ARTIFICIAL NEURAL NETWORK BASED LOSSES MODEL DUE TO ENERGY EXCHANGE

J. M. ORTEGA, HELDER Z. MAIA, FAETE J. T. FILHO, J. O. P. PINTO, L. MULLER

Electrical engineering department, Federal University of Mato Grosso do Sul
Cidade Universitária ¤ Caixa Postal 549 ¤ CEP: 79070-900 Campo Grande, MS – Brasil
E-mail: jmortega@del.ufms.br; helderzmaia@batlab.ufms.br, ffilho@utk.edu, joaonofre@gmail.com, lucas.muller@enbr.com.br

Abstract— Due to the characteristic of the Interconnected Brazilian National System and to the location of Mato Grosso do Sul state, the connection between the South and Southern power systems has become a fragile link, making this state susceptible to significant changes, as far as losses and damages are concerned, when submitted to different levels of energy exchanges between those regions. The determination of these losses permits the reevaluation of the costs, which are normally transferred to the consumer. This work presents a model that allows the determination of the so called, technical losses related only to the energy internally used by the Mato Grosso do Sul State, and the one related to the energy exchange between South and Southern regions. The work presents the methodology of the elaboration of database and the neural network modeling. The results show that the model does match with the actual physical system.

Keywords — Artificial Neural Networks, Technical Losses, Principal Component Analysis.

1 Introduction

The Mato Grosso do Sul State Electric Power System is the power exchange path between the power systems of the Southeast and South regions of the Brazilian National Interconnected Power System. Those regions constitute the stronger power systems in the country, both in generation capability and load concentration. Furthermore, there is an active and dynamic change of the direction and intensity of the power flux between these systems.

This constitutes a significant source of losses to the Mato Grosso do Sul State. Owing to market regulations in Brazil the costs associated with the exchange losses are charged only to the state. The proposition of a policy of sharing of the costs between the systems make necessary to develop a methodology capable to establish what parcel of the total losses is consequence of the power exchange.

In this context, the development of one such methodology is presented along with the motivations for the use of the specific techniques, its implicit assumptions, and the obtained results.

The input data consists of the total monthly energy expenditure in each bus of the power system in the period from 2002 to 2008. These measurement data permits to establish the total system losses.

Tough, to determine how much of the losses were due to the power exchange between the neighboring regions was necessary to develop a suitable model. The approach chose was to use the official software package to simulate the power flux condition for each month to separate the proper losses from the exchange ones.

These model's outputs were fed to a Multiple Layers Perceptron (MLP) neural network. The MLP role was to infer from the data the relationship between the expenditure in each system bus and the losses.

Due to reasons to be discussed later, the direct application of the ANN's (Artificial Neural Network) was unfitting. The use of a data conditioning technique was necessary. The specific techique employed was the Principal Component Analysis (PCA).

2 Data set acquisition

The data set was provided by the utility company, which allowed to evaluate and determine the critical conditions of the system and also the losses due to exchange energy between south and southeast. However, the data available was from its new topology and it did not reflect the system configuration from 2002 to 2008. To correct this it was needed to adapt the topology to reflect the same characteristics at the point of measurement. The starting point was the Enersul system loads that were a data measured and available for analysis for the state of Mato Grosso do Sul (MS). The data set with each substation MWh information from the Enersul system was used to start the research using two parallel paths: one to elaborate a dataset for the separated losses from the utility and those from energy exchange and a second line of study to evaluate Artificial Neural Networks (ANN) performance and its applicability to this project.

2.1 Enersul losses and exchange losses data set determination.

The main idea used here is that the losses can be separated in two type of losses based on the problem definition. Given a certain condition with no exchange losses between South and Southeast systems there are losses that occur due to the Enersul system itself. In this case if it is added a power flow to exchange energy between South and Southeast systems then there will be an increase in the energy losses due to this exchange. This increment in losses is due to the exchange itself.

To obtain a dataset with Enersul losses and exchange losses it was elaborated case studies for each month where a simulation software were used to simulate the operation point for that month and then the result were compared with the actual data available from measurement to separate what is from Enersul from what is due to exchange losses. The simulation software provides results that consider a scenario where no exchange losses exists so if the software is correctly adjusted for each case then the losses can be separated within a precision range.

Before the simulation process can be initiated it is needed to adjust the system configuration (load, line, topology, etc) to run the power flow. To determine the topology configuration a historical it was used data of the utility's configuration along the years that was recorded since 2002 and also help of the utility's experts. Using this approach it was possible to setup the system configuration for each month in the range from 2002 to 2008 for power flow simulation in the software ANAREDE.

2.2 ANN performance study and applicability.

While the approach to separate the losses was applied in parallel an Artificial Neural Network (ANN) approach was also tested in order to determine if a data based model could be found. Data based model have some advantages as they can be based in a model ruled by the intrinsic linear and/or non-linear properties of the dataset. In this case it is also possible to used a ANN model to separate the losses without doing assumptions.

ANN's are one of many artificial intelligence techniques that are based on data. It utilizes mathematical approaches to model neurons and simulate brain behavior in decision dependent and learning applications. Once the ANN topology is defined the next step is to properly train the ANN in order to have a correct input/output mapping. This step is called training where the ANN learns from the dataset to do input/output mapping with minimum error to the problem in study.

At the end of this process the ANN has established a good model, which is based on the input/output data. However, a generalization test needs to be applied to the ANN to evaluate its performance. This procedure is done by separating the data in two sets that are used for training and then for testing in order to evaluate the ANN performance using data never presented before to the ANN. Using this procedure it is possible to detect over fitting and analyze ANN.

For the case of Multilayer Perceptrons (MLP) it is possible to show that they can represent any kind of real function of multiple input/output variables if the number of neurons and layers are correctly adjusted [2].

However, this is only possible if enough and consistent data is available. The MLPs need an amount of data large enough for a good performance that is proportional to the number of inputs and outputs, which determine the dimension of the problem. The more complex is the mapping the more data will be needed for a good performance, which will also require a bigger ANN (with more neurons and layers).

After a first training with the raw data set a generalization error of 49% was observed. This error can be attributed to an insufficient number of data as well as to corrupted data.

As a consequence it was needed first to do an analysis of the raw data to determine if either a preprocessing was required or a bigger data set. A preliminary data analysis has shown that most of the substations were following collinear trends that could be observed in the data and is presented in Fig. 1. As it can be seen in this figure the substations have a similar trend but different amplitudes. This redundancy in the data set actually decreases ANN performance.

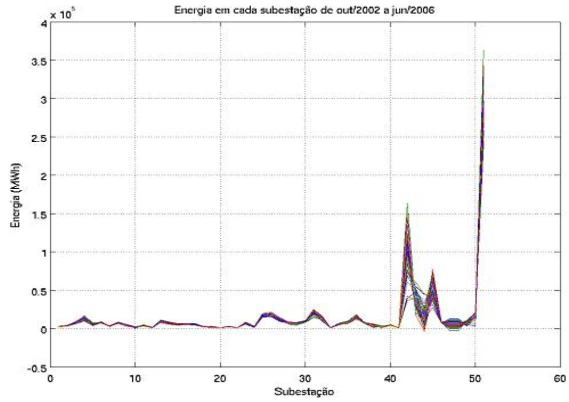

Fig. 1- Substation's energy over a 45 month range.

This feature found in the dataset brought attention to a statistic tool called Principal Component Analysis (PCA). This tool transforms the original data to orthogonal components showing which variables are important and those who are redundant. By using PCA as a data pre-processing technique the amount of substations used can be greatly reduced which is the same as a reduction in the problem's dimension as shown in Fig. 2. In this figure it is possible to note that 15 buses in the system have 95% of all information of the original dataset of 45 buses. Also, 90% of the information in the original data set can be reduced to a data set with only seven inputs.

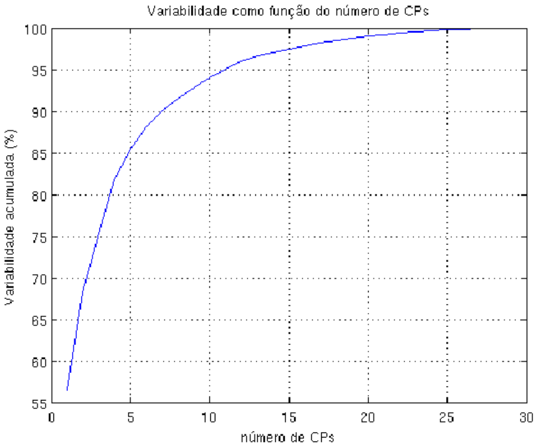

Fig. 2 – Cumulative data information along the principal components.

PCA has shown its applicability in this case by reducing the number of inputs needed. This affected ANN performance by decreasing generalization error to 3.6% (2.8% standard deviation). Also, there was no need to increase the data set for ANN training.

3 Results

An artificial intelligence based simulation model was developed on MATLAB environment and tested using real data provided by the electric utility company Enersul. The performance parameters were derived from the model to validate the proposed methodology.

The ANN training and generalization performance tests were done based on a data set available from October of 2002 until June of 2008. This data set has the Enersul system monthly consumption in MWh from its substations, the respective total percent losses and utility simulated losses. Simulated losses were obtained using ANAREDE according to the cases previously presented.

As it can be seen in Fig. 3 there was a considerable change in the system total losses and Enersul losses due to a change of the place were the measurements were done in March of 2005 (28th month). As a consequence the data set were split in order to facilitate the ANN training and produce more accurate results.

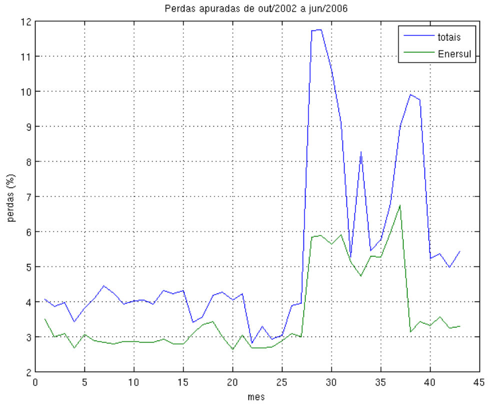

Fig. 3 – Total and Enersul losses profile.

For each one of the periods two ANN's were trained and validated. One was trained to estimate total losses and other to estimate Enersul losses. For each case a Principal Component Analysis (PCA) were done in order to determine the number of components to be utilized, topology and parameters in order to minimize the Mean Relative Error (MRE). Fig. 4 shows the ANNs simulated under different stop criteria and training algorithms but under the same topology and same number of components.

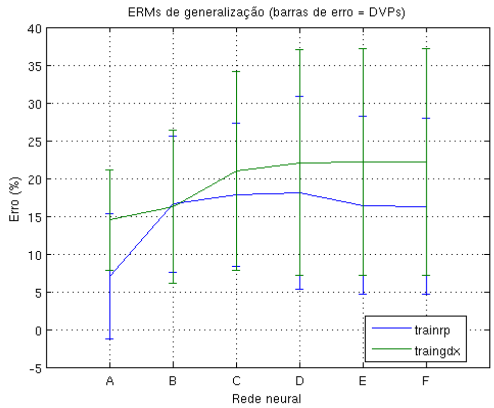

Fig. 4 – Neural Network performance for total losses generalization test.

It can be observed in Fig. 4 that different conditions will have a considerable impact on the ANN performance. The best result for the generalization test obtained is shown in Fig. 5 for the total losses estimation.

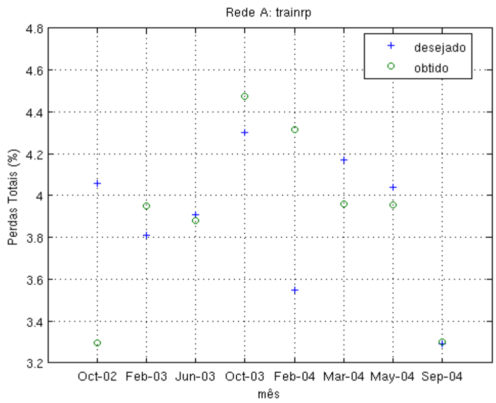

Fig. 5 – Total losses estimation for the ANN with best generalization test.

Table 1 shows the generalization performance for the top ANNs trained for each period. It can be noted that the worst case condition has a mean relative error of 8.87%.

Table 1. ANN's generalization performance.

Period	Losses	MRE (%)	SDV (%)
Oct/2002 to Feb/2005	Total	7,03	8,32
	Enersul	8,87	6,69
Mar/2005 to Jun/2006	Total	7,30	6,52
	Enersul	8,02	6,81

4 Conclusion

This work presented a methodology that allows Artificial Neural Networks to separate the total losses of the Enersul system in the losses that come from the energy exchange between the south and the Southeast of the country from the losses coming from Enersul itself.

The worst-case error obtained by the ANN was 8.87% with a 6.69% standard deviation. Considering the amount of data available and data points that do not follow the trends the results have shown to be consistent. The generalization tests show that the relative error is higher only for points that have an unusual behavior. For the remaining points the error is considerably low. Those unusual points worsen the error and as a consequence the generalization performance is reduced.

References

[01] S. Haykin. Redes Neurais, Princípio e Prática. Bookman, 2001.

[02] G. Cybenko. Approximation by superpositions of a sigmoidal function. Mathematics of control, Signal and Systems, 1989, 2:303-314.

[03] I. T. Jolliffe, Principal Component Analysis. Springer, 2002