View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Repositório Aberto da Universidade do Porto

SECURITY EVALUATION OF INTERCONNECTED SYSTEMS WITH LARGE WIND POWER PRODUCTION USING ARTIFICIAL INTELLIGENCE SYSTEMS

J. A. Peças Lopes[1] **Helena Vasconcelos**
INESC Porto – Instituto de Engenharia de Sistemas e Computadores do Porto
FEUP – Faculty of Engineering of the University of Porto
Portugal

KEYWORDS

Interconnected Systems, Wind Generation, Dynamic Behavior, Security Assessment, Artificial Neural Networks, Feature Subset Selection, Regression Trees.

ABSTRACT

This paper presents a new methodology to evaluate, within the framework of on-line security assessment, the dynamic behavior of interconnected power systems having an increased penetration of wind power production. This approach exploits functional knowledge generated off-line, the Regression Tree (RT) automatic learning method to perform Feature Subset Selection (FSS) and Artificial Neural Networks (ANN) to provide a way for fast evaluation of the security degree.

INTRODUCTION

It is well know that, nowadays, the number of cross border power transactions is increasing due to an electricity market liberalization trend. At the same time, the need to decrease CO2 emissions is leading to an increase of wind power penetration and of other renewable dispersed generation (DG) technologies in power systems. The conjunction of these two facts creates, in interconnected systems, an increased use of the main transmission lines that may lead to very stressed operating conditions.

In fact, changes in wind power production may result from unexpected wind speed variations or from sudden disconnection of a large number of wind generators (as well as other DG units), due to the triggering of their protection relays following grid disturbances. Although Automatic Generation Control (AGC) takes care of interchange power flow deviations, it will take some time to eliminate these changes. Moreover, AGC will create a new dispatch solution and therefore will generate a new power flow solution inside the control area. Therefore these wind power disturbances may lead to quasi-steady-state overloads in transmission lines that may provoke a set of undesired cascading events that afterwards may involve load curtailment or even system collapse.

Transmission System Operators (TSO) have been defining the levels of acceptance of wind generation and other DG on the basis of deterministic (n-1) steady-state security studies for worst case scenarios. More recently, TSO have started conducting also dynamic behavior and stability analysis studies following grid disturbances and subsequent operation of DG protection relays [1]. Again, these studies have been conducted for worst case scenarios, leading to severe limitations on system wind generation integration.

In order to increase these acceptable wind penetration limits, interconnected systems with large wind power production require on-line system security assessment tools, able to make prediction of electrical current flow behavior in transmission lines following the occurrence of system disturbances or changes in wind power production, and, based on those predictions, provide preventive control measures if undesired line overload conditions are detected. Such prediction, for current or alternative operating conditions, requires full dynamic simulations of the interconnected system, including AGC operation, which is incompatible namely with the time frame requirements for the management of secondary reserves or the acceptable time period for overload in transmission lines. In fact, in available power systems simulation tools, several minutes are required to obtain some seconds of a large interconnected system dynamic behavior. On the other hand, TSO usually accept no more than 20 minutes of 20% overload in transmission lines. This means that TSO dispatching centers require new tools able to provide fast and accurate forecast of interconnected systems security.

For this purpose, an ANN based approach was designed to emulate a set of security indices, characterizing the level of security of a two area control interconnected system, following a disturbance that provokes the disconnection of a large share of wind power. The main concerns of this approach were to obtain a security structure that can:

[1] e-mail: jpl@fe.up.pt; address: INESC Porto, Campus da FEUP, Rua Dr. Roberto Frias, 378, 4200-465 Porto; Fax: +351 22 209 4150; Phone: +351 22 209 4230.

- provide fast and accurate prediction of system security;
- be exploited to provide preventive control measures, in order to reach system security if insecurity is detected.

In order to gather enough information about the power system behavior, an algorithm able to automatically generate all possible operating conditions for this kind of power systems was developed. This includes a full dynamic simulation stage where system behavior is computed and the variables of interest are kept in a knowledge Data Set (DS). In a next stage, this data was analyzed in order to understand and characterize the security problem. From this stage, the candidate features to be used for ANN inputs and the security indices to be predicted were selected based on engineering judgments. In this procedure, the ANN preventive control purposes introduced some restrictions to the type of input feature to be considered. Namely, controllable variables (like dispatch conditions) were preferred, where non-directly controllable variables (like pre-fault power solution) were avoided.

An automatic FSS procedure was also conducted, based on the RT growing algorithm, in order to eliminate irrelevant features from the ANN input layer.

POWER SYSTEM MODELING

In this research, a test system was created based on the Portuguese – Spanish interconnected system. The single-line diagram of the created system is presented in Figure 2.

Figure 1 presents the installed capacity, for each type of generating power, and the minimum and maximum load values considered for the studied power system.

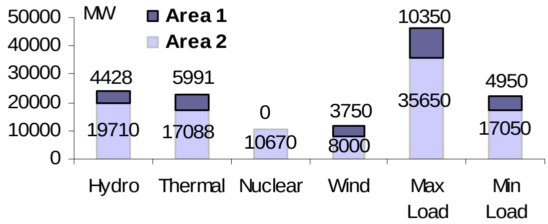


Figure 1 – Considered values for installed capacity and minimum and maximum load

Control area 1 corresponds to an approximation of the Portuguese transmission system, and control area 2 represents an equivalent of the Spanish/European UCTE system. In order to reduce the power system dimension, without losing relevant information, all the generating units of control area 1 are equivalent machines modeling similar generators operating in parallel in the same power plant. Regarding that this research was focused on the security of control area 1, the neighboring system was modeled by one busbar with equivalent generation units. In each control area, thermal, hydro and wind generating units were considered. In control area 2, nuclear units were considered as must run units and not participating in secondary frequency control. All the hydro and thermal units were considered to participate in primary and secondary frequency control.

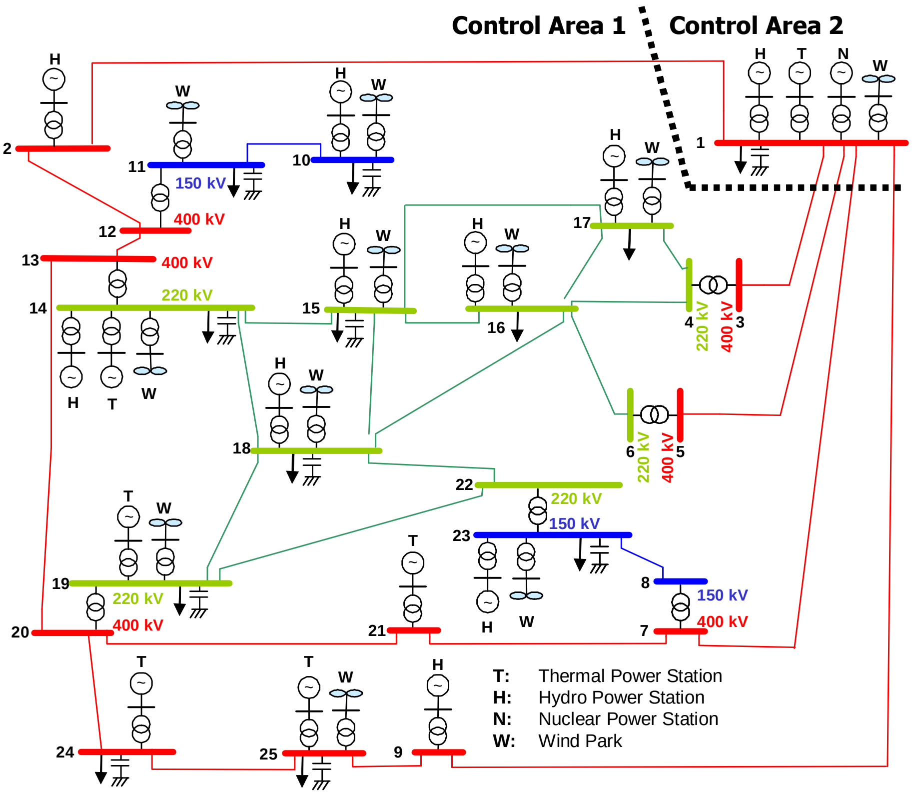


Figure 2 – Single-line diagram of the interconnected transmission system

In order to obtain the dynamic behavior of thermal, hydro and nuclear units, the usual corresponding local frequency regulator models as described in [2] were used, including also the voltage regulator behavior adopting an IEEE1 model type. Wind generators were modeled by a classical third order asynchronous machine model.

The AGC system response is also modeled, adopting the traditional integral control approach and using the configuration described in [3]. According to this configuration, besides keeping the system frequency and interchange power between control areas in the specified value, changes in power production are distributed among generators through participation factors in order to maintain generating units at the most economic operating conditions. For all the parameters of the power system model, typical values were considered, and extracted from the full Portuguese - Spanish interconnected system.

DATA SET GENERATION

A special care was devoted to the DS generation step, in order to cover all possible operating conditions of the power system. In addition, several operating restrictions were mandatory to be included in order to filter out unrealistic scenarios, and therefore decrease computational time without compromising the knowledge data quality. Besides, this also avoids load flow convergence problems, in face of unfeasible conditions sampling. The developed algorithm includes the following main steps:

- Sampling:

  Based on typical operating conditions, a structured Monte Carlo sampling method [4] is applied in order to produce all possible operating scenarios of the system, characterized by different settings of: system load level, wind power production levels, and import levels. These conditions were defined regarding their potential influence on the transmission lines power flow values during pre-fault and post-fault periods.

  Based on available statistical data about wind power production of the Portuguese power system, the following dependencies were considered among wind parks of area 1:

  $$CF(Wind\ Park) = b + CF(Wind\ Production\ in\ area\ 1) \times m \qquad (1)$$

  where CF is the capacity factor, being defined has the power production divided by the installed capacity of the units in operation; m and b are the slope and the y-intercept of the best linear regression relating $CF(Wind\ Park)$ to $CF(Wind\ Production\ in\ area\ 1)$. In order to introduce diversity in the DS, in each sampled scenario, m is randomly sampled between $\pm 3 \times standard\ deviation(m)$ and b between $\pm 3 \times standard\ deviation(b)$.

- Scheduling of conventional units:

  For each sampled scenario, the identification of several unit commitment solutions is performed, based on specified scheduling settings. To obtain a solution, in each control area the units are sequentially connected, until the load supply is satisfied, constrained by:

  - a pre-defined connecting order among power plants of each control area;
  - the minimum and maximum number of available units in each power plant, being the availability of some specified units sampled in order to provide diversity among scheduling solutions;
  - the minimum and maximum technical limits of each unit;
  - a primary control reserve (PR) criteria for each control area, namely:

  $$PR > \text{capacity of the largest unit in operation} \qquad (2)$$

  - a secondary control reserve (SR) criteria for each control area, namely the one presented in [5]:

  $$SR > \sqrt{a \times L\,max + b^2} - b \qquad (3)$$

  being $L\,max$ = maximum estimated load in that period; a = 10 MW and b = 150 MW.

- Dispatch of conventional units:

  For each created scheduling scheme, a dispatch module randomly distributes the insufficiency of power production by the conventional units that were defined to be in operation, considering again their production limits.

- Power Flow:

  For each dispatch solution, a load flow is solved in order to identify all the system pre-fault operating conditions.

- Feasible Steady-State Solution:

  Before starting the dynamic simulation, the feasibility of the power flow solution is checked regarding the minimum and maximum allowed voltage values in the transmission system, and if those limits are violated the voltage of PV synchronous generators is changed.

- Dynamic Simulation:

  For each feasible steady-state operating point, the time simulation of system dynamic behavior is computed in order to characterize the system security following specified disturbances that affect wind power generation.

- Data Set Recording:

  After each dynamic simulation analysis, a pattern is added to the DS, being characterized by all the features needed to describe the system pre-fault operating conditions and the dynamic behavior of power flow in transmission lines after the disturbance. From this set, the most relevant pre-fault features will be used as the ANN input set and the violated post-fault operating conditions will be selected for the ANN output set.

In order to characterize the system security problem, the following condition was analyzed: quasi-steady-state post-fault load in transmission lines, $Ifim$ (120 s after the disturbance, involving therefore AGC operation).

Besides being used for ANN training, recorded features must also enable the analysis of the generated patterns quality regarding the feasibility of the generated operating conditions, namely the following must be observed:

- the steady-state pre-fault and post-fault transmission voltages are within ± 10% the nominal value;
- in pre-fault scenarios, conventional units and transmission lines are not overloaded.

Moreover, the feature set must also include the necessary input data in order to perform the dynamic simulation analysis of the generated scenarios.

- Learning and Testing set:

  Finally, the DS gathered in this way is afterwards randomly divided in two sets, creating a learning set (LS), with 3/5 of the patterns used for training, and a testing set (TS), with the remaining patterns used for overfitting control during training and performance evaluation.

FEATURE SUBSET SELECTION WITH RT

Previous to ANN training, the growing algorithm of a Regression Tree structure, as described in [6], is performed in order to identify the most relevant features for the variation of the electrical current post-fault value in each critical transmission line.

The design of a RT is made by applying a recursive partitioning algorithm, which successively divides the learning knowledge data into mutually exclusive subsets, aiming to minimize knowledge dispersion. Each tree node is divided by the application of a splitting test of the following form:

$$\{feature_k(OP) > u_k\}? \tag{4}$$

where $feature_k(OP)$ is the value of feature k in operating point OP, and u_k is the optimal threshold value for the chosen feature. By applying this test to all the set of operating points in node t, two successor nodes are created, t_L and t_R, which correspond to the two possible instances of the test. This must be performed according to an "optimal" splitting test, which corresponds to the one that most reduces the security index variation in nodes t_L and t_R. By considering the mean value as the predicting function in the tree nodes, the goal becomes into reducing knowledge variance. Therefore, the quality of each splitting test s in node t is measured by the obtained variance reduction, given by:

$$\Delta var(s,t) = var(t) - \frac{N(t_L)}{N(t)} var(t_L) - \frac{N(t_R)}{N(t)} var(t_R) \tag{5}$$

where $N(t)$ is the number of operating points stored in node t, and $var(t)$ is the variance of the security index in node t.

In this research, each feature relevance was measured by the maximum variance reduction, Δvar, obtained for the root node division. The obtained Δvar values are then divided by the variance in the root node, giving therefore a percentage variance reduction. From this ranking, only the features that provided more than 2% of variance reduction were selected for the ANN input set.

This approach was inspired by the procedure presented in [7], where Decision Trees (DT) were suggested to provide a feature ranking regarding its contribution to the total tree information. The main differences of the applied approach to this previous work are the following:

- Regression trees (RT) are used instead of decision trees (DT). In fact, RT provide numerical forecasting, where DT only provide classification forecasting which depends from the considered security boundary. Therefore, RT are more suitable for the security problem under analyzed, since the acceptable over current limits in transmission lines change with weather conditions. This way, the numerical line current forecast can be compared with these limits.
- Only the gain provided for dividing the root node was considered, based on the knowledge that only the first split of a tree structure is optimal. In fact, all the other splits are conditioned by the first split.

The applied FSS method with RT does not identify associations among the tested features. Therefore, the Pearson correlation among each pair of features was calculated in order to identify strong linear relationship among them. For a correlation higher than 0.99, one of the features was removed from the ANN input set in order to remove irrelevant information from the DS.

DESIGN OF THE ANN STRUCTURE

In this research, an ANN based tool was chosen, since it performs generally better than concurrent tools in the fast dynamic security evaluation of power systems [8]. For ANN training, the MATLAB Neural Network Toolbox tool was used [9]. ANN parameters (i.e. the network weights and biases) were found through the Levenberg-Marquardt backpropagation algorithm. Before starting the training stage, training and testing patterns are normalized to have zero mean and a standard deviation of one. To perform overfitting control, besides considering a maximum epochs number, when the testing error increases for a specified number of iterations, the training is stopped. The used ANN was a two-layer feedforward network, with a tan-sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. To choose the number of hidden units, a rule described in [10] was used. According to this rule, the number of hidden units of a single hidden layer network is given by:

$$\frac{N_{LS}}{A} \times \frac{1}{n+r+1} \tag{6}$$

where N_{LS} is the number of learning patterns; A is a constant factor $\in [5;10]$; n is the number of input features; and r is the number of output variables.

NUMERICAL RESULTS

Data Set Results

For the DS generation, the load was considered to change from light load scenarios to peak load scenarios and wind parks from disconnected to connected with a maximum capacitor factor of 0.9. The import level from control area 2 was considered to change from 0 to 1700 MW. For the scheduling and dispatch of the conventional units, 2 different situations were taken into account – a thermal based and a hydro based dispatch scenario – which defined two different connecting order solutions among conventional units inside the scheduling procedure. Also special care was taken in order to generate scenarios with higher spinning reserve mainly provided by hydro power plants. Such approach was adopted having in mind the results presented in [11], where it was concluded that in systems with very high wind power production where secondary control is mainly provided by thermal power plants, the operation of additional or faster secondary control (like pumped hydro production) is required in order not to compromise the quality of generation control.

System security was evaluated relatively to a short-circuit that takes place in a sensitive line of the transmission system (one of the two parallel lines installed between buses 15 and 16 of Figure 2). A duration of 300 ms was considered before the disconnection of the faulty line, leading to the lost of the nearest wind generating units due to the triggering of the under-voltage protection relays (which operate if the voltage drops bellow 0.9 p.u.). The system was considered to be insecure if, 2 minutes after the disturbance, any transmission line current is 20% above the electrical current technical limit. A minimum value was considered for the electrical current limit of transmission lines in thermal dispatch scenarios and a maximum value was considered for the hydro dispatch scenarios.

After applying the data set generation procedure earlier described, 4596 patterns were generated where the 7 transmission lines mentioned in Figure 3 were identified has losing security for some of the generated patterns. A total number of 983 patterns were classified as insecure. Obviously, the connection between bus 15 and 16 is the one with the major number of insecure scenarios, since it loses one of the two parallel installed lines.

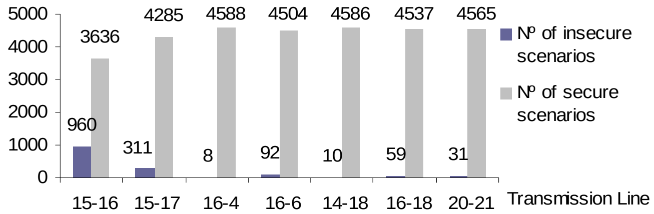


Figure 3 – Number of insecure/secure scenarios in the DS

ANN output set

For the ANN output set, two hypotheses were considered:

- to use only one ANN to predict security for all the critical lines;
- to use a different ANN to predict security for each critical line.

ANN input set

Before ANN training, the generated data was analyzed in order to understand the security problem under analysis. From this analysis the authors of this paper concluded that the security of transmission lines strongly depends from the generation scheduling solution, the interconnection import flows and from wind generation level. Also, the AGC operation avoids system insecurity for much of the generated patterns. In fact, after approximately 20 s the beginning of the disturbance, i.e., when the primary control as already actuated and the AGC is just starting to be activated, 1339 of the generated patterns have transmission lines with more than 20% overload.

From this analysis, 82 features were selected as candidate features for ANN inputs, describing:

- Total active load;
- Dispatch conditions in each conventional power plant of the system;
- Dispatch conditions in each wind park of control area 1;
- Initial pre-fault value of the electrical current in the 7 critical transmission lines.

In order to analyze the influence of removing non-directly controllable variables from the ANN input set, a smaller set with 63 features was selected from the previous one by removing the following variables:

- Dispatch conditions in each conventional power plant of control area 2;
- Initial pre-fault value of the electrical current in the transmission lines.

In the hypothesis of designing a different ANN to predict security for each critical line, the feature subset selection procedure earlier described may be applied. In this paper, the results for the prediction of $Ifim$ for line 15-16 are presented. By applying the FSS procedure to this security problem, another 8 features were eliminated from the input set with 63 variables, remaining, therefore, a set of 55 inputs.

ANN results

The next figures and tables present the obtained testing set prediction errors, from designing a ANN to predict security for the critical line 15-16. The following four different ANN structures were designed:

- $ANNp$: to predict $Ifim$ for line 15-16 with all the 82 candidate features;
- $ANNc$: to predict $Ifim$ for line 15-16 with the 63 directly controllable features;
- $ANNc, fss$: to predict $Ifim$ for line 15-16 with the 55 features selected from the FSS applied procedure;
- $ANNc(Global)$: to predict $Ifim$ for all the 7 critical lines, with the 63 directly controllable features.

Figure 4 presents the obtained regression error (RE) for the trained ANN, namely, the mean squared error divided by the output variance. This figure includes the number of considered input, hidden and output units. The obtained classification errors are presented in Figure 5, namely the following:

- Global Classification Error, given by:

$$\frac{N^o \{OP\ of\ the\ TS\ incorrectly\ class.\}}{N^o \{OP\ of\ the\ TS\}} \times 100\% \tag{7}$$

- False Alarm Error, given by:

$$\frac{N^o\{\text{"secure" OP of TS class. as "insecure"}\}}{N^o\{\text{"secure" OP of the TS}\}}\times 100\% \tag{8}$$

- Missed Alarm Error, given by:

$$\frac{N^o\{\text{"insecure" OP of TS class. as "secure"}\}}{N^o\{\text{"insecure" OP of the TS}\}}\times 100\% \tag{9}$$

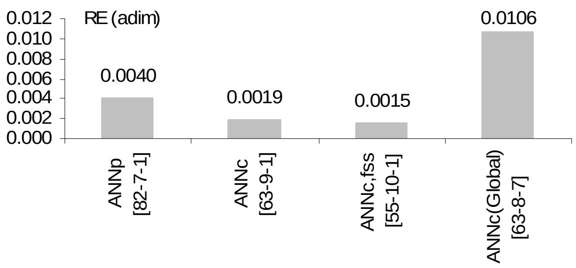
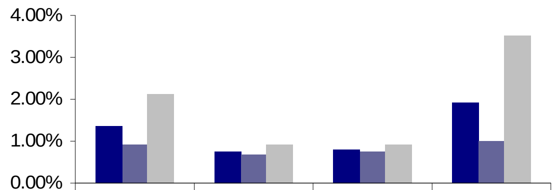


Figure 4 – ANN TS regression errors – prediction of *Ifim* for line 15-16



	ANNp [82-7-1]	ANNc [63-9-1]	ANNc,fss [55-10-1]	ANNc(Global) [63-8-7]
Global (%)	1.36%	0.76%	0.82%	1.90%
False Alarm (%)	0.93%	0.68%	0.76%	1.01%
Missed Alarm (%)	2.14%	0.92%	0.92%	3.51%

Figure 5 – ANN TS classification errors – prediction of *Ifim* for line 15-16

From these results, we may conclude that removing non-directly controllable variables from the input set, besides simplifying the preventive control algorithm that is under development, it provides more accurate predictions of *Ifim* for line 15-16. In addition, the implemented FSS method was capable of increase ANN accuracy, even further. These results also showed that the hypothesis of considering a unique ANN to predict *Ifim* for all the seven critical lines is not the best solution, because an evident increase in prediction error can be observed.

From Figure 6 and Figure 7 we may get a better view of the predicting errors provided by the *ANNc, fss* structure. Namely, from Figure 7 we may see that, in the worst situation, the prediction of *Ifim* for line 15-16 may reach 15% below or 20% above the real value. However, Figure 7 also shows that these are outlier results, and therefore these maximum errors may be reduced by re-training the ANN with additional reproduction of these outliers in the Data Set. Namely, Figure 8 and Table 1 present the obtained error results after re-training the ANN with 20 replicas, in the LS, of each pattern of Figure 7 with more that 0.05 p.u. of prediction error. We may see that, by doing so, the maximum ANN error is reduced to $\pm 10\%$ the nominal current value in summer limit. The only problem is that a slight increase is observed in the total regression error.

In order to choose the best ANN structure for the remaining critical lines, the same type of analysis was implemented.

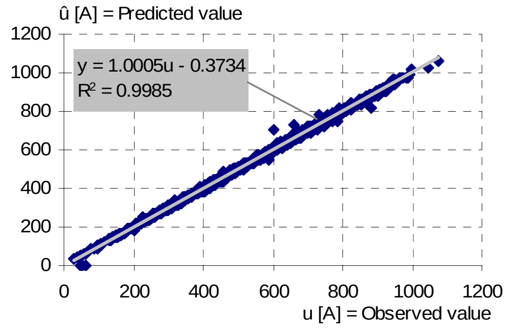


Figure 6 – Linear regression between the predicted values with ANNc,fss and the observed values – prediction of Ifim for line 15-16

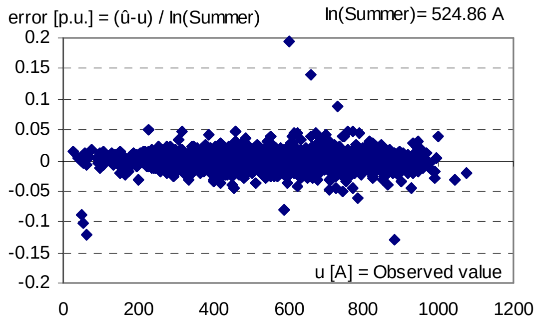


Figure 7 – Error values of Figure 6, in p.u. regarding the nominal current value in summer time

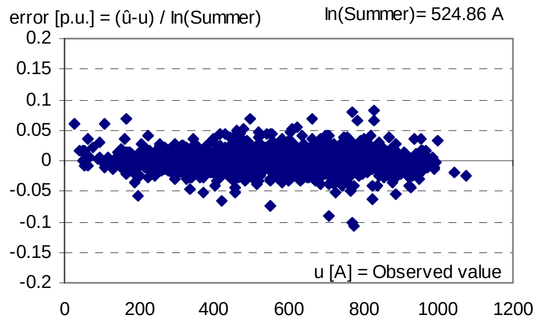


Figure 8 – Error values after re-training the ANN with 20 replicas of each pattern with more that 0.05 p.u. of prediction error

	ANNc,fss, with 20*outliers [55-10-1]
RE (adim)	0.0018
Global (%)	1.09%
False Alarm (%)	0.85%
Missed Alarm (%)	1.53%

Table 1 - ANN TS classification errors – for the perditions presented in Figure 8

CONCLUSIONS

The approach described in this paper develops a new dynamic security assessment concept and provides a new tool able to deal with the impact of the presence, in multi control area systems, of large shares of wind and other DG generation following system disturbances. The reduced testing errors obtained confirm the feasibility and quality of the approach and of the derived security assessment tool. Further research is being developed in order to exploit the ANN structure for the derivation of preventive control measures.

ACKNOWLEDGMENT

The authors would like to thank the financial support of FCT within the "Programa Operacional Sociedade do Conhecimento (POS_C) do III Quadro Comunitário de Apoio".

REFERENCES

[1] IST, REN (the Portuguese TSO), "Transient Stability Study in the Portuguese Transmission Network with High Penetration of Wind Power " (translated from Portuguese), Presented at the CIGRE Symposium ERAC, Paraguay, May 2004.

[2] P. Kundur, "Power System Stability and Control", New York: McGraw-Hill, 1994.

[3] A.J. Wood, B.F. Wollenberg, "Power Generation Operation and Control", ISBN 0-471-58699-4, John Wiley & Sons, Second Edition, 1996.

[4] H. Vasconcelos, "Application of Hybrid Automatic Learning Techniques for Fast Dynamic Security Assessment of Isolated Power Systems with Wind Power Production", Master Thesis, FEUP, 1999.

[5] UCTE Operation Handbook – Policy 1 – Load-Frequency control and Performance, Prelim v1.2, 3/10/2002.

[6] Breiman, L., Friedman, H. F., Olshen, R. A., Stone, C. J., "Classification and Regression Trees", Wadsworth International, 1984.

[7] L. Wehenkel et. al., "Automatic Learning Approaches for On-line Transient Stability Preventive Control of the Hydro-Quebec System – Part I. Decision tree approaches", SIPOWER'95, Mexico, December 1995.

[8] J. N. Fidalgo, J. A. Peças Lopes, V. Miranda, "Neural Networks Applied To Preventive Control Measures For The Dynamic Security Of Isolated Power Systems With Renewables", IEEE Transactions on Power Systems, Vol. 11, No. 4, November 1996.

[9] MathWorks, Inc., "MATLAB® Neural Network Toolbox User's Guide", Version 3.0, 1998.

[10] Louis A. Wehenkel, "Automatic Learning Techniques in Power Systems", Kluwer Academic Publishers, ISBN 0-7923-8068-1, 1998.

[11] G. Dany, "Power Reserve in Interconnected Systems with High Wind Power Production", Proc. Of the IEEE Porto Power Tech., Porto, September 2001.

BIOGRAPHIES

João A. Peças Lopes was born in Portugal in May 1958. He graduated in Electrical Engineering from the Engineering Faculty of Porto University (FEUP) in July 1981, and obtained the Ph.D. and Aggregation degrees also from FEUP in October 1988 and November 1996 respectively. Dr. Peças Lopes is presently Associate Professor with Aggregation at FEUP (http://www.fe.up.pt/) and Assistant Coordinator of the Power Systems unit at INESC Porto (http://power.inescn.pt/).

Maria Helena Vasconcelos was born in Porto, Portugal, on March 19, 1973. She graduated in Electrical Engineering from the Engineering Faculty of Porto University (FEUP) concluded in July 1996, and obtained the MSc. from FEUP in October 1999. Since September 1996 she works as a researcher in the Power System Unit of INESC Porto. Currently, she is a Ph.D. student and Teaching Assistant at FEUP.