A validation set for verifying overfitting was not defined due to the limited availability of historical data.

3.2 Model identification and adjustment

Input patterns were built considering lags of the same consumption as well as lags of the respective exogenous variables. Several configurations were adjusted, varying the number of lags for each variable, as well as different values for the subtractive clustering hyperparameters. The selection of the final model was performed considering the one-step ahead root mean square error over the testing data set.

Based on this procedure, subtractive clustering considered parameters $r_a = 0.50, 0.55, 0.25$ for the residential, industrial and commercial models, respectively and $r_{ba} = 1.0$ for all the cases. The final input patterns configuration that provided the best result for each model over the testing set is summarized in Table 1, where t denotes the time instant in months. Therefore, if we pretend to forecast industrial demand for instant t, then, the input pattern will be composed by the first two lags of the regional industrial production index and the last two records of the industrial demand.

Tabela 1: Final input patterns configuration for demand forecasting at instant t.

Residential		
	Lags	
Total employment rate	$t-1$	$t-2$
Consumption[a]	$t-1$	$t-2$
Industrial		
	Lags	
Regional Industrial Production Index	$t-1$	$t-2$
Consumption	$t-1$	$t-2$
Commercial		
	Lags	
Commerce and Service employment rate	$t-1$	$t-2$
Consumption[b]	$t-1$	$t-2$

[a] Residential consumption per capita and per billed day.

[b] Commercial consumption per billed day.

3.3 Analysis of the results

In order to verify the model performance, three performance metrics were evaluated: the root mean square error RMSE (MWh), mean absolute percentage error MAPE (%) and mean absolute error MAE (MWh).

The first task considered was the one step ahead forecasting of the regional monthly electric energy demand. Results achieved for the residential demand are depicted in Figure 3-(a), whereas results for industrial and commercial demand are illustrated in Figure 3-(b) and 3-(c), respectively. All these figures show the results achieved over the in-sample (2003-2007) and the out-of-sample (2008) data sets.

As observed, results obtained specially for the residential and the commercial demands are very promising. In the case of the industrial demand, we can observe a slightly higher deviations for some months, affecting in general the total demand estimate for 2008. Numerical results for the performance metrics are presented in Table 2, where the higher level for one step ahead forecasting errors in the industrial demand is evident.

Table 2 also shows results achieved for a twelve steps ahead forecasting task over 2008. The most important issue to observe is that, comparing both results for each model, the performance is preserved, which may be understood as an evidence of the model robustness, since it is capable of dealing with errors related to previous forecasts samples that are fedback through the input vector. Results for a twelve steps ahead forecasting task are depicted in Figure 4.

Observing carefully the behavior of the industrial demand, we can see that the FIS adjusted for this type of consumption estimated the industrial demand of 2008 according to the behavior observed until 2007, which was well captured given the errors achieved in the in-sample data set. However, in 2008 the industrial consumption was lower that the expected, producing errors in the out-of-sample data set higher than the ones obtained in the residential and the commercial types. This behavior is observed more specifically from mid-2008 when the subprime crisis was at its highest, causing a slowdown in the economy of the region. In that sense, we can conclude that the regional (ponderate) industrial production index was not enough to explain the industrial electric energy demand behavior.

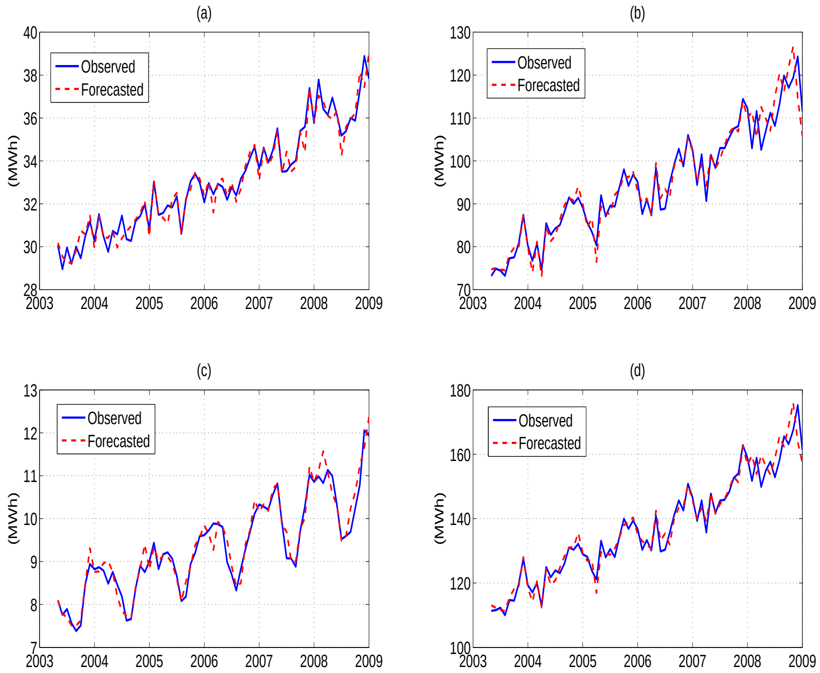

Figura 3: One step ahead forecasting results: (a) residential; (b) industrial; (c) commercial; (d) total.

Tabela 2: Performance metrics for one and twelve steps ahead prediction.

	Residential		
	In-sample	Out-of-sample	
		One step ahead	Twelve steps ahead
RMSE	0.438	0.775	0.552
MAPE	0.981	1.674	1.273
MAE	0.313	0.622	0.469
	Industrial		
	In-sample	Out-of-sample	
		One step ahead	Twelve steps ahead
RMSE	1.806	6.770	6.174
MAPE	1.613	5.742	4.898
MAE	1.442	6.415	5.552
	Commercial		
	In-sample	Out-of-sample	
		One step ahead	Twelve steps ahead
RMSE	0.228	0.380	0.250
MAPE	1.758	2.782	1.705
MAE	0.158	0.301	0.188
	Total		
	In-sample	Out-of-sample	
		One step ahead	Twelve steps ahead
RMSE	2.027	7.964	7.126
MAPE	1.248	8.732	8.158
MAE	1.625	12.689	11.777

In terms of annual results, Table 3 presents RMSE and MAPE achieved over 2008. These results show that the aggregation of the forecasting results to estimate the annual expected demand is affected by the lead time, since the one-step ahead annual result is more accurate than the one estimated considering the time origin in January 2008. As noticed, the error in the forecasting of the total demand is directly affected by the industrial results. Therefore, it is necessary to study other macroeconomic variables with explanatory power for modeling the industrial demand, due to the structural break presented in 2008.

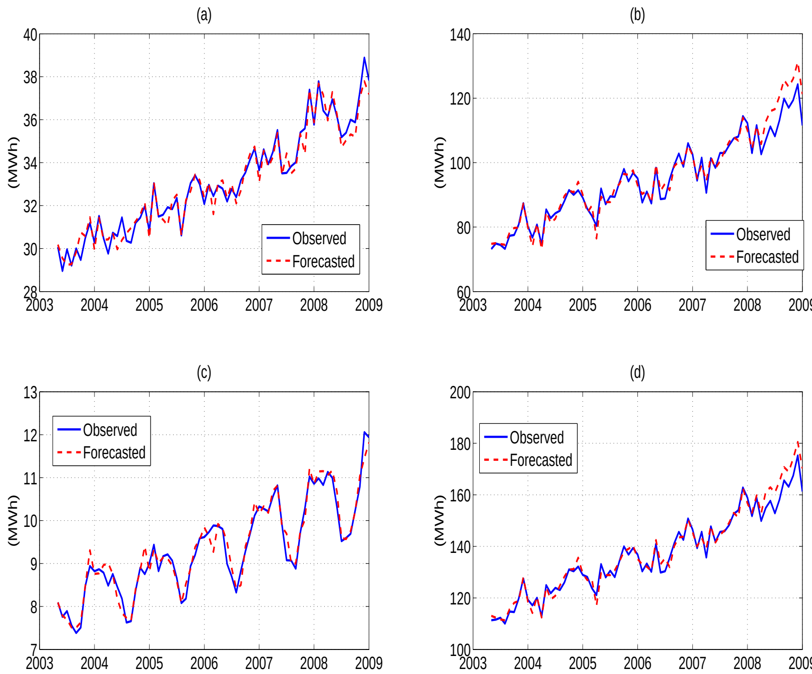

Figura 4: Twelve steps ahead forecasting results: (a) residential (b) industrial (c) commercial (d) total

Tabela 3: Performance metrics for annual forecast (2008).

	Residential		Industrial		Commercial		Total	
	One step ahead	Twelve steps ahead	One step ahead	Twelve steps ahead	One step ahead	Twelve steps ahead	One step ahead	Twelve steps ahead
RMSE	0.994	3.259	18.215	66.313	1.777	0.542	18.998	63.596
MAPE	0.226	0.741	1.350	4.916	1.387	0.423	0.991	3.318

4. CONCLUSIONS AND FUTURE WORK

This paper proposes a fuzzy system model for mid-term electric energy demand forecasting. The model is based on an offline model structure definition and parameter adjustment. Some advantages of the model proposed are that it provides a compact structure and a fast learning, which are great advantages in terms of time process and computational effort. The fuzzy inference system was applied to energy demand forecasting of an enterprise in the Southeastern region of Brazil. The total demand was divided in three groups of consumption: residential, industrial and commercial; one model is adjusted for each group. The results obtained specially for the residential and the commercial demands are very promising. In the case of the industrial demand, we can observe a slightly higher deviations for some months, affecting in general the total demand estimate. Numerical results showed the adequacy of the models adjusted, achieving annual absolute percentage errors of 3% in average. For future works, we would like to compare the fuzzy inference system with econometric models for mid-term electric energy demand forecasting and analyze the influence of climatic and other macroeconomic variables in the electric energy consumption.

ACKNOWLEDGMENT

The last author thank the Brazilian National Research Council, CNPq, for grants 302407/2008-1.

REFERENCES

[1] C. Balocco and G. Grazzini. "A statistical method to evaluate urban energy needs". *International Journal of Energy Research*, vol. 21, no. 14, pp. 1321–1330, 1997.

[2] G. E. Nasr, E. A. Badr and G. Dibeh. "Econometric modeling of electricity consumption in post-war Lebanon". *Energy Economics*, vol. 22, pp. 627–640, 2000.

[3] C.-L. Hor, S. Watson and S. Majithia. "Analyzing the impact of weather variables on monthly electricity demand". *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 2078 – 2085, 2005.

[4] S. Mitra and Y. Hayashi. "Neuro-fuzzy rule generation: survey in soft computing framework". *IEEE Transactions on Neural Networks*, vol. 11, no. 3, pp. 748–768, May 2000.

[5] M. Jordan and R. Jacobs. "Hierarchical Mixtures of Experts and the EM Algorithm". Technical Report AIM-1440, 1993.

[6] S. Chiu. "A cluster estimation method with extension to fuzzy model identification". In *Proceedings of The IEEE International Conference on Fuzzy Systems*, volume 2, pp. 1240–1245, June 1994.

[7] R. Jacobs, M. Jordan, S. Nowlan and G. Hinton. "Adaptive Mixture of Local Experts". *Neural Computation*, vol. 3, no. 1, pp. 79–87, 1991.

[8] H. Widiarta, S. Viswanathan and R. Piplani. "Forecasting aggregate demand: An analytical evaluation of top-down versus bottom-up forecasting in a production planning framework". *International Journal of Production Economics*, vol. 118, no. 1, pp. 87–94, March 2009.

[9] L. Lapide. "Top-down & bottom-up forecasting in S&OP". *Journal of Business Forecasting*, vol. 25, no. 2, pp. 14–16, 2006.

[10] Cadastro Geral de Empregados e Desempregados. "http://granulito.mte.gov.br/portalcaged/paginas/home/home.xhtml", 2010.

[11] Instituto Brasileiro de Geografia e Estatística. "http://www.ibge.gov.br/", 2009.