Paper accepted for presentation at 2009 IEEE Bucharest Power Tech Conference, June 28th - July 2nd, Bucharest, Romania

An Hybrid Aggregate Model Applied To the Short-Term Bus Load Forecasting Problem

Ricardo Menezes Salgado, Rosangela Ballini and Takaaki Ohishi

***Abstract*—In this paper we present a hybrid methodology built on a combination of clustering and forecasting techniques used to solve the short-term bus load forecasting problem. The proposed method was made in two phases: In the first phase a clustering algorithm is used to identify buses clusters with similar daily load profile and in the second phase is proposed an aggregate structure for to foresee each bus using a conventional prediction model. The methodology was applied on bus load data from the Brazilian North/Northeast system and the results showed that the model was efficient with 2% to 3.6% of the mean percentage error level on the buses.**

***Index Terms*—Short-Term Bus Load Forecasting, Aggregate Forecasting Model, Clustering Algorithm, Artificial Neural Networks, Multiple Linear Regression, Time Series Analysis.**

I. Introduction

THE the generation, in an electric power system, is distributed to the consumption area through the transmission network using points of the network, called bus node. In this network points (buses) the load is delivered to the consumers fed by this node, in this sense, the operation of the transmission network is influenced by total consumption at each node. In this case, to evaluate the impact of a given generation dispatch, it is necessary to know the load power at each bus node [15], [21]. In real power system operation these kind of analysis must be performed in short-term basis, because they support several steps in the real time system operation. The knowledge about bus load demand is also important for decision-making such as reliability analysis [3], congestion analysis, system operation, commercial strategies, tariffs definition [19], [22] and energy prices in the electricity markets [17] in all cases the bus load forecasting must be done in short-term basis.

In general short-term bus load forecasting can be viewed as a traditional short-term global load forecasting, in which for each bus can be developed a specific forecasting model. However, due to the great number of buses, the execution of all these different models requires considerable computational time and it can be inadequate to short-term decisions.

The most of the bus load forecasting methods reported in the literature make the forecasting individually for each bus of the system. Nevertheless, in a real system the bus load forecast must be performed in short-term basis because it is support of system operation. Moreover, in a system with many buses, it is very difficult to obtain an individual bus load forecast in short-time.

There are a few papers treating the bus load forecasting in the literature and the majority of papers considers only global load demand or individual bus load forecasting. A linear model has been proposed for bus load prediction in distribution networks in [9]. This model relates load measurements with normalized load curves (state variables) for the individual customer groups. In [6], application of periodic time series for bus load forecasting has been presented.

In [13], a hybrid approach utilizing a fuzzy system and artificial neural network has been presented for bus load forecasting and this approach was applied in one residential bus from the town of Hinton, West Virgina. In [19] was proposed a dynamic state estimator in the form of extended Kalman filter (EKF) with inclusion of second-order terms for bus load forecasting. In this paper also the authors incorporate a neural network (ANN) model in the prediction step of the state estimator. However, computational effort of the approach was higher, and so it is not adequate for on-line applications in the real power networks. In [20], the authors proposed an hierarchical model at the filtering step of their state estimator. It is important to note that the previous papers treat each bus load forecast individually.

In [1] a hybrid bus load forecasting method composed of the forecast-aided state estimator (FASE) and the multilayer perceptron neural network (MLP) was applied in three buses in Iranian power system. The FASE forecasts hourly loads of each bus by means of its previous data, then the inputs and outputs of the FASE are fed to the MLP. The results obtained show that the hybrid method has better prediction accuracy than MLP, FASE and the periodic auto-regression (PAR) models.

To solve the problem of the bus load forecast making possible short-time decisions this paper proposes an aggregate model to short-term bus load forecasting. The main idea is to use the similarity existent in the load profile in each bus and then to make an aggregate forecasting for each group of the buses with same profile. Thus, the proposed model combines two techniques and is built in two phases: The first one uses a clustering algorithm to obtain a set of data clusters where the main goal is to cluster the buses with similar daily load profile, in this context we use the Self Organized Map (SOM) algorithm [14] and [11] to find the buses clusters. The second phase applies a forecast model developed specifically for each buses clusters, the prediction model can be specific characteristic (linear or nonlinear) that

Ricardo Menezes Salgado is with Department of Exact Sciences, Federal University of Alfenas, Alfenas - Minas Gerais/Brazil — e-mail: ricardomenezes@unifal-mg.edu.br

Takaaki Ohishi is with Department of System Engineering, School of Electrical and Computer Engineering, University of Campinas, Campinas - São Paulo - Brazil — e-mail: taka@cose.fee.unicamp.br

Rosangela Ballini is with Institute of Economics, University of Campinas, Campinas - São Paulo - Brazil — e-mail: ballini@eco.unicamp.br

assist each cluster profile.

To evaluate the performance of the aggregate process in the real case, the methodology was applied on real bus load data from the Brazilian North/Northeast system. The results showed that the proposed model was efficient with mean percentage error about of the 2% to 4% in most of the buses, this error level shown that the proposed model can be applied successfully in the short-term bus load forecasting. Moreover, we compared the performance of this aggregate approach with individual bus load forecasting to compare the approaches in the solution of the bus load forecast as well as to verify the time computacional demanded by each approach.

The remaining sections of this paper are organized as follows. A brief review of the problem and its difficulties are presented in section I. In section II the bus load forecasting problem and the prediction/clustering models are detailed. The section III shows the proposed bus load forecasting model explaining its attributes. In section IV the bus load data used in this paper are presented and discussed. An application case of the method for bus load forecast and the obtained numerical results are presented and discussed in section V. Section VI concludes the paper and summarizes issues that deserve further developments.

II. Problem and Models Description

This section make the short description of the the bus load forecasting problem and details the models of the forecasting and clustering used in this paper.

A. Short Description: Bus Load Forecasting Problem

As previously observed the bus load forecasting is an important input to auxiliary the electric power system operation. In this context, an important stage is the short-term operation planning, which determines an electric power production scheduling for the next days, usually in hourly basis. This scheduling must determine the operational requirements of generation and transmission systems. Particularly, for the transmission system operation analysis a load balance is executed at each bus and, consequently, it is necessary to have a estimative of the bus load demand for each time interval.

In practical way, the bus electrical load can be defined as the sum of energy consumption of all feeders of a given bus. In Fig. 1 is possible see an example of the a simple electrical network diagram where, even in a simple example, there is a decision problem associate to the load flow in the buses from 1 to 4. This decision depends of the line limits and other factors associate to the load flow and in both cases an bus load forecasting with good precision is essential to provide informations to support this operational decisions.

The bus load forecasting process can be defined as the conventional time series forecasting in which for each bus can be developed a specific forecasting model. But, due to the great number of buses, the execution of the one model to the each bus requires considerable computational time and it can be inadequate to short-term decisions. In general way, the prediction of bus loads usually may be more complex than forecasting of the system total demand, this fact happens

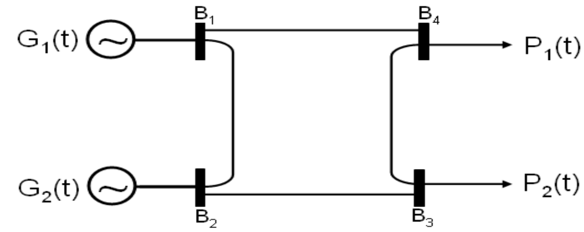

Fig. 1. Example: Simple Electrical Network Diagram.

due the characteristics of the bus loads time series. Generally the bus load data has more outliers, nonlinearity, and high-frequency components than the system global demand. In many cases, the historical load demand data of a given bus may have significant distortion due to change on network configuration. Another relevant factor is the diversity of the profiles that increase the complexity of problem.

The Brazilian electric system is composed by more than 4000 buses with different levels of the load and of voltages. In Fig. 2 is possible see a small region, in Brazil, with great number of buses that assists several consuming areas. This great numbers of information turns the problem complex and difficult of solving through conventional methods. At the Table I is possible see that there are many buses with different voltage levels, and observing the Fig. 3 we can see the load curve in two diffrent buses where there are significant differences in the profile of the load curve.

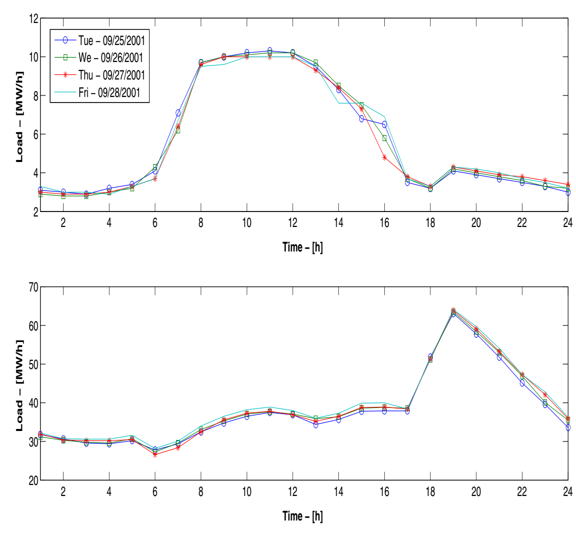

Fig. 3. Example: Bus Load Profiles.

In general, the bus load demands present a varied daily load profiles, depending on the predominant consumers of each bus. Observing the Fig. 3 we can see two buses with different load profiles and consumption levels. This fact may be associate to the consumer type fed by each bus and thus is necessary a model capable to detect these differences of profile providing forecasts of good quality.

A way to do the bus load forecasting is to foresee each

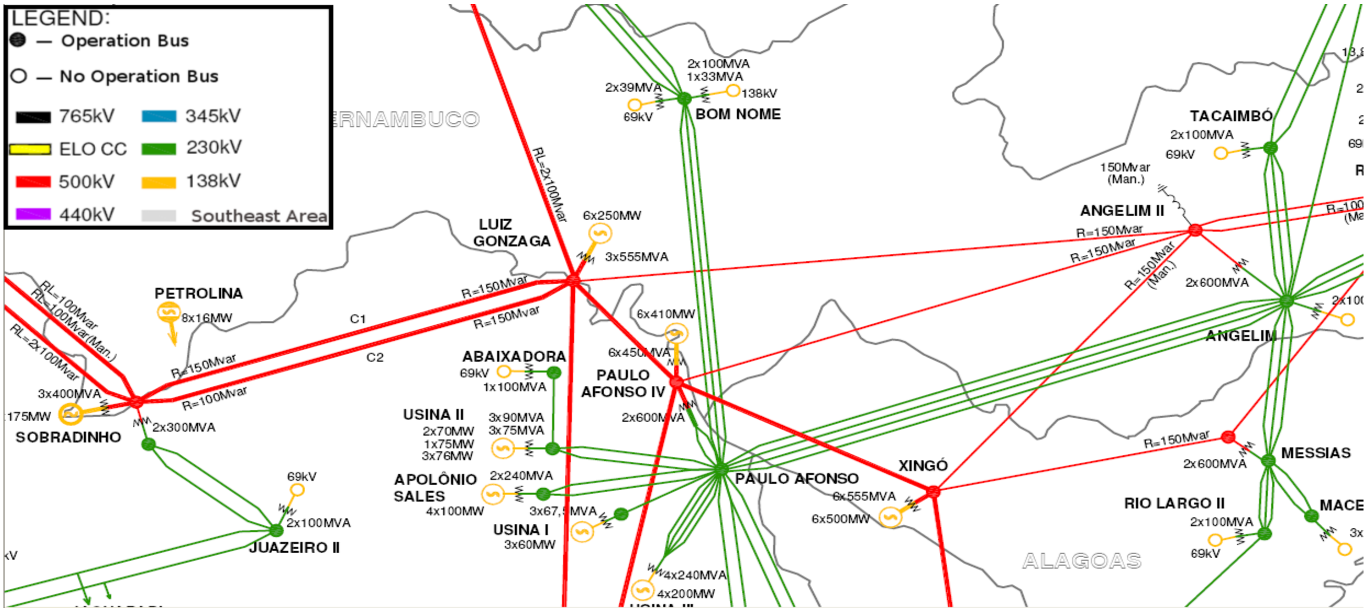

Fig. 2. Example: Electrical Network in Brazil.

TABLE I
BUSES BY VOLTAGE - TRANSMISSION SYSTEM/BRAZIL.

Voltage - Kv	*Buses*
[1 – 13.8]	901
(13.8 – 69]	540
(69 – 138]	1351
(138 – 500]	568
(500 – 765]	37
Total	*3397*

bus of the system individually but, in the systems with a high number of buses, the individual treatment is difficult because such task demands a long time to obtain of the forecasts [19], [22] and [12]. The alternative proposal to solve the bus load forecasting problem is to foresee the bus load using an aggregate model (aggregation of the buses by consumption profile). In this paper, the use of such forecast approach intends to provide a smaller number of the forecasts, thus reducing the computational effort demanded for the bus load forecasts.

B. Clustering and Forecasting Techniques

The model introduced in this paper is built in two phases. The first one uses a clustering algorithm to obtain clusters with similar daily load profile. In the second phase a forecasting model is adjusted for each cluster. The prediction model for each cluster can be either linear (typically linear regression) or nonlinear (neural networks, fuzzy predictors). These phases are briefly described below.

1) Clustering Algorithm: The main goal in this step is to cluster time series data with similar behavior. To detect such behaviors, we use the Kohonen Self-Organizing Feature Map (SOM) algorithm [14] to cluster the bus daily load profile.

The Kohonen SOM is part of a group of networks usually called networks based on competition, or simply competitive networks [14] and [11]. These networks combine competition with an unsupervised learning strategy to make the adjustment of the weights, based on existing similarities in the input patterns. The main goal of the SOM is to group similar input data into clusters.

Each output unit may represent a cluster, limiting the number of clusters to the number of output units. During the training process, the net determines the output unit (called winner) that is "closer" to the input vector; the weight vector associated to the winner is adjusted according to the learning algorithm. The units within a certain neighborhood of the winner may have their weight vectors similarly adjusted. In the next code is possible to see Algorithm 1 where is presented the the SOM training process.

Algorithm 1 SOM Algorithm

1. Initialization and parameter definition
- Initialize the weights w_{jk}.
- Define the neighborhood parameters.
- Define the learning parameters.

2. While the stopping criterion is false, do.

2.1 For each l determine:

2.1.1 $J = \underbrace{\arg min}_{j} \{\|w_j - x_l\|\}$

2.1.2 $\forall j \in Nc(J)$ and $\forall k$

$$w_{ij}(new) = w_{ij}(old) + \alpha\,[x_{lk} - w_{ij}(old)]$$

2.2. Update the learning rate.

2.3. Reduce the neighborhood radius.

The neighborhood radius $Nc(J)$ is important to update the weights. In the training process, $Nc(J)$ is decreased after a pre-specified number of iterations until $Nc(J)$ reaches zero. In this case, one iteration is assumed to be associated with the presentation of each input pattern.

Figure 4 presents a typical architecture for the multi-dimensional Kohonen SOM. The number of input units depends on the training data set, but the output grid, in principle, can have a varied number of nodes, arranged in a well-defined topological structure.

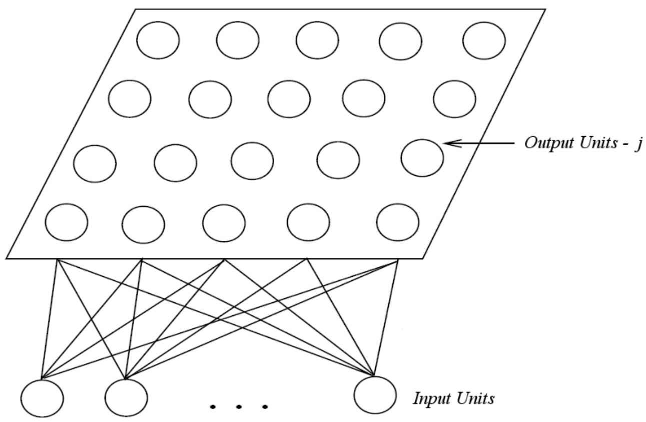

Fig. 4. Typical Multi-Dimensional SOM.

The application of SOM algorithm in buses clustering is showed in section V.

2) Forecasting Model: After the process of the detection of the buses clusters the next step is adjusted a predictive model for each cluster. In this paper we use a nonlinear model, more specifically, a multi-layer perceptron (MLP) neural network. This choice was done because the MLP is a tool frequently used in load forecasting with good results.

Architectures MLP represent the most used and known neural network models. An MLP consists of n inputs nodes, h hidden layer nodes, and m output nodes connected in a feed-forward fashion by multiplicative weights W. The MLP must be trained with historical data to find the appropriate values for W and the number of required neurons in the hidden layer. In this paper, the learning algorithm employed is the well-known error back propagation method [10].

Forecasting performance requires an appropriate selection of the input variables of the forecasting model. Here, the number of inputs was found using the partial autocorrelation function. This approach was applied for each bus and we construct a input-output matrix used to adjust the forecasting models.

III. Forecast Methodology

The bus load forecast can be seen as a problem of foreseeing n different time series in a short time interval. One in the ways of solving this problem is to treat each bus in an individualized way. This approach tends to produce good results, because the forecast model will be specifically developed to learn the characteristics of each bus. The inconvenience of this approach is the high computational effort to adjust each forecast model. Moreover, this approach is unviable for the short-term daily operation programming of the electric power systems with great number of the buses.

In a forecast model is necessary an adjustment phase where the parameters of the model are configured according to the data series presented. Usually the adjustment of a model is made by presentation of patterns of $(input \times output)$ in an iterative process. The forecast methodology developed in this paper proposes a forecast aggregate process in which the $(input \times output)$ buses data are presented in an only way using clustering information. The main idea adopted in this technique consists in creating a mapping of $(input \times output)$ through the matrix formed by blocks of specific buses. Thus, we creating the input/output matrices I and O that will just be formed by buses that possess similar profiles.

To minimize these inconveniences the found alternative it was to perform a clustering procedure to dividing the initial buses set in subsets of the buses with similar characteristics. To execute the bus forecasting by cluster reduces the complexity of the problem and it increases the correlation among the buses blocks of the matrix I that will just be formed by buses that possess similar behaviors. Another interesting factor is that with the clustering process the dimension of the matrices I and O will be smaller, reducing the complexity of the predictor.

The matrix I is a structure formed for n submatrices $\left\{I_{b(1)}, I_{b(2)}, \ldots, I_{b(n-1)}, I_{b(n)}\right\}$ that contains the input patterns for each bus in the cluster. In a similar way, the matrix O is a vectorial structure formed for n subvectors $\left\{O_{b(1)}, O_{b(2)}, \ldots, O_{b(n-1)}, O_{b(n)}\right\}$ that contains the desired output for each bus block $I_{b(i)}$. For each bus i exists an association $\left(I_{b(i)} \times O_{b(i)}\right)$ that represent the mapping $(input \times output)$ of the bus i, where $i = 1, \ldots, n$.

$$I = \begin{bmatrix} I_{b(1)} & 0 & \cdots & 0 & 0 \\ 0 & I_{b(2)} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_{b(n-1)} & 0 \\ 0 & 0 & \cdots & 0 & I_{b(n)} \end{bmatrix}$$

$$O = \begin{bmatrix} O_{b(1)} \\ O_{b(2)} \\ \vdots \\ O_{b(n-1)} \\ O_{b(n)} \end{bmatrix}$$

The matrix I was used to minimize the influences of the data of different buses and to guarantee that the informations of the bus i will influence only itself. The matrix I contains the inputs used for each bus of the same cluster and the output matrix O is formed by the respective desired output. With the matrices I and O, previously defined, it is possible to execute the bus load forecast adjusting a single predictor with the mapping $(I \times O)$.

An example of the description above can be seen in the Fig. 5. In this case the buses i and j are in the same cluster (possess same profile) and the aggregate structure is composed with 2 inputs for the bus i and one input to the bus j associates to each exit respectively. The strategies used to built the aggregate matrix is ideal to the bus load forecasting process.

A. Input/Output Blocks

From the buses clusters information are formed blocks $I_{b(i)}$ and $O_{b(i)}$ that composes the matrices I and O, respectively.

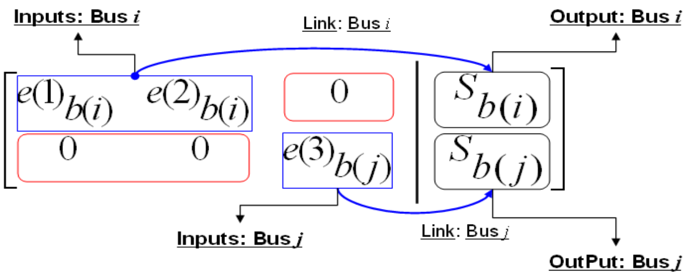

Fig. 5. Example: Aggregation Process.

Each block $I_{b(i)}$ it is composed by the inputs of the bus i and its respective desired output $O_{b(i)}$. The submatrix $I_{b(i)}$ contains m patterns discriminated by line and n inputs in each pattern, the element $b(i)^{(j)}_{[h,(d_n)]}$ represents the electric load of the bus i in a hour h for a give day d_n in the pattern j. Each input $b(i)^{(j)}_{[h,(d_n)]}$ corresponds the load of a same hour in different days.

$$\overbrace{\begin{bmatrix} b(i)^{(1)}_{[h,(d_n)]} & b(i)^{(1)}_{[h,(d_{n-1})]} & \cdots & b(i)^{(1)}_{[h,(d_2)]} & b(i)^{(1)}_{[h,(d_1)]} \\ b(i)^{(2)}_{[h,(d_n)]} & b(i)^{(2)}_{[h,(d_{n-1})]} & \cdots & b(i)^{(2)}_{[h,(d_2)]} & b(i)^{(2)}_{[h,(d_1)]} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b(i)^{(m-1)}_{[h,(d_n)]} & b(i)^{(m-1)}_{[h,(d_{n-1})]} & \cdots & b(i)^{(m-1)}_{[h,(d_2)]} & b(i)^{(m-1)}_{[h,(d_1)]} \\ b(i)^{(m)}_{[h,(d_n)]} & b(i)^{(m)}_{[h,(d_{n-1})]} & \cdots & b(i)^{(m)}_{[h,(d_2)]} & b(i)^{(m)}_{[h,(d_1)]} \end{bmatrix}}^{I_{b(i)}}$$

The structure $O_{b(i)}$ contains m desired outputs discriminated by line respectively linked to the m patterns of the submatrix $I_{b(i)}$.

$$\overbrace{\begin{bmatrix} b(i)^{(1)}_{[h,(d_{(0)})]} \\ b(i)^{(2)}_{[h,(d_{(0)})]} \\ \vdots \\ b(i)^{(m-1)}_{[h,(d_{(0)})]} \\ b(i)^{(m)}_{[h,(d_{(0)})]} \end{bmatrix}}^{O_{b(i)}}$$

In the aggregate process, the forecast is made with the information of the matrix I and O. This method create news matrices I and O for each cluster of the buses, i.e., if a buses set were clusterized in seven clusters, seven structures I and O must be formed one for each cluster. The buses submatrices ($I_{b(i)}$, $O_{b(i)}$) that composes the structure I and O were determined by number of buses in the cluster, i.e, if a cluster contains four buses the matrices I and O will be formed by four blocks corresponding each bus in the cluster.

The individual forecasting was made through the same submatrices bocks $I_{b(i)}$ and $O_{b(i)}$ used to create the mapping $(input \times output)$ in aggregate process. The blocks $I_{b(i)}$ and $O_{b(i)}$ are the input and desired output for each bus used to adjust the individually predictor and to execute the forecasting.

To validate the mapping I/O proposed in this paper we choose an artificial neural network (like MLP) [16] and a model based on multiple linear regression [8]. These models were chosen because they are success techniques used in the resolution of the problem of electric load forecasting. As in this paper the goal is the forecasting of the 24 hours of next day we opted to perform a forecasting hour-to-hour, i.e., for every hour of the day the matrices I/O they are updated.

IV. Bus Load Data

In this paper, it was used historical data of 15 buses to test the proposed model. The bus load measurements were done from 06/01/2001 to 10/03/2001 in the Brazilian North/Northeast system. These buses operation voltage varying from 38KV to 230KV.

The bus load data series are represented as $b(i)_{[h,d]}, d = 1, ..., 125; h = 1, ..., 24; i = 1, ..., 15$ where: i is the bus index, d the day index and h the hour index. Table II shows the maximum, minimum, medium, standard deviation and the variation coefficient level (%) to each bus before normalization process.

TABLE II
Bus Load Analysis.

Buses	*Max (Mw)*	*Min (Mw)*	*Average (Mw)*	*Std (Mw)*	*Variation C. (%)*
1	65.50	22.10	35.60	2.42	6.80%
2	2.80	0.70	1.39	0.11	7.91%
3	75.0	16.80	49.03	4.42	9.01%
4	8.70	4.50	7.20	0.71	9.86%
5	160.30	83.10	117.64	10.89	9.26%
6	259.50	122.40	190.29	20.19	10.61%
7	13.40	3.20	9.10	1.41	15.49%
8	22.60	8.50	14.82	1.52	10.26%
9	261.50	89.50	179.57	21.41	11.92%
10	114.70	53.50	77.72	8.56	11.01%
11	53.10	13.10	33.22	5.04	15.17%
12	37.30	11.70	20.92	1.86	8.89%
13	4.00	1.10	2.27	0.23	10.13%
14	55.80	17.80	32.23	5.43	16.85%
15	10.60	1.80	4.81	1.06	22.04%

A. Buses Normalization

The bus load level is very different for distinct buses, for example, the load level of bus 6 is approximately 140 times the load level of the bus 2. In order to eliminate this differences the bus time series were normalized. The objective normalization is to extract consumption profile, independent of the load levels. We used the normalization according to equation 1.

$$\widehat{b(i)}_{[h,d]} = \frac{b(i)_{[h,d]}}{mean(b(i)_{[:,d]})} \tag{1}$$

Fig. 6 shows the normalization effect on load curves in which the first graphic shows the original bus load curve and the second graphic shows the normalized curves.

V. Case Study

This section shown the numerical results of the simulations performed in this work. Initially is present results associates at the clustering process where the main objective is found the buses clusters in some days. Soon afterwards is shown the forecasting results using the bus load method proposed in this work presenting the comments and advantages of the methodology.

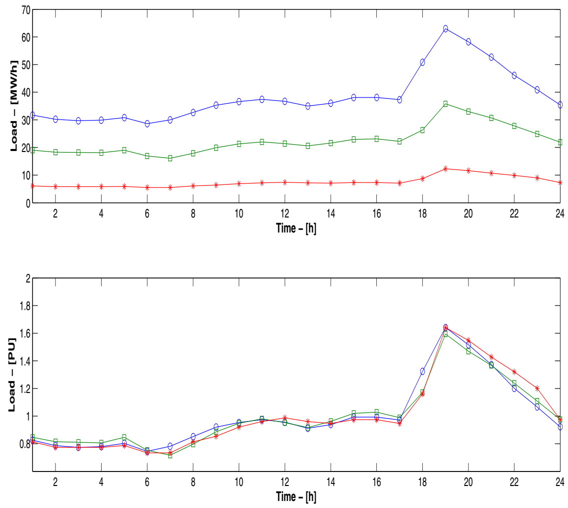

Fig. 6. Buses Normalization Motivation.

A. Clustering Results

The SOM algorithm was configured to clustering the buses in four subsets. This cluster number was chosen based in the typical Brazilian consumption profiles of according to the presented results showed in papers [18]. The SOM algorithm was tested to find the buses on seven consecutive days. The clusters obtained present similar features and in general the clusters had good classifications in relation to the consumption profiles and buses with different profiles are not allocated in the same group. It is important to say when the number of clusters is high, there is more chance to form cluster with a single bus.

The Table III show the clustering results during period of seven days, from Thursday (9/19/2001) to Wednesday (9/25/2001). The clustering results presented are very regular and the bus clusters do not present much variations from one day to another day. The results are compatible with cluster results showed in [18].

TABLE III
CLUSTER BUSES - 9/19/2001 TO 9/25/2001.

	Cluster-1	**Cluster-2**	**Cluster-3**	**Cluster-4**
Wed	1 2 11 13 14	3 5 6 7 8 9 10	4	12 15
Thu	1 2 11 13 14	3 5 6 7 8 9 10	4	12 15
Fri	1 2 3 11 13 14	5 6 7 8 9 10	4	12 15
Sat	1 2 3 6 11 13	5 7 8 9 10 14	4	12 15
Sun	1 2 3 11 10 13 14	5 7 8 6 9	4	12 15
Mon	1 2 10 11 13 14	3 5 6 7 8 9	4	12 15
Tue	1 2 11 13 14	3 5 6 7 8 9 10	4	12 15

In a real power system with great number of the buses to find the ideal number of the cluster it is possible use an algorithm of automatic number cluster detection (like as *subtractive clustering* [5]) or an expert power system operator with great experience to provide the buses cluster information for this model.

B. Forecasting Results

The proposed model was applied on a real power system data composed of fifteen buses from Brazilian North/Northeast system. The training data set was composed by load from 01 June to 19 September of 2001 and the test data set was composed by loads measured from September 20 to 26 of 2001.

The ANN used in this work (individual and aggregate way) is a well-know multilayer perceptron trained with back propagation (gradient method). The number of the neurons in a hidden layer and the moment term was found by an exhausting search in the domain [1,15] and [0, 0.99] respectively, the neurons limits values are founds using Baum-Haussler metric [2]. The initial learning rate initial is 0.9 in each iteration we use the unidimensional line-search to found the next value of the learning rate [7] and [4]. The number of the neurons in the input/output layers depends of the I and O matrix respectively, where the inputs are determined by partial autocorrelation function.

The models used in this work were named: Individual Model by Multiple Linear Regression (IM - MLR), Individual Model by Artificial Neural Network (IM-ANN), Aggregate Model by ANN with 4 clusters (AM - ANN/4 Clusters).

The forecasting performance was evaluated by the *Mean Absolute Percentage Error* ($MAPE(\%)$) between the observed x_j and estimated loads in the prediction $\widehat{x}_j$. *MAPE* is well-known statistical indexes for evaluation of forecast methods, defined by equation 2:

$$MAPE(\%) = \frac{100}{24} \sum_{j=1}^{24} \frac{|x_j - \widehat{x}_j|}{x_j} \tag{2}$$

To verify the models performance, we tested the method to foresee seven consecutive days. Using this forecast approach we can test the model in several situations, including the forecast of types of different days (Sunday to Saturday). Thus we will know which the situations the model was effective and also where the model can be improved.

1) Seven Consecutive Days Forecasting: The models used in this work was configured to foresee seven days from Thursday (9/20/2001) to Wednesday (9/26/2001). In aggregate model the matrices I and O was computed using the 4 clusters the results is showed in Table III. To foresee the hourly load of the next day the clusters obtained for the previous day was used, i.e., to forecast the load of Thursday (9/20/2001) the clusters of Wednesday (9/19/2001) was used. This cluster information was used to compute the matrices I and O in the aggregate model.

Analyzing the Fig. 7 we can see the mean error *MAPE* for the forecast of the hourly load curve seven days ahead in the individual and aggregate model. The aggregate models presented competitive performance when compared to the individual bus load forecasting methodology and in several cases its results is better than the individual models.

In terms of the aggregate model we can see that the forecasting results are showed in the following way: **1)** Buses [1, 2, 10 and 13] with error about 0 to 2%; **2)** Buses [2, 3, 5, 6, 8 and 11] with error about 2 to 4%; **3)** Buses [4, 7, 12, 14 and 15] with large error than 4%. We can see that the buses with high forecasting error also show high level of the variation coeficient. If we observe the variation coefficient values presents in the Table II it is possible to see that the buses with high variation coefficient tends present high forecast error.

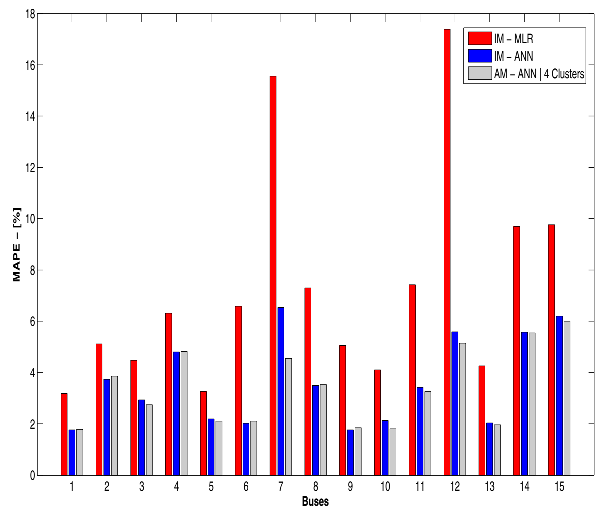

Fig. 7. Average MAPE - Sytem buses.

Analyzing Fig. 8 we can observe that in general the models have a certain stability in terms of the error. The results of the models IM-ANN and AM are compatible during the whole week, with advantage for the aggregate model in most of the days. Only the model IM-MLR did not present good results and the reason of this high error can be associated of the number with the inputs or number of the pattern to adjust the coefficients of the multiple linear regression. In general, for the test period, the model aggregate model showed smaller error with mean MAPE of 3.40%. The results showed in the proposed bus load forecasting model are compatible with the safe/security load operation levels of the Brazilian power system.

According to 8 we can see that the aggregate model was better in 67% of the buses and 85% better in week days, the difference between aggregate and individual model, in terms of results, is very low. In other words, the aggregate model is competitive in relation to an indivdual models in terms of the forecasting performance.

Fig. 9 shows the predicted and actual load curve and the absolute percentage hourly error from the Wednesday 09/26/2001 in the bus 1. It is possible to note that the aggregate and individual ANN model was able to forecast the curve in a efficient way. In terms of the error both models are compatible where IM-ANN has error level of the 2.21% and the aggregate model of the 1.94% in all day.

Table IV show the computational time to foresee the bus load in all the models. To provide equality in the tests all models were running in the unique hardware system using same operational system, softwares and mathematical library.

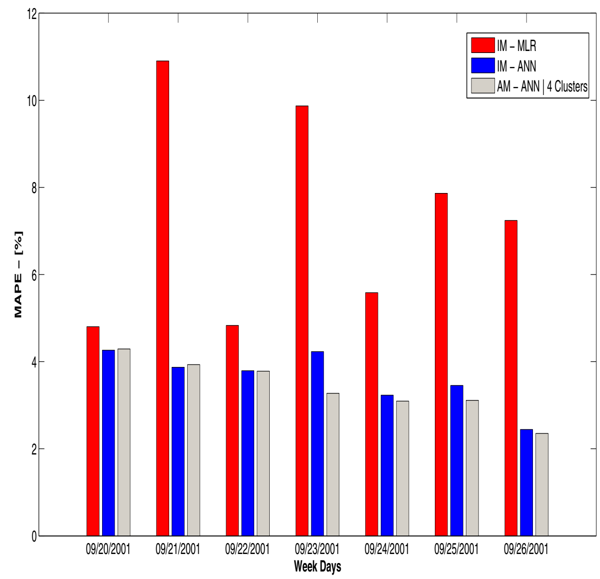

Fig. 8. MAPE Seven Consecutive Days - System buses.

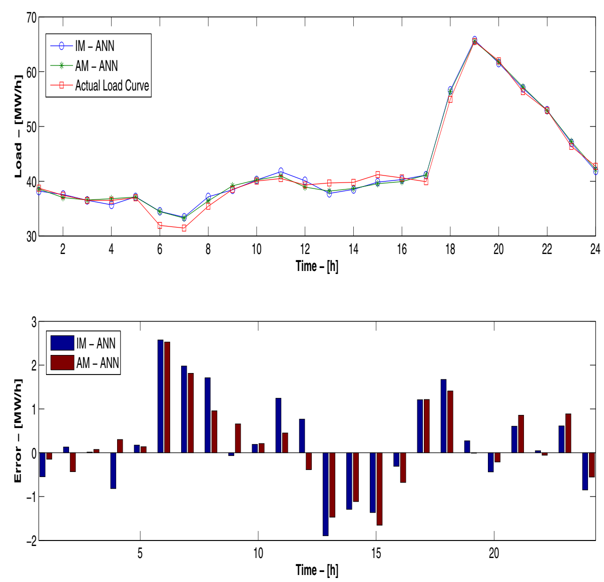

Fig. 9. Actual × Predicted Curve and Hourly Error - Bus 1 — 09/26/2001.

Table IV shows that aggregate model was faster than IN-ANN model. The model IM-MLR was the fastest in all of the approaches, however its results is not good.

TABLE IV
COMPUTATIONAL TIME - FORECAST MODELS.

	IM MLR	**IM ANN**	**AM - ANN 4 Clusters**
Time - [min]	***0.030***	***12.30***	***0.60***

Using the aggregate model the results are so good or better than the individual model with lower computational effort. The aggregate model is about 95% more faster than individual model.

Another advantage present in use matrix I and O it is the possibility of the choose any predictor to predict the bus load forecasting. We chose an ANN/MLP as the main model to forecast the bus load because this model is a classic method in electric load forecast, but without loss of generality it is possible to use any other type predictors, liner or nonlinear, (like as PAR or fuzzy predictors) to learn the relationships of the mapping I and O.

We know that there are other important variables that can improve the quality of the load forecasts (climatic information and others exogenous variables). This work did not use these information because do not have available climatic or exogenous data for the bus area used in this work. The proposed model can use climatic an exogenous information in their structures without change the main idea of aggregate bus data in the sparse matrix I. Is possible to insert these information in each block $I(b(i))$ and to create a input matrix I with climatic influence. We believed that inserting climatic information in the input matrix the results can be improved.

VI. Conclusion and Analysis

In this paper an aggregate bus forecast model was proposed to solve the short-term bus load forecast problem. According to the obtained results the aggregate model presented robustness with good generalization capability to obtain a more accurate forecast. The model was effective in the solution of the problem with low errors in the most buses.

When we compared to the individual model, the aggregate model was robust presenting better performance in 67% of the analyzed buses. In other words, the aggregate model was compatible with the individual model presenting advantages in the time and computational effort.

Even with efficient results the proposed model can be improved to present forecasts more efficiently. We can add climatic-related variables and other exogenous information that tends to improve the bus load forecasting. Thus, this technique emerges as a promising alternative to deal with short-term bus load forecasting problems. Perspectives for further research include a deeper comparative analysis with other proposals for short-term bus load forecasting, and an extension of the proposed methodology to handle other prediction tasks.

Acknowledgment

This research was supported by the Research Foundation of the State of São Paulo (FAPESP) and the Brazilian National Research Council (CNPq).

References

[1] N. Amjady, "Short-term bus load forecasting of power systems by a new hybrid method," *Proceedings of the IEEE Transactions on Power Systems*, vol. 22, pp. 333–341, February 2007.

[2] E. B. Baum and D. Haussler, "What size net gives valid generalization?" *Neurocomputing*, vol. 1, pp. 151–160, 1989.

[3] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems*, 2nd ed. New York: Plenum, 1996.

[4] M. Bromberg and T. S. Chang, "One dimensional global optimization using linear lower bounds," *C. A. Floudas & P. M. Pardalos (Eds.) - Recent advances in global optimization*, pp. 200–220, 1992.

[5] S. Chiu, "Fuzzy model identification based on cluster estimationk," *Journal of Intelligent & Fuzzy Systems*, vol. 2, no. 3, Sept 1994.

[6] M. Espinoza, C. Joye, R. Belmans, and B. DeMoor, "Short-term load forecasting, profile identification, and customer segmentation: A methodology based on periodic time series," *Proceedings of the IEEE Transactions on Power Systems*, vol. 3, no. 20, pp. 1622–1630, August 2005.

[7] C. d. Groot and D. Würtz, "Plain backpropagation and advanced optimization algorithms: A comparative study," *Neurocomputing, NEUCOM 291*, vol. 6, pp. 153–161, 1994.

[8] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558–1573, December 1987.

[9] E. Handschin and C. Dornemann, "Bus load modelling and forecasting," *Proceedings of the IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 627–633, May 1988.

[10] S. Haykin, *Neural Networks, A Comprehensive Foundation*. Macmillan Publishing Company, 1999.

[11] S. Kaski and T. Kohonen, "Exploratory data analysis by the self-organizing map: Structures of welfare and poverty in the world," *Proceedings of the Third International Conference on Neural Networks in the Capital Markets*, pp. 498–507, 1995.

[12] A. Keyhani and S. Miri, "On-line weather sensitive and industrial group bus load forecasting for microprocessor-based application," *Proceedings of the IEEE Transactions on Power Apparatus and Systems*, vol. PAS-102, no. 12, pp. 3868–3876, December 1983.

[13] H. R. K. A. Keyhani, T. Woung, and M. Rahman, "A hybrid fuzzy, neural network bus load modeling and prediction," *Proceedings of the IEEE Transactions on Power Systems*, vol. 14, no. 2, May 1999.

[14] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Proceedings of the Biological Cyberneticys*, vol. 43, pp. 59–69, 1982.

[15] G. Lambert-Torres, L. E. B. da Silva, B. Valiquette, H. Greiss, and D. Mukhedkar, "A fuzzy knowledge-based system for bus load forecasting," *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 1211–1218, March: 8-12 1992.

[16] D. C. Park, El-Sharkawi, M. A., and R. J. Marks II, "Eletric load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 442–449, May 1991.

[17] A. J. R. Reis and A. P. A. da Silva, "Feature extraction via multiresolution analysis for short-term load forecasting," *Proceedings of the IEEE Transactions on Systems*, vol. 20, no. 1, pp. 189–198, February 2005.

[18] R. M. Salgado, T. Ohishi, and R. Ballini, "Clustering bus load curves," *Proceedings of the IEEE-PES Power Systems Conference and Exposition*, vol. 3, pp. 1251 – 1256, October: 10-13 2004.

[19] A. K. Sinha and J. K. Mandal, "Dynamic state estimator using ann based bus load prediction," *Proceedings of the IEEE Transactions on Systems*, vol. 14, no. 4, pp. 1219–1225, November 1999.

[20] ——, "Hierarchical dynamic state estimator using ann based dynamic load prediction," *Proceedings of the IEE Generation, Transmission and Distribution*, vol. 146, no. 6, pp. 541–549, November 1999.

[21] R. L. Sullivan, *Power System Planning*. New York: McGraw-Hill Inc, 1977.

[22] L. Wenyuan and R. Bi1linton, "Effect of bus load uncertainty and correlation in composite system adequacy evaluation," *Proceedings of the IEEE Transactions on Systems*, vol. 6, no. 4, pp. 1522–1529, November 1991.