Applications of Neural Networks and Decision Trees to Energy Management System Functions

E. Borioli, E. Ciapessoni, D. Cirio, E. Gaglioti
Power System Development Dept.
ERSE
Milano, Italy
Enrico.Borioli@erse-web.it

***Abstract*— Artificial neural network (ANN) and decision tree (DT) applications are proposed to contribute to control room Energy Management System (EMS) functions respectively aimed to assess critical conditions leading to inter-area oscillations, to evaluate the loading margin, and to support state estimation. To facilitate the design, the setting of the parameters, the training of ANNs and the growing of DTs, a tool has been developed within Matlab environment, that allows the user to design and train in an effective way ANNs and DTs with no need to write any procedure or line of code. The tool has been applied in the analysis of the above problems providing a valid support in training and testing both ANNs and DTs.**

***Keywords*- Neural Networks, Decision Trees, Oscillatory Stability Assessment, Loading Margin, State Estimation**

I. Introduction

In order to take advantage of the many opportunities offered by the open electricity market and the increasing penetration of renewable energy sources, the network must assure long distance power transfers from regions with surplus of competitive generation to regions that are short of it. To this aim it is very important that all the security-related operations be performed in a fast and reliable way to prevent instability phenomena thus guaranteeing the integrity of the system. Therefore it is necessary to have very fast Energy Management System (EMS) functions and in particular Dynamic Security Assessment (DSA) tools helping in taking the proper decisions. This requires the use of innovative tools to run in parallel with the traditional ones, thus obtaining the information more rapidly, or to support them by carrying out specific sub-functions e.g. in order to increase the overall speed. From this point of view Artificial Intelligence (AI) may represent a valid alternative in making fast on-line analyses [1] [2] [3]. In the frame of the activities promoted by the Italian RdS[1] fund this need has been emphasized, thus leading to the application of AI tools, like neural networks and decision trees, to three well-known problems with power systems, namely: to prevent critical conditions leading to inter-area oscillations in transferring large amounts of power; to determine the loading margin; to support the state estimation process. To facilitate the design and the training of the ANNs and the growing of the DTs and to optimize the choice of their parameters, a tool has been developed in the Matlab environment, that allows the user to operate efficiently the required analyses and processes with no need to write computer procedures or lines of code, thus saving time and resources.

The paper presents the main features of the tool in Section II. Section III describes the AI approaches adopted to deal with the operational problems mentioned above, and the results obtained by the AI tools trained with ARTIST. Conclusions are reported in the end.

II. "ARTIST" TOOL

The application of artificial neural networks requires to define the network structure and key parameters like the number of hidden layers and nodes, training algorithm, transfer functions, learning coefficient, training error, maximum number of epochs etc. At the time being there is not a well sound rule to identify the design solution that guarantees the best performances, for a given application. As a consequence, several different combinations of settings should be tested, because ANN performances may significantly change. This activity, however, requires considerable efforts and time.

Matlab environment is provided with a specific ANN toolbox, containing a large number of algorithms and functions as well as a Graphical User Interface (GUI) utility to build user friendly interfaces. Nevertheless, there is the need to develop specific procedures e.g. to manage I/O files and function calls, draw plots, perform statistical calculations etc. These needs suggested to develop a specific tool, named ARTIST (ARTificial Intelligence Software Tool), that supports the overall process of ANN and DT design, training, selection, and final application, by means of proper functions managed by a graphic interface. Specific analyses and sensitivity studies are also supported. Fig. 1 shows ARTIST's user interface.

[1] This work has been financed by the Research Fund for the Italian Electrical System under the Contract Agreement between CESI RICERCA and the Ministry of Economic Development - General Directorate for Energy and Mining Resources stipulated on June 21, 2007 in compliance with the Decree n.73 of June 18, 2007

One major feature of the tool is the creation of several ANNs, obtained by the automatic iteration of the training process, in order to select the ANN with best performances.

In fact, the training outcome of an ANN depends, with all other parameters fixed, on the initial values of the synaptic weights and biases. The training algorithm tries to minimize the mean squared error between the outputs calculated by the ANN and the corresponding target values by exploring, after each epoch, the error surface and searching for the greater gradient of reduction in the mean square error. The error surface may have many local minima, therefore there is the possibility to fall into one of these minima with no possibility to get out of it, instead of falling into the absolute minimum. Starting with a different set of weights and biases, the training will proceed along another direction coming to different results. Usually, initial weights are assigned using random numbers that change at every instance of the training process. Therefore, it is useful to repeat the training more than once and to select the one that provides the best results. For this purpose, ARTIST gives the possibility to repeat automatically the training for a number of times defined by the user, starting each time with a different set of initial weights. At the end, it is possible to select automatically the best training as the one presenting the minimum error in the testing set. Shuffling of training data is also possible, as it may affect the training process outcome.

Looking at Fig. 1, the left part of the interface includes the parameters and all the information required to define the structure of the ANN or the DT, such as the number of hidden layers and nodes, the training algorithm and parameters, the transfer functions, the number of epochs, the target training error, the location of the training and testing data files, as well as the ability to enable advanced processing functions.

The latter include the Principal Component Analysis (PCA) and the K-means analysis, both aimed to reduce the number of input variables without loosing accuracy. PCA indicates a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables named principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. Therefore, it is possible to consider only the components that account for the most of the variability neglecting the remaining whose account is negligible. K-means clustering is an algorithm aimed to classify or to group objects into k groups (clusters) obtained by minimizing the sum of the squares of the distances between the data and the corresponding cluster centroid.

III. APPLICATIONS

In the following, applications to significant power system analysis problems are described, and main results are presented.

A. *Oscillatory stability*

The wide extension and heavy exploitation of interconnections among power systems has several economic and operational advantages, but also drawbacks because disturbances may spread thoroughout the system. Further, disturbances may arise, specifically related to the dimensionality of the system: low frequency (0.1÷1 Hz), weakly damped inter-area oscillations may appear, associated to synchronous machines swinging between each other in different areas of the system, with related large excursions of power and voltage, eventually leading to protection intervention and widespread disturbances [4].

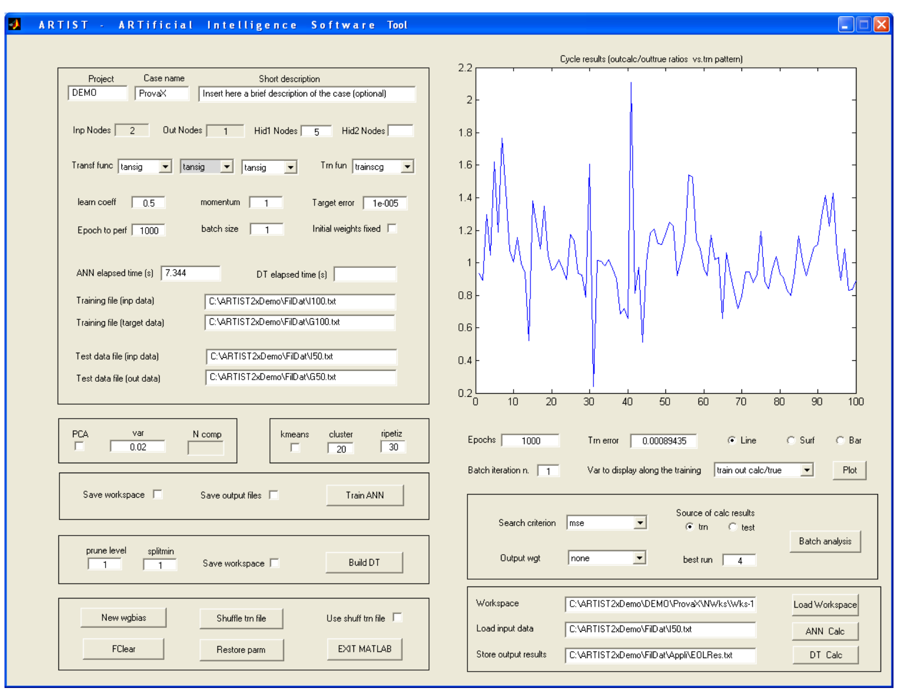

Figure 1. ARTIST user interface

The oscillatory stability of a power system is its capability to maintain a stable operating point when subjected to a small disturbance, such as load fluctuations. A system is stable when all the eigenvalues associated with its state matrix (both real σ_i ,and complex $\sigma_j \pm \omega_j$) have negative real part, weakly stable if some of the eigenvalues have negative real part and some have it null, unstable if at least one of the eigenvalues has positive real part. The damping factor of oscillatory mode j is defined as $\zeta_j = -\sigma_j / \sqrt{\sigma_j^2 + \omega_j^2}$ while the oscillation frequency is $f_j = \omega_j / 2\pi$. Oscillatory stability implies positive damping of all oscillatory modes. Generally a power system is considered sufficiently damped when all the corresponding eigenvalues have a damping factor not below 5%, otherwise it is considered insufficiently damped.

For large networks, oscillatory stability assessment is a complex and time consuming process consisting of the linearization of the operating point and the computation of the eigenvalues. Further, the use of on-line tools may be complicate as the Transmission System Operators (TSO) usually exchange only part of the information needed to make accurate evaluations. Artificial intelligence tools, instead, are fast running and may provide a valid help for the evaluation of the oscillatory stability [5], [6]. Application of ANNs to oscillatory stability assessment has therefore been investigated.

Among the approaches proposed in the literature, the work [5] suggests a mapping of the eigenvalues instead of their direct calculation. To this aim, an observation area and a number of "sampling points" are defined in the region of the complex plane, characterised by critical damping. For each sampling point, an "activation" is defined as the sum of contributions from neighbouring eigenvalues. The contribution of each eigenvalue depends on its distance from the sampling point [5]. The values of the activations are used to train the ANNs. When trained, the ANNs provide the values of the activation function at the sampling points, hence the predicted position of the eigenvalues. The sampling points defined in [5] are equally spaced according to Cartesian variables (i.e. respectively along the σ and the ω axes).

In the work herein described, different sampling alternatives have been considered. Recalling that the iso-damping curves consist of straight lines passing through the origin of the (σ, f), and hence also of the (σ, ω) plane, a radial sampling seemed particularly suitable to predict damping. Therefore an approach similar to the previous one, but with sampling points equally spaced according to polar coordinates, was adopted. Two different solutions were tested, with activation function depending either on the distance between the eigenvalues and the sampling points, calculated in terms of polar coordinates components, or on the number of the eigenvalues falling in a region (circular sector). Finally, also the direct calculation of the eigenvalues has been considered, results are reported here.

The power system considered is shown in Fig. 2 [4] and consists of 16 generators, 68 nodes and 86 longitudinal components (lines and transformers).

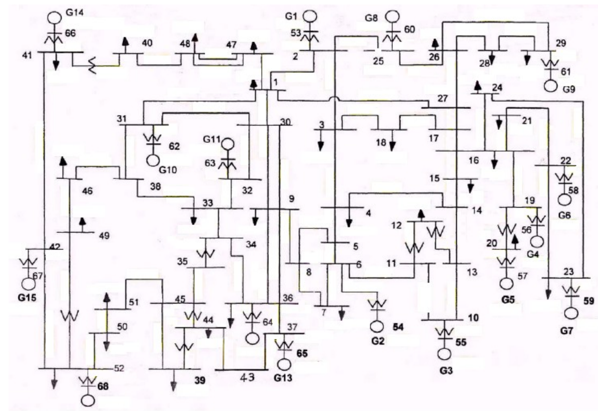

Figure 2. One-Line Diagram of the test Power System

In order to create the data required for training and testing, a procedure has been set up to automatically generate many situations starting from a base case. The procedure has been coupled with the load flow and modal analysis computation engines of Matlab PST Toolbox [7]. Modal analysis results are assumed as the reference for oscillatory stability evaluation with the ANN training and testing stages.

From the whole set of cases analyzed a subset of 10000 samples has been selected, part of them correspond to sound network and part to N-1 and N-2 contingencies. In the following, only the main results obtained with the direct calculation of the real and imaginary part of the eigenvalues are reported.

Input data have been selected as a set of quantities mostly representative of the small-signal dynamics behaviour, such as voltage phase angles, active power flows, according to engineering judgment.

For a power system with n synchronous machines, there exist $n-1$ electromechanical oscillation modes. Since each mode is defined by two parameters, σ and ω, or alternatively by ζ and f, it follows that $2\,(n-1)$ quantities need to be predicted. Therefore in the considered power system model, $2 \cdot 15 = 30$ quantities have to be predicted. The output consists of the real and imaginary part of the eigenvalues.

The analyses have been carried out using both ANNs and DTs and ARTIST has been used to design, train and find the best ANN, and to grow the DTs. Besides, a sensitivity study has been performed consisting in many trainings, each one repeated 5÷10 times, starting each time from different values of the weights. Each training differs from the others in the number of hidden nodes, training algorithm, transfer functions and other parameters related to the training algorithm. In fact, after a certain number of epochs the error value decreasing, either referred to the training set or to the test set, was very slow and it was almost impossible to reduce it to very low values. Further, in many cases it appeared that even if the training error was decreasing, the error on one or more output nodes was increasing, hence revealing an overfitting phenomenon. This suggested to limit the number of epochs of the training. The best compromise was obtained in the order of a few tens of thousands epochs.

The DT approach consisted in growing a set of 30 trees, one for each output node, nevertheless it did not require a large amount of time as the growing of DTs needs only a few parameters to assign.

The definition of the ANN parameters value needs a large number of trainings and the time required for each training is higher (1÷2 orders of magnitude) than the time required to grow the trees. Therefore the setting of ANNs requires large amounts of time, especially with a large number of input and output nodes.

The best results obtained from ANNs are summarized in Table I, where the rows represent the frequencies of the oscillations in the range 0.1 ÷ 1 Hz and the columns represent the damping ratio in the range -3% ÷ 5%. The table reports the error defined as the absolute value of difference between the number of elements calculated and expected in each cell divided by the total number of the elements. Values of zero in cells means that the number of elements predicted for that cell is the same as the number of elements expected, otherwise there is an overprediction or an underprediction of number of the elements.

B. *Loading margin*

Given an operating point, the loading margin is the amount of additional load that would cause a voltage collapse, according to a specified pattern of load-generation increase. The loading margin is the most basic and widely accepted index of proximity to voltage collapse. Voltage collapse is a system instability that occurs on power systems when they are heavily loaded, faulted or have reactive power shortages. In a PV curve, the loading margin to the collapse is the change in loading between the operating point and the nose of the curve, as shown in Fig. 3.

Classical methods for voltage stability margin evaluation rely on special power flow techniques (continuation power flow) or simplified dynamic approaches [8]. Because the computation by these kinds of analyses is currently not so fast as desired, solutions based on AI tools are considered attractive [1], [2], [3], [9].

TABLE I
ANN ERROR

ζ	Frequency (Hz) 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
-3%	0	0	0.002	0.007	0	0	0.002	0	0	0
-2%	0	0	0.015	0.037	0.005	0	0	0	0	0
-1%	0	0	0.012	0.042	0.007	0.002	0	0	0	0.002
-0%	0	0	0.012	0.072	0.005	0.005	0.002	0	0	0
1%	0	0	0.037	0.087	0.012	0.025	0.015	0.002	0.002	0.002
2%	0	0.005	0.040	0.097	0.010	0	0.027	0	0	0
3%	0	0.002	0.005	0.020	0.037	0	0.045	0.005	0.007	0.007
4%	0	0.001	0.017	0.372	0.350	0.055	0.165	0.017	0.010	0.020
5%	0	0.027	0.015	0.032	0.025	0.007	0.415	0.037	0.047	0.092

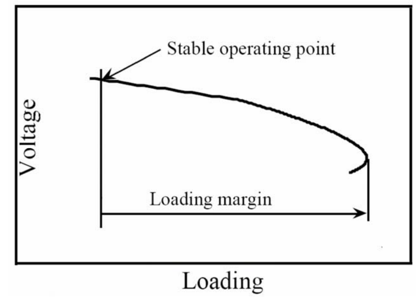

Figure 3. Loading margin

The proposed approach aims to obtain fast indications of the loading margin by ANNs or DTs, based on the values of key quantities throughout the system. Such quantities are identified, taking advantage of the properties of the “pilot nodes” defined within the Area Voltage Regulation. Pilot node voltage is, by definition, representative of the voltage behaviour of the area, hence it is very much correlated with loading margins.

The AI application was based on an archive of operating points of the Italian power system sampled from the field every 15 minutes. Part of these data (700 situations) have been used in the present activity.

From each sample, 54 variables are drawn (i.e. for each of the 18 voltage regulation areas: pilot node voltage, active power unbalance, generator reactive power margin). These are used as input to the ANN. The output is the loading margin computed with a classical Continuation Power Flow tool.

Table II reports the results obtained from the ANN and the DT. The values in the table are presented as the percentage of the test cases whose error falls in one of the three classes 5%, 15%, 25%. From the table it appears that the results provided by the ANN are better than those provided by the DT.

The time required to train the ANN for loading margin problem is much less than the time required for oscillatory stability problem, therefore the influence of the parameters has been analyzed in a deeper detail. Each training has been repeated from 20 to 50 times to select the best one. Networks with one hidden layer provided results better than those with two hidden layers and the best results regarding the hidden nodes were obtained for values in the range 10-15. Therefore the best ANN is characterized by one hidden layer with 10 hidden nodes, training algorithm “scg” (scaled conjugate gradient) and the tan-sigmoid transfer function “tansig”.

As there is only one output variable, this case requires only one decision tree.

TABLE II
LOADING MARGIN – ANN AND DT RESULTS

	% of elements in error classes of ±-2.5%,	±7.5%,	±12.5%
ANN	48.50	93.99	99.14
DT	39.91	84.55	94.85

C. State estimation

The real-time knowledge of the power system operating condition is essential to ensure its safe operation and to avoid dangerous situations like overload and over and undervoltages. To this aim, a basic tool in the EMS environment is state estimation. State estimation consists in the evaluation of the variables that describe the state of the system, i.e. voltage magnitude and phase angle in each node [10]. The process involves measurements that are redundant and imperfect because of noise, malfunctioning of transducers, communication problems, loss of data, etc.

State estimation is a computational-intensive process where algorithms such as Weighted Least Squares (WLS) are applied to large dimension systems. The algorithms may lack to converge under stressed conditions.

An interesting support to state estimation may come from AI techniques, taking advantage of their speed of response and of their property to cope with (limited) missing data. The present approach consists of three sequential steps: (1) pre-filtering of the measured data, (2) reconstruction of the missing data, and (3) state estimation without resorting to conventional algorithms. It must be noticed that steps (1)-(2) can be used to support conventional state estimation procedures, within the pre-processing functions, or as preliminary activities to step (3). The filtering and the noise reduction process in step (1) is performed using a technique based on wavelets. Steps (2) and (3) are based on ANNs and DTs. To perform step (2), i.e. to recover the data that are unreliable or missing, an autoassociative ANN structure has been proposed, as illustrated in Fig. 4 [11], [12].

The first and the third hidden layers of the ANN in Fig. 4 have the same number of nodes; the second hidden layer, named bottleneck layer, has a smaller number of nodes. The input and output have the same number of nodes, as the network is trained to provide in output a replica of the input data. When trained, this network is able to provide a value for input data missing or unreliable, on the basis of the rules underlying the data, learned during the training phase.

Step (3) consists in the evaluation of the state variables, using ANNs, starting from the filtered values of voltage magnitude in each node of the system, and active and reactive power flows at longitudinal terminals (lines, transformers).

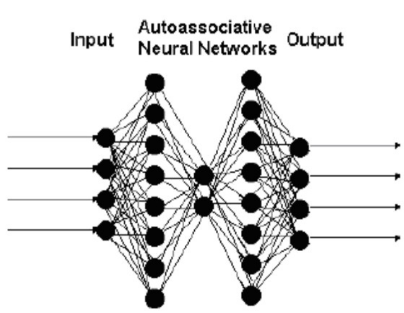

Figure 4. Three hidden layer neural network

The data for this activity have been obtained by running a large number of load flow situations, changing the generation and load power level and distribution, on the same power system model presented in Fig. 2. Of the overall cases, 3000 have been used for the training and 500 as test cases. Then, power flow results have been degraded to simulate measurement inaccuracy by adding on them random errors uniformly distributed, with a mean value of 10%.

Regarding step (2), due to the complexity of the structure that produces a large number of weights, the defined ANN requires a considerable amount of time to be trained and this makes it difficult to run more than a few trainings in a reasonable time. Therefore, a simplified version has also been used with only one hidden layer. The two structures provided similar results and the network with one hidden layer required a much lower time for the training. Then, the capability in evaluating the missing values has been tested setting to zero three input values of V, three values of P, three values of Q and comparing the outputs provided by the ANN with the original values of V, P, Q.

Regarding step (3), an ANN and a DT set have been defined and trained for direct evaluation of the state variables. Using a unique ANN for all of the output nodes proved excessively complex, resulting in unsatisfactory results mainly for the phase angles, also after many different ANN design solutions tried with ARTIST.

The training problem has therefore been simplified by splitting the outputs into two homogeneous sets of variables: one set for the voltages magnitudes and one for the phases. The results obtained in this way were better, thus, to enhance them, a further splitting was decided. Therefore, the groups have been divided into 10 subsets: 3 subsets for the voltages magnitudes and 7 for the phases, associating and training an ANN for each subset. This procedure increased the number of ANNs and consequently the computing time, but provided a good improvement of the results.

Fig. 5 and 6 show the results obtained from the ANN for voltages and phases. For the voltages, they are presented as the histogram of the percent errors between the values calculated and expected; each bar has a width of 0.1 % and the height is proportional to the number of elements contained. For the phases, the results are presented as the difference in degrees between the angles calculated and predicted; each bar has a width of 0.25° and the height is proportional to the number of elements that are contained.

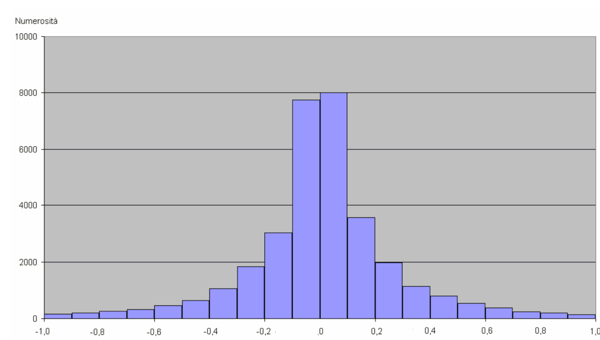

Figure 5. ANN Results for voltages

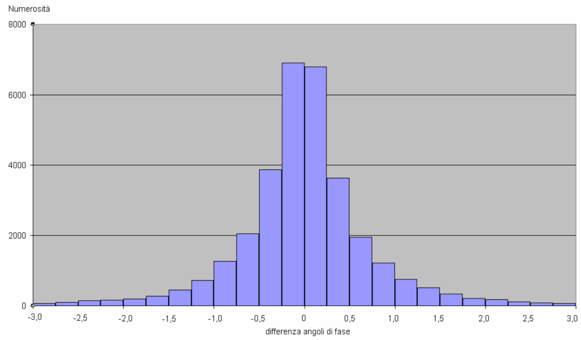

Figure 6. ANN Results for phases

Results provided by the DTs are near the same but their growing required much less time than ANN training.
All the ANNs have been trained with the "scg" algorithm.

IV. CONCLUSIONS

To enhance performances from ANNs several architectures and algorithms need to be tested to identify the best solution for the considered problem. For this reason, the training process should be performed several times, changing every time one or more of the design settings such as ANN architecture and training algorithm. Further, after having set all the other parameters it is a good practice to repeat each training more than one time as the training results change depending on the initial values of the weights, and to select the one that provides the best results. Unfortunately this requires considerable efforts and time, mainly if it is not performed in an automatic way. To support this process, a tool has been developed in Matlab environment, named ARTIST, that allows to easily perform a large number of sensitivity studies for the designing and training of ANNs, and to grow a single DT or a structure of DTs.

ARTIST has been extensively used within several research activities, providing a valid help in the AI-related activities. In particular, applications related to loading margin evaluation, oscillatory stability assessment, and state estimation have been presented. The systematic use of ARTIST emphasized some important points that have impact on the final results. Among the other things, it has been confirmed that performing training with a large number of epochs to reach a small training error does not guarantee to have satisfactory results as it may produce overfitting. To address this problem it is important, during training, to evaluate the behaviour of the error on the test data rather than the error on the training data, keeping track of it to optimize the number of epochs.

The partitioning of the output nodes into groups and the association of a network to each group improved the results but increased the total amount of time required for the training.

The time required to grow DTs is much shorter than the time required to train ANNs.

ANNs and DTs are two independent methods and proved to be valid tools to be used in the analysed problems.

REFERENCES

[1] T. S. Dillon, D. Niebur (editors), "Neural Networks Applications in Power Systems", CRL Publishing, London, 1996

[2] M. Tarafdar Haque, A.M. Kashtiban, "Application of Neural Networks in Power Systems; A Review", Proceedings of World Academy of Science, Engineering and Technology, Volume 6, June, 2005

[3] L. A. Wehenkel, "Automatic Learning Techniques in Power Systems", Kluwer Academic Publisher, 1998

[4] G. Rogers, *Power System Oscillations*, Kluwer Academic Publishers, 2000

[5] S. P. Teeuwsen, I. Erlich and M. A. el-Sharkawi "Small-Signal Stability Assessment based on Advanced Neural Network Methods", IEEE PES General Meeting, Toronto, Canada, July, 2003

[6] S. P. Teeuwsen, I. Erlich and M. A. el-Sharkawi "Neural Network based Classification Method for Small-Signal Stability Assessment", IEEE PowerTech, Bologna, Italy, June, 2003

[7] J. H. Chow and K, W. Cheung, "A toolbox for power system dynamics and control engineering education and research", *IEEE Trans. Power Syst.*, vol. 7, no. 4, pp. 1559–1564, Nov. 1992

[8] T. Van Cutsem, C. Vournas, "Voltage Stability of Electric Power Systems", Kluwer Academic Publishers, 1998

[9] A. J. Cifuentes, C. A. Castro, "Voltage Stability Security Margin Assessment via Artificial Neural Networks", Power Systems Computation Conference (PSCC), 2005, Liége, Belgium

[10] A. Abur, A. Gómez Expósito, "Power System State Estimation: Theory and Implementation", CRC Press, 2004

[11] P. F. Fantoni "A Neuro-Fuzzy Model Applied to a Full Range Signal Validation of PWR Nuclear Power Plant Data", Int. J. General Systems, Vol.29(2), pp. 305÷320, 2000 OPA (Overseas Publishers Association)

[12] G. A. Zanetta, "Sviluppo e Sperimentazione di Metodi Basati su Tecniche di Intelligenza Artificiale per la Validazione e la Riconciliazione Dati nella Misura delle Emissioni Inquinanti", CESI RICERCA A5004836, February 2005 (in Italian)