Composite Reliability Assessment Based on Monte Carlo Simulation and Artificial Neural Networks

Armando M. Leite da Silva, *Fellow*, *IEEE*, Leonidas C. de Resende, Luiz A. da Fonseca Manso, and V. Miranda, *Fellow*, *IEEE*

***Abstract*– This paper presents a new methodology for reliability evaluation of composite generation and transmission systems, based on non-sequential Monte-Carlo simulation and artificial neural network concepts. Artificial neural network (ANN) techniques are used to classify the operating states during the Monte Carlo sampling. A polynomial network, named Group Method Data Handling (GMDH), is used and the states analyzed during the beginning of the simulation process are adequately selected as input data for training and test sets. Based on this procedure, a great number of *success* states are classified by a simple polynomial function, given by the ANN model, providing significant reductions in the computational cost. Moreover, all types of composite reliability indices (i.e. loss of load probability, frequency, duration and energy/power not supplied) can be assessed not only for the overall system but also for areas and buses. The proposed methodology is applied to the IEEE Reliability Test System (IEEE-RTS), to the IEEE-RTS 96 and to a configuration of the Brazilian South-Southeastern System.**

***Index Terms*–Composite reliability, Monte Carlo simulation, pattern analysis, artificial neural networks, group method data handling.**

I. Introduction

Methodologies based on probability concepts can be extremely useful in assessing the performance of power systems [1]. They have been successfully applied to many areas including generation capacity planning, operating reserve assessment, distribution systems, etc. The application of these methodologies to evaluate composite reliability indices [2]-[10] makes it mandatory to know the status of the operating states of the system, which are defined by the availabilities of the generation and transmission equipment, and by the corresponding load conditions. Usually, a power flow analysis is used to define whether this state is or is not adequate, i.e. if it is success or failure. The computational cost of such analyses depends on several system characteristics, including the network dimension and how rare the failure states are. Thus, the evaluation of a huge number of operating states, especially in reliability studies of large power systems, may become extremely expensive or even unfeasible.

Many works have been proposed to reduce the computational cost of composite reliability evaluation. Some of these techniques are: (*i*) variance reduction [11]; (*ii*) pseudo-simulations [2], [6], [10]; (*iii*) fuzzy optimal power flow [12]; (*iv*) state space pruning [13]; (*v*) parallel and distributed computation [14]. Recently, ANN based on Self Organizing Map (SOM) combined with Monte Carlo simulation has been proposed to evaluate composite reliability [15], [16]. Most recently, these methods were extended to assess adequacy and security calculations [17], [18]. These methodologies classify the operating states of the system into failure or success, through the use of SOM [15]-[17] or fuzzy SOM [17]. A common characteristic among these methods is the fact that the analyses (adequacy and/or security) of the system operating states are directly assessed by those ANNs. The use of such techniques can produce good results for overall or system reliability indices, but they become limited, or even unfeasible from the computational point of view, for assessing composite reliability indices per areas and/or buses. Another constraint of these methodologies is the assessment of power and energy not supplied indices (consequently, loss of load cost indices), for system/area/buses, unless several specific ANNs are built during the simulation process.

Another interesting application of ANN in composite reliability evaluation is described in [19]. This model uses a Multilayer Perceptron trained through a backpropagation strategy. It calculates only system indices, including energy not supplied, but unfortunately it spends huge CPU times, although tolerable, bearing in mind the discussed applications.

This paper presents a new methodology to evaluate the main composite reliability indices. Based on the load model described in [2] and in the *one-step forward transition system* [3], the proposed method combines non-sequential Monte Carlo simulation with an ANN based on Group Method of Data Handling (GMDH) [20], [21]. For such, a pre-classification of the operating states is performed through this ANN type, where the states analyzed during the beginning of the simulation process are selected as input data for training and test sets. The constructive artificial neural network based on the GMDH algorithm performs a pattern analysis to identify only success states (i.e. those states with no load curtailments). With this procedure, a great number of states are classified by a simple polynomial evaluation, providing significant reductions in the computational cost required. The main reason for this choice is that, whenever the classes associated with the training samples are available, models based on supervised learning provide better performance when faced with new cases than others that

This work was partially supported by the National Council for Research and Development – CNPq, Ministry of Science and Technology, by CAPES – Ministry of Education, Brazil, and also by the Science and Technology Foundation – FCT, Portugal

A.M. Leite da Silva and L.C Resende are with the Institute of Electrical Engineering, Federal University of Itajubá – UNIFEI, Brazil.

L.A.F. Manso is with the Electrical Eng. Department, Federal University of São João del-Rei – UFSJ, Brazil.

V. Miranda is with the Institute for Systems and Computer Engineering of Porto – INESC Porto, and with the Faculty of Engineering of the University of Porto – FEUP, Portugal.

use unsupervised learning. Also, the ANNs based on GMDH algorithm usually leads to simpler architectures and faster training processes than it would be obtained with others approaches [15]-[19], [22], [23]. Moreover, due to its self-organized nature, GMDH algorithms have already been successfully applied to the reliability field but in the framework of flexible manufacturing systems [24].

II. Composite Reliability

The estimates of composite reliability indices are obtained through algorithms based on two distinct representations: state space and chronological modeling. Usually, state space based algorithms follow three major steps [4]:

a) Select a system state (i.e. load level, equipment availability, etc.);
b) Analyze the performance of the selected states (i.e. check if available generating units and circuits are able to satisfy the associated load without violating any operating limits; if necessary, activate corrective measures such as generation redispatch, load curtailment, etc.);
c) Estimate reliability indices (i.e. loss of load indices, etc.); if the accuracy of the estimates is acceptable, stop; otherwise go back to step (a).

State enumeration and non-sequential Monte Carlo simulation methods are examples of state space based algorithms, where Markov models are used for both equipment and load states transitions. Therefore, states are selected and evaluated without considering any chronological connection. The necessary steps to evaluate reliability indices considering a chronological representation (sequential Monte-Carlo) are conceptually the same as those described for the state space representation [5]. The basic difference is that the sequential approach moves chronologically through the systems states, while the non-sequential approach selects the states randomly. Sequential simulation, however, requires more computational effort than other approaches [2], [6], [10].

Considering that Monte Carlo simulation is more attractive than state enumeration method to select the system states for large scale systems, composite reliability worth indices have to be assessed by an efficient and accurate Monte Carlo algorithm. Loss of load indices can be estimated by non-sequential simulation techniques, as the mean over N sampled system state values x^k of the test function $F(x^k)$, i.e. [4], [6], [7]:

$$\tilde{E}[F] = 1/N \sum_{k=1}^{N} F(x^k)$$

All the basic reliability indices can be estimated by the previous equation, depending on the definition of the test function F. The estimate uncertainty is given by the variance of the estimator, and this uncertainty is usually represented as the *coefficient of variation* β [4]-[7]. In this paper, the system state is sampled by a non-sequential Monte Carlo simulation, and the frequency and duration indices are estimated through the process called *one-step forward transition system* [3].

The major portion of the computational effort required by the composite reliability algorithm is concentrated on step (b). The objective is to identify if the system, in the sampled state, is capable of meeting the demand without violating its operating limits, from the static point of view. This requires the solution of a power flow algorithm followed by the monitoring of some system variables. If this analysis identifies operating violations or a potential load curtailment, a large-scale optimization problem must be solved in order to avoid or, at least, minimize the necessary load shedding.

Observe that patterns for success states are more easily achieved, due to general system characteristics related with adequate generation and transmission reserve capacities. Conversely, failure states are related with rare events emerging from specific system operation conditions due to, for instance, local transmission constraints. Therefore, it would be necessary to build specific ANNs in order to measure, in terms of reliability indices, all consequences of these failures. Unfortunately, the computational benefits of avoiding several optimization evaluations would be vanished by the computational effort of building several ANNs.

In the next section, it will be shown a methodology which uses a polynomial network (i.e. GMDH) for a faster evaluation of system operating states. Such methodology will be applied to the composite reliability algorithm to reduce the computational effort spent in step (b). The goal is to obtain a pattern only for the success states, since the failure or *at risk* states will be fully analyzed to determine all consequences of load curtailments. This procedure will ensure the assessment of all reliability indices at all levels, i.e. system/area/buses.

III. The Polynomial Network - GMDH

A. Basic Algorithm

The GMDH [20]-[24] can be interpreted as a feed forward neural network with supervised learning that performs a polynomial mapping between input data and the desired output. Each neuron output can be expressed by a second-order polynomial function as shown in Fig. 1, where x_i and x_j are inputs and A, B, C, D, E and F are the polynomial coefficients, which are equivalent to the network weights, and Y is the neuron output value.

During the learning phase, a train and test procedure is used, i.e. two different data sets are employed: one for estimating the network weights (polynomial coefficients), and the other for testing which neurons should survive. This procedure allows network architecture to be determined automatically in the training process. ANN layers are constructed one by one, and each new generated neuron is, in fact, an estimate of the desired output. The training process continues until no better estimates are obtained with the generation of new layers. In this case, only one neuron is saved in the last layer (the one that provides best estimates), and only those neurons that are necessary to generate the output neuron are preserved in the previous layers. An important feature of the GMDH is that only the relevant input variables are preserved in the remaining reduced network.

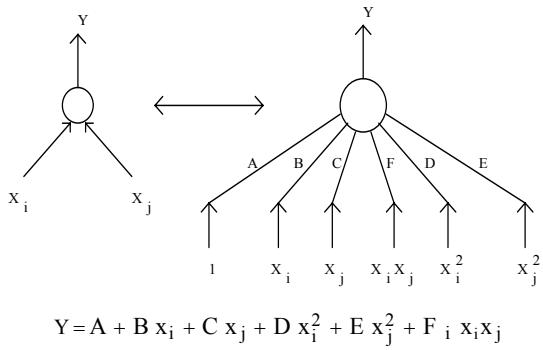

Fig. 1. Neuron model.

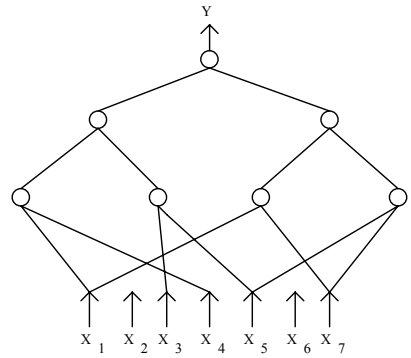

Fig. 2. Polynomial network.

Fig. 2 shows a generic example where 7 input variables were initially presented, but only 5 of them were found to be relevant, for the output class Y, after the training process has been completed. The GMDH is described in Appendix.

B. Proposed GMDH Network

1) Input Variables: The capacity of the neural network to distinguish between success and failure states is based on the fact that, for each group of sampled states, the corresponding variables must present a well defined characteristic pattern. The network should, during the training phase, capture these patterns as to perform correct classifications, when applied to the system or to the new sampled states.

There are many combinations of variables that may be considered as input of the ANN for state classification: e.g. system load; available generation reserve; bus power injections; circuit flows; equipment unavailabilities, etc. In this paper, the following input variables are first adopted: (i) capacity reserve per area; and (ii) unavailable generation per area. Clearly, these input variables are not able to supply information about the transmission network. In order to add such information, the unavailable transmission capacity (in MW) will also be included in the set of input variables of the ANN. The adoption of this variable may look simplistic, since second order or superior transmission contingencies tend to cause more severe impacts to the system. To avoid this problem, and also because transmission lines have low failure rates, the trained ANN will be used *only* in those states in which there is, at most, a transmission contingency. For those states with double or superior transmission contingencies, adequacy will be assessed by the power flow algorithm.

2) Data Selection: The data for network training and testing are obtained from initial states sampled by the Monte Carlo simulation. The size of the data sample can be determined by the regularity criterion [20]-[21], or by the convergence of the coefficient of variation "β", given by (3). In this work, it is used a coefficient $\beta = 20\%$ to define the data set that will classify the success states. Another important aspect is that the ratio between failure and success states tends to be proportional to the rate between failure and success probability. Thus, in systems with low probabilities of failure, which is usually the case, the data sample will have a number of success cases much greater than the number of failure cases. This might make it difficult to capture the correct pattern of the system by the ANN, especially those related to failure states. With the objective of obtaining more balanced samples, data from two success states are selected for each failure state.

To illustrate these concepts, the IEEE-RTS [25] is initially used, although the results of the estimated indices are detailed in Section IV.A. Therefore, for this particular system, when β has reached the value of 20%, data from 396 operating states will have been collected, being 264 of success and 132 of failure. In these data, information on load, generation reserve, equipment availability and states classification (success or failure) are duly captured. Such data are used as input for training and testing of the proposed GMDH neural network. The numerical values 10 and 20 were adopted to represent the desired outputs of the failure and success states, respectively.

For this particular system, it is worth noting that the number of input variables for the ANN was equal to 5 (i.e. unavailable generation capacity per area, generation reserve per area and unavailable transmission capacity). Observe that the IEEE-RTS has two well defined areas, i.e. 138kV and 230kV, where the first one is dominated by load while the second by generation.

3) ANN Performance: Considering the IEEE-RTS, the obtained GMDH network has only one layer. Thus, from the 5 variables selected for training, only two were considered relevant. The visualization of the input/output data of the ANN network is illustrated in Fig. 3. It was assumed that the success states (character "×", being attributed the value of 20, during training) are identified when the classifications obtained are greater than specified threshold. The status of the remaining states, i.e. with classification inferior to the threshold, will be defined by the conventional adequacy analysis (e.g. optimal power flow). In this way, it will be possible to determine not only if the state is of failure but also the amount of load shed. Character "+" represents the failure states (being attributed a value of 10, during training). The success states not classified by the network are under the line of each threshold. Finally, for the threshold 12, it is possible to note that two classification errors were made in this case (character "o"). All performance aspects of the proposed GMDH considering the IEEE-RTS are discussed as follows.

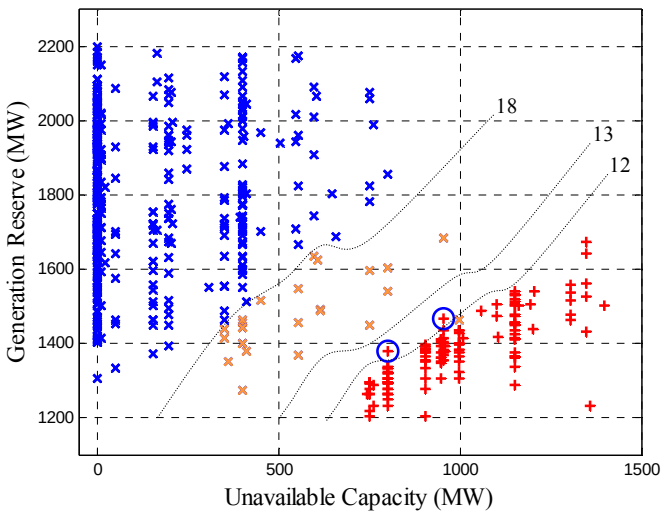

Fig. 3. Input data projection: Unavailable Capacity (area 2) × Generation Reserve (area 2) – thresholds 18, 13 and 12.

Table I presents the performance of the ANN network in the training and test sets, for thresholds (thres.) of 18, 13 and 12, considering the IEEE-RTS. The necessary CPU time for training and testing the proposed GMDH network was inferior to 1 second, in a Pentium IV processor with 2.80 GHz.

The maximum number of correct classifications that the ANN network could achieve is 132 for the training and test, and 264 for the total set (values between brackets). The "correct" classifications are related to the number of success states present in the selected data. Note that the fundamental principle of the proposed methodology is to avoid adequacy analyses only for success states. Therefore, the GMDH network succeeds if a success state is correctly classified, and it does not if a failure state is classified as a success state. Moreover, if success or failure states are not classified, the adequacy analysis (e.g. optimal power flow) will be carried out for these states. Since there are 66 failure states in the selected data set, this should be the minimum number of not classified states in the training and test set, and 132 in the total set. The parameter "Error rate" indicates the percentage of wrong state classifications in relation to 66 or 132, since, as previously mentioned, the ANN network only makes a mistake if a failure state is classified as a success one.

From Table I, note that, with the proposed methodology, 152, 133 and 131 power flow analysis are performed (i.e. states not classified) for thresholds 18, 13 and 12, respectively, contrasting with the 396 analysis that are necessary if the ANN network were not to be used. Thus, one can conclude that the objective for using the network was achieved and, in a simple analysis with the thresholds 13 and 12, about 99.6% of the success states (i.e. 263/264) would be identified by the ANN network. Therefore, this would be the percentage of adequacy analyses of success states avoided in the composite reliability evaluation, once the GMDH network is set (i.e. for β = 20%) until the Monte Carlo simulation reaches the pre-specified convergence (e.g. for β = 5%).

TABLE I
PERFORMANCE OF THE ANN NETWORK – IEEE-RTS SYSTEM

		Training	Test	Total
No. of Selected States		198	198	396
Thres. 18	Correct classifications	126 (132)	118 (132)	244 (264)
	Wrong classifications	0	0	0
	Not classified	72 (66)	80 (66)	152 (132)
	Error rate (%)	0.00	0.00	0.00
Thres. 13	Correct classifications	132 (132)	131 (132)	263 (264)
	Wrong classifications	0	0	0
	Not classified	66 (66)	67 (66)	133 (132)
	Error rate (%)	0.00	0.00	0.00
Thres. 12	Correct classifications	132 (132)	131 (132)	263 (264)
	Wrong classifications	1	1	2
	Not classified	65 (66)	66 (66)	131 (132)
	Error rate (%)	1.51	1.51	1.51

Also note that, in reducing the threshold, the no. of success states identified by the ANN network tends to increase. However, one must take care so that the error rate does not also increase. Undoubtedly, the ANN network proposed to analyze the operating states adequacy has presented very good results for this particular system. This performance analysis will be extended to other systems in the next section.

IV. RESULTS

The systems used to test the proposed methodology are: the IEEE-RTS, the Modified IEEE Reliability Test System (IEEE-MRTS), the IEEE-RTS-96 and a configuration of the Brazilian South-Southeastern System (BSS). The adequacy analysis of each sampled state is performed by DC power flow and by an interior point linear optimization algorithm, whose main objective is to minimize load curtailments. In all tests, a convergence criterion of $\beta \leq 5\%$ (for all system reliability indices) is adopted for the non-sequential Monte Carlo simulation. Besides that, the size of the training and test sets of the ANNs is also defined by coefficient β = 20%. All cases were performed in a Pentium IV processor with 2.80 GHz. All CPU times include only the computational effort required by step (b) of the composite reliability algorithm (i.e. load flow calculation and/or optimization).

A. IEEE-RTS System

Figure 4 shows the diagram of the IEEE-RTS [25], which has 24 buses, 38 circuits and 14 plants (32 generating units). The total installed capacity is 3405 MW. The original hourly load curve with 8736 levels is utilized, with a peak load of 2850 MW. The multi-level non-aggregate Markov load model proposed in [2] is used. Although it is assumed that all the bus loads follow the same pattern as the system load curve, this assumption is not mandatory for the proposed methodology. Different load patterns per area or bus could also be used [3].

Two case studies are performed for the IEEE-RTS system: in the first one (Case 1), the results are presented without using neural networks (reference); in the second one (Case 2), the proposed methodology is used.

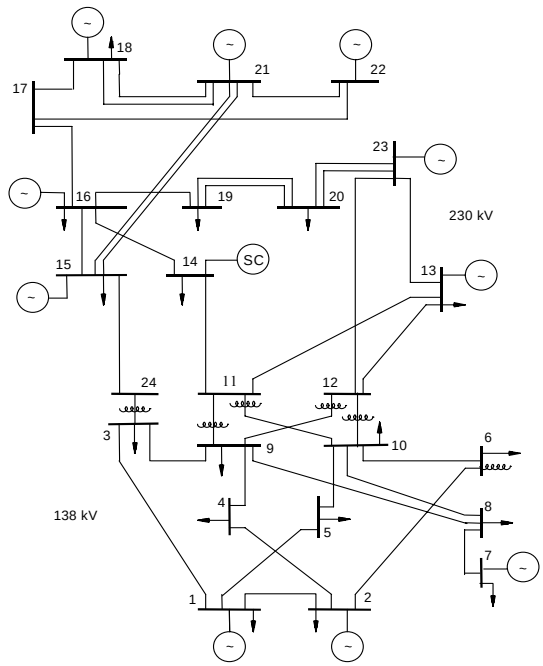

Fig. 4. IEEE RTS.

The performance and characteristics of the resulting neural network obtained (threshold 13) for this system are presented in Section III.B.3. Table II shows for both cases the results achieved for the following system reliability indices: EENS (Expected Energy Not Supplied), LOLP (Loss of Load Probability), LOLF (Loss of Load Frequency) and LOLD (Loss of Load Duration).

For Case 1, the total amount of sampled states was of approximately 1.94×10^6, being necessary to perform 1.93×10^6 power flows and 5083 optimizations (corrective measures such as redispatching and/or load shedding). The CPU time (step-b) necessary for the convergence of the indices was 13.70 minutes. Through Case 2, it is studied the impact of using neural networks in order to avoid adequacy analyses of the success states. The number of sampled states was 1.93×10^6, very close to the one presented in Case 1. However, the total of power flow analysis was 0.15×10^6, a number much lower as compared to Case 1. The number of corrective measures was 4354. Therefore, approximately 1.79×10^6 load flows plus 5083-4354=729 optimizations were replaced by a simple polynomial evaluation. The reduction in static adequacy analyses resulted in the lessening of the CPU time down to 0.91 minutes, which represents a speedup of 15.1 in relation to Case 1. It is possible to note that the average percentage error was 0.46%, staying within the uncertainty margin ($\beta \leq 5\%$) of the indices.

Table III shows the same composite reliability indices for buses 1 and 18, belonging to areas 138kV and 230kV, respectively, considering both cases 1 (without ANN) and 2 (with ANN). As it can be seen, the performance of the proposed GMDH methodology is very good in terms of accuracy, with overall average errors less than 1%.

TABLE II
RELIABILITY COMPOSITE SYSTEM INDICES – IEEE-RTS SYSTEM

Indices:	Case 1	Case 2	Error (%)
EENS (MWh)	1095	1099	0.33
LOLP	0.000998	0.000998	0.02
LOLF (occ./y)	1.97	1.98	0.79
LOLD (h)	4.43	4.40	0.80

TABLE III
RELIABILITY COMPOSITE BUS INDICES – IEEE-RTS SYSTEM

Indices:	Case 1	Case 2	Error (%)
EENS Bus 1 (MWh)	199.4	199.5	0.04
EENS Bus 18 (MWh)	109.0	109.5	0.40
LOLP Bus 1	0.0010	0.0010	0.00
LOLP Bus 18	0.0009	0.0009	0.00
LOLF Bus 1 (occ./y)	1.97	1.98	0.79
LOLF Bus 18 (occ./y)	1.77	1.78	0.46

B. IEEE-MRTS System

The IEEE-MRTS results from modifications made in the IEEE-RTS [25], with the objective of stressing the transmission network. For such, the generation capacity and load are doubled in each bar of the system, providing 6.8 GW of installed power capacity and 5.7 GW of annual load peak. The original hourly load curve with 8736 levels is used. Two study cases are performed: in the first one (Case 3), the results are presented without using the proposed GMDH network; and in the second one (Case 4), the proposed methodology is used.

The set of input/output variables used for training the ANN network in this system is the same of the IEEE-RTS system, apart from the modifications mentioned above. The trained GMDH network contains 3 layers and uses 4 of the 5 variables considered in the training and test sets. An algorithm of data projection in a multidimensional space, such as shown in [23], was used to better understand the clustering tendency of patterns. The performance of this network with threshold 15 is shown in Table IV. The values between brackets for "Correct classifications" and "Not classified" have the same meaning previously explained in Table I. As it can be seen, the total accuracy performance of the proposed GMDH network is again very good, with error percentages of 2%.

Table V shows the results obtained for the two study cases performed. The total of states sampled in the Monte Carlo simulation, the number of power flow analysis and the number of optimizations with corrective measures were 436,790; 423,172 and 23,481, respectively, for Case 3, and 457,172; 46,527 and 19,108, respectively, for Case 4. With the assistance of the GMDH neural network, 376,645 (89.0%) power flow analyses and 4373 (18.6%) corrective measure optimizations were avoided. The CPU times were 2.85 minutes for Case 3, and 0.78 minutes for Case 4, resulting in a speedup of approximately 3.7 in relation to Case 3. The percentage errors are within the uncertainty margin of parameter β (i.e. 5%), and the obtained results showed, again, a good performance of the proposed methodology, even considering very stressful transmission conditions.

TABLE IV
ANN NETWORK PERFORMANCE – IEEE-MRTS SYSTEM

	Training	Test	Total
No. of Selected States	225	225	450
Correct classifications	144 (150)	148 (150)	292 (300)
Wrong classifications	3	0	3
Not classified	78 (75)	77 (75)	155 (150)
Error rate	4.00%	0.00%	2.00%

TABLE V
RELIABILITY COMPOSITE SYSTEM INDICES – IEEE-MRTS SYSTEM

Indices:	Case 3	Case 4	Error (%)
EENS (MWh)	6121	6064	0.93%
LOLP	0.004975	0.004727	4.98%
LOLF (occ./y)	8.73	8.33	4.56%
LOLD (h)	4.98	4.96	0.44%

C. IEEE-RTS-96 System

The IEEE-RTS-96 system (IEEE Reliability Test System 1996) [26] results from modifications performed in the IEEE-RTS. This system has 73 buses, 120 branches and 96 generating units, distributed in 42 power plants, with a total generating capacity of 10.21 GW for a system peak load of 8.55 GW. In order to create a more stressing operating condition for this system, the peak week (week 51) load curve of the IEEE-RTS system is repeated 52 times in a year period.

The neural network obtained is made up of two layers. Four variables were considered relevant from a total of 17 used during the training and test data. The performance of the GMDH neural network obtained for this system is shown in Table VI. The CPU time necessary for training and testing the proposed ANN was also inferior to 1 second, and the threshold adopted was 13.

Table VII shows the results obtained for the two study cases performed. Case 5 presents the results without using the ANN and Case 6 uses the proposed GMDH methodology. As it can be seen in Table VII, the differences presented between the two cases are very small, with an average error of 1.91%, within the uncertainty margin of coefficient β. The total of states sampled in the Monte Carlo simulation, the no. of power flow analysis and the no. of corrective measure optimizations are 1,711,969; 1,717,594 and 4840, respectively, for Case 5, and 1,657,307; 139,699 and 3046, respectively, for Case 6. With the use of the neural network, 1,577,895 (91.9%) power flow analyses and 1794 (37.1%) optimizations were avoided. The CPU times (step-b) were 23.26 minutes for Case 5 and 2.21 minutes for Case 6, resulting in a speedup of 10.6 in relation to Case 5.

Two new study cases were performed with the IEEE-RTS-96 system. In these cases, the original load curve of the IEEE-RTS was used. The obtained results indicated that: the average error in the reliability indices considering the proposed methodology was 4.17%; the CPU times were 143 minutes to reference case and 29 minutes to the case using the proposed GMDH, with a speedup of 4.9. All tests with the IEEE-RTS-96 showed good accuracies for bus indices as well.

TABLE VI
ANN NETWORK PERFORMANCE – IEEE-RTS-96 SYSTEM

	Training	Test	Total
No. of Selected States	316	317	633
Correct classifications	211 (211)	211 (211)	422 (422)
Wrong classifications	4	6	10
Not classified	101 (105)	100 (106)	201 (211)
Error rate	3.81%	5.66%	4.74%

TABLE VII
RELIABILITY COMPOSITE SYSTEM INDICES – IEEE-RTS-96 SYSTEM

Indices:	Case 5	Case 6	Error (%)
EENS (MWh)	1160	1152	0.63%
LOLP	0.000720	0.000714	0.81%
LOLF (occ./y)	2.49	2.58	3,41%
LOLD (h)	2.52	2.42	4,08%

D. BSS System

The configuration used for the Brazilian South-Southeastern System (BSS) contains 413 buses, 685 circuits and 255 generating units. The installed capacity and annual peak load are equal to 46 GW and 41 GW, respectively. This was the configuration used in planning studies during the 90's. A typical annual curve, with 8736 levels, is used to represent the behavior of the hourly load in all buses of the system. The estimated reliability indices refer to the area of the Minas Gerais state which contains 20 buses (15 load buses). The BSS system is subdivided into 33 areas, thus the number of input variables supplied for network training was 67. The neural network obtained for this system is made up of two layers and four variables were considered relevant. The performance of this network is presented in Table VIII. The CPU time for the training and testing of the ANN was inferior to 1 second.

Two case studies were performed. The first one (Case 7) presents the results without using ANNs. The second one (Case 8) uses the proposed GMDH neural network methodology. The results obtained for the BSS system are presented in Table IX. Considering Cases 7 and 8, the total of sampled stated in the Monte Carlo simulation, the number of power flow analysis and the number of optimizations were 484,503; 445,668 and 43,406, respectively, for Case 7 and 505,075; 47,764 and 16,008, respectively, for Case 8. The CPU times were 52.08 minutes for Case 7, and 9.05 minutes for Case 8, resulting in a speedup of 5.8. Table X shows the same composite reliability indices for the two most important buses of this specific area (numbers 1280 and 1293), considering both cases 7 (without ANN) and 8 (with ANN). The performance of the proposed methodology is very good in terms of accuracy, with overall average errors around 2%.

Finally, it is possible to observe that the speed-up, obtained through the proposed methodology, is directly proportional to the computational cost of the adequacy analysis. Moreover, since the optimization cost of corrective measures is greater than the cost of the power flow analysis, greater speedups can be achieved in those systems in which the number of heavy optimization calls are avoided by the use of ANNs.

TABLE VIII
ANN NETWORK PERFORMANCE – BSS SYSTEM

	Training	Test	Total
Number of cases	474	474	948
Correct classifications	308 (316)	312 (316)	620 (632)
Wrong classifications	1	1	2
Not classified	165 (158)	161 (158)	326 (316)
Error rate	0.63%	0.63%	0.63%

TABLE IX
RELIABILITY COMPOSITE AREA INDICES – BSS SYSTEM

Indices:	Case 7	Case 8	Error (%)
EENS (MWh)	1319	1266	4.05%
LOLP	0.010423	0.010119	2.91%
LOLF (occ./y)	11.36	10.85	4.44%
LOLD (h)	8.02	8.15	1.60%

TABLE X
RELIABILITY COMPOSITE BUS INDICES – BSS SYSTEM

Indices:	Case 7	Case 8	Error (%)
EENS Bus 1280 (MWh)	721.5	720.0	0.21
EENS Bus 1293 (MWh)	563.0	560.2	0.50
LOLP Bus 1280	0.0101	0.0101	0.49
LOLP Bus 1293	0.0102	0.0102	0.49
LOLF Bus 1280 (occ./y)	11.18	11.84	5.91
LOLF Bus 1293 (occ./y)	11.27	11.94	5.93

V. FINAL REMARKS

The assessment of composite generation and transmission reliability indices for large power systems is a difficulty task, involving lots of computational calculations. The major portion of these evaluations comes from adequacy analyses, based on power flows and optimizations related with actions taken to correct the system operation conditions.

This paper has presented a new methodology for evaluating the reliability of large composite power systems based on artificial neural networks (ANNs), specifically the *Group Method Data Handling* (GMDH) polynomial network, to reduce the computational effort required during the adequacy analyses of the system operating states. The proposed approach also uses the non-sequential Monte Carlo simulation (MCS), based on the *one-step forward state transition* process and a non-aggregate Markov load model. It appropriately captures, therefore, all equipment and load level transitions.

The proposed GMDH neural network is extremely efficient, from the computational point of view, and it is only used to avoid adequacy analyses of success states, whose patterns are discovered during the first sampling assessments of the MCS. The idea is to provide the simulation process with an *intelligent memory*, based on polynomial parameters, to speed up the evaluation of operating states. Failure states exhibit much more complex patterns and, therefore, are fully evaluated through power flow and/or optimization algorithms. This simple idea allows the assessment of all types of composite reliability indices, including power/energy not supplied and loss of load costs, at all levels, i.e. system, areas and buses.

Case studies with several systems, including the IEEE-RTS, IEEE-RTS-96 and a configuration of the Brazilian South-Southeastern System (BSS), revealed the good performance of the proposed methodology, not only in terms of CPU time savings but also in relation to the accuracy of the reliability indices. Some additional tests applied to the BSS system, which combined the proposed GMDH methodology with a simple and effective network reduction proposed in [27], have shown final speed-ups around 45. This proves the effectiveness of all recent ideas, including ANNs, network reductions, distributed processing, etc. to make the assessment of composite reliability feasible for large power systems.

APPENDIX: THE GMDH ALGORITHM

Consider N input/output patterns divided into two different data sets: N_t samples for training and $N\text{-}N_t$ samples for testing. The GMDH algorithm [20]-[21] is summarized as follows:

(i) Combine the l columns of input variables, two by two, and solve $N_L = l\,(l-1)/2$ linear systems of equations, as shown bellow, in order to generate N_L neurons in the next layer.

$$\begin{aligned} y_1 &= A_{ij} + B_{ij}x_{1i} + C_{ij}x_{1j} + D_{ij}x_{1i}^2 + E_{ij}x_{1j}^2 + F_{ij}x_{1i}x_{1j} \\ &\vdots \\ y_m &= A_{ij} + B_{ij}x_{mi} + C_{ij}x_{mj} + D_{ij}x_{mi}^2 + E_{ij}x_{mj}^2 + F_{ij}x_{mi}x_{mj} \\ &\vdots \\ y_{N_t} &= A_{ij} + B_{ij}x_{N_t i} + C_{ij}x_{N_t j} + D_{ij}x_{N_t i}^2 + E_{ij}x_{N_t j}^2 + F_{ij}x_{N_t i}x_{N_t j} \end{aligned}$$

for $i = 1,\ldots, l$; $j = 1,\ldots, l$; with $i < j$

where y_m, x_{mi} and x_{mj} represents the desired output and the observed input variables x_i and x_j for the mth pattern, respectively. A_{ij}, B_{ij}, C_{ij}, D_{ij}, E_{ij}, and F_{ij} are the polynomial coefficients (weights) to be estimated for the neuron generated by x_i and x_j.

(ii) Evaluate the coefficients in step (i) (in the sense of least squares) and estimate the generated neurons outputs z as:

$$z_{mk} = \hat{A}_{ij} + \hat{B}_{ij}x_{mi} + \hat{C}_{ij}x_{mj} + \hat{D}_{ij}x_{mi}^2 + \hat{E}_{ij}x_{mj}^2 + \hat{F}_{ij}x_{mi}x_{mj}$$

for $i = 1,\ldots, l$; $j = 1,\ldots, l$; $i < j$; $m = 1,\ldots,N$ and $k = 1,\ldots,N_L$

(iii) Evaluate each of the generated neurons using the test set.

$$r_k^2 = \sum_{m=1}^{N-N_t}(y_m - z_{mk})^2 \Bigg/ \sum_{m=1}^{N-N_t} y_m^2 \;; \qquad k = 1,\ldots,N_L$$

(iv) Sort the r_k^2 values. Compare the minimum value with the minimum r_k^2 of the previous layer. If the current error is greater than the previous one, stop, keeping only the best neuron in the previous layer and the neurons that were necessary to generate it, from the original inputs. Otherwise, save the l fittest neurons, transform their outputs into new inputs and go back to step (i) to generate a new layer.

ACKNOWLEDGMENT

The authors would like to thank Prof. Julio Cesar Stacchini de Souza (Fluminense Federal University – UFF) and Prof. Alexandre P. Alves da Silva (Federal University of Rio de Janeiro – COPPE/UFRJ), for their comments, and also to Andrea de Mattos Rei (Brazilian Electric Power Research Center – CEPEL) for providing part of the system data.

REFERENCES

[1] R. Billinton and R.N. Allan, "*Reliability Evaluation of Power Systems*", Plenum Press, New York, 1996.

[2] A.M. Leite da Silva, L.A.F. Manso, J.C.O. Mello and R. Billinton, "Pseudo-chronological simulation for composite reliability analysis with time varying loads", *IEEE Trans. on Power Systems*, vol. 15, no. 1, pp. 73-80, Feb. 2000.

[3] A.M. Leite da Silva, L.C. Resende, L.A.F. Manso and R. Billinton, "Well-being analysis for composite generation and transmission systems", *IEEE Trans. on Power Systems*, vol. 19, no. 4, pp. 1763-1770, Nov. 2004.

[4] M. V. F. Pereira and N. J. Balu, "Composite generation/transmission reliability evaluation", *Proc. of IEEE*, vol. 80, no. 4, pp. 470-491, 1992.

[5] L. Salvaderi, "Monte Carlo simulation techniques in reliability assessment of composite generation and transmission systems", IEEE Tutorial Course 90EH0311-1-PWR, (1990).

[6] J. C. O. Mello, M. V. F. Pereira and A. M. Leite da Silva, "Evaluation of reliability worth in composite system based on pseudo-sequential Monte Carlo simulation", *IEEE Trans. Power Systems*, vol. 9, no. 3, pp. 1318-1326, Aug 1994.

[7] A.C.G. Melo, M.V. Pereira and A.M. Leite da Silva, "Frequency and duration calculations in composite generation and transmission reliability evaluation", *IEEE Trans. on Power Systems*, vol. 7, no. 2, pp. 469-476, May 1992.

[8] A.C.G. Melo, M.V. Pereira and A.M. Leite da Silva, "A conditional probability approach to the calculation of frequency and duration indices in composite reliability evaluations", *IEEE Trans. on Power Systems*, vol. 8, no. 3, pp. 1118-1125, Aug. 1993.

[9] R. Billinton and W. Li, "A system state transition sampling method for composite system reliability evaluation", *IEEE Trans. Power Systems.*, vol. 8, no. 3, pp. 761-770, Aug. 1993.

[10] J.C.O. Mello, A.M. Leite da Silva and M.V.F. Pereira, "Efficient loss-of-load cost evaluation by combined pseudo-sequential and state transition simulation", *IEE Proc. Pt. C*, vol. 144, no. 2, pp. 147-154, Mar. 1997.

[11] G.J. Anders, J. Endrenyi, M.V. Pereira, L.M.V.G. Pinto, C.G. Oliveira and S.H.F. Cunha, "Fast Monte-Carlo simulation techniques for power system reliability studies", in *Proc. 1990 CIGRE Meeting*, Paper 38-205, Paris, September 1990.

[12] J. Tome Saraiva, V. Miranda, L.M.V.G. Pinto, "Generation/ transmission power system reliability evaluation by Monte-Carlo simulation assuming a fuzzy load description", *IEEE Trans. on Power Systems,* vol. 11, no. 2, pp. 690-695, May 1996.

[13] C. Singh and J. Mitra, "Composite system reliability evaluation using state space pruning", *IEEE Trans. on Power Systems*, vol. 12, no. 1, pp. 471-479, Feb. 1997.

[14] C.L.T. Borges, D.M. Falcao, J.C.O. Mello and A.C.G. Melo, "Composite reliability evaluation by sequential Monte Carlo simulation on parallel and distributed processing environments", *IEEE Trans. on Power Systems*, vol. 16, no. 2, pp. 203-209, May 2001.

[15] X. Luo, C. Singh and A.D. Patton, "Loss-of-load state identification using self-organizing map", in *Proc. Power Engineering Society Summer Meeting*, vol. 2, pp. 670-675, 18-22 Jul. 1999.

[16] X. Luo, C. Singh and A.D. Patton, "Power system reliability evaluation using self organizing map", in *Proc. Power Engineering Society Winter Meeting*, vol. 2, pp. 1103-1108, 23-27 Jan. 2000.

[17] Y. Song, G. Bu and R. Zhang, " A fast method probabilistic reliability assessment of bulk power systems using FSOM neural network as system states filters", in *Proc. Transmission and Distribution Conf. and Exhibition: Asia and Pacific*, 2005 IEEE/PES, pp. 1-6, Aug. 2005.

[18] C. Singh, X. Luo and H. Kim, "Power system adequacy and security calculations using Monte Carlo simulation incorporating intelligent system methodology", in *9th Int. Conference on Probability Methods Applied to Power Systems*, Stockholm, Sweden, June 11-15, 2006.

[19] M.Th. Schilling, J.C.S. Souza, A.P. Alves da Silva and M.B. Do Coutto Filho, "Power systems reliability evaluation using neural networks", *Engineering Intelligent Systems*, vol. 9, no. 4, pp. 219-226, Dec. 2001.

[20] A.G. Ivakhnenko, "Polynomial theory of complex systems", *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-1, no. 4, pp. 364-378, Oct. 1971.

[21] S. J. Farlow, "*Self-Organizing Methods in Modeling - GMDH Type Algorithm*", Marcel Dekker, New York, Basel, 1984.

[22] J.C.S. Souza, A.M. Leite da Silva and A.P. Alves da Silva, "Data debugging for real-time power system monitoring based on pattern analysis", *IEEE Trans. on Power Systems*, vol. 11, no. 3, pp. 1592-1599, Aug. 1996.

[23] J.C.S. Souza, A.M. Leite da Silva and A.P. Alves da Silva, "Data visualisation and identification of anomalies in power system state estimation using artificial neural networks", *IEE Proceedings – Generation, Transmission and Distribution*, vol. 144, no. 5, pp. 445-455, Sept. 1997.

[24] P. Yuanidis, M.A. Styblisnki, D.R. Smith and C. Singh, "Reliability modeling of flexible manufacturing systems", *Microelectronics and Reliability*, vol. 34, no. 7, pp. 1203-1220, July 1994.

[25] IEEE APM Subcommittee, "IEEE reliability test system", *IEEE Trans. on PAS*, vol. PAS-99, pp. 2047-2054 Nov./Dec. 1979.

[26] IEEE APM Subcommittee, "IEEE reliability test system – 1996", *IEEE Trans. on PWRS*, vol. 14, no. 3, pp. 1010-1020, Aug. 1996.

[27] A.M. Leite da Silva, L.C. Resende and L.A.F. Manso, "Application of Monte Carlo simulation to well-being analysis of large composite power systems", in *9th Int. Conference on Probability Methods Applied to Power Systems*, Stockholm, Sweden, June 11-15, 2006.

Armando M. Leite da Silva (S'77, M'78, SM'91, F'00) was born in Rio de Janeiro, Brazil, in 1954. He received the B.Sc. degree from the Catholic Univ. of Rio de Janeiro (PUC-Rio) in 1975, the M.Sc. from the Federal Univ. of Rio de Janeiro (COPPE-UFRJ) in 1977, and the Ph.D. degree from Univ. of Manchester (UMIST), U.K., in 1980.

He worked at the Electrical Engineering Department, PUC-Rio, as a Professor until 1994. From 1990 to 1991, he was Visiting Researcher at the Research Division of Ontario Hydro, Canada. From 2003 to 2004, he was a Visiting Researcher at the Power System Unit, INESC Porto, Portugal. Since 1994, he has been a Professor at the Institute of Electrical Eng., Federal University of Itajubá (UNIFEI), Brazil.

Prof. Leite da Silva received the Sebastian Z. de Ferranti Premium Award from the Power Division of the IEE, U.K., in 1992.

Leonidas Chaves de Resende was born in São João Del Rei, Brazil, in 1974. He received the B.Sc. from Federal University of São João del-Rei (UFSJ), in 2000, and M.Sc. and Ph.D. degrees from UNIFEI, in 2002 and 2006, respectively. He is now a post-doctoral researcher, at the Institute of Electrical Eng., UNIFEI, Brazil.

Luiz Antônio da Fonseca Manso was born in Bias Fortes, Brazil, in 1961. He received the B.Sc. degree from the Federal Univ. of Juiz de Fora - UFJF, in 1985, the M.Sc. from PUC-Rio, in 1989 and Ph.D. from UNIFEI, in 1999. Since 1988, he has been an Associate Professor at the Electrical Eng. Department, UFSJ, Brazil.

Vladimiro Miranda (M'90, SM'04, F'05) received the graduation, Ph.D., and Agregado degrees in electrical engineering from the University of Porto, Porto, Portugal in 1977, 1982, and 1991, respectively.

In 1981, he joined the Faculty of Engineering, University of Porto and currently holds the position of Professor Catedrático. He is currently Director of the Institute for Systems and Computer Engineering of Porto, INESC-Porto, Portugal. He has authored many papers and has been responsible for many projects in areas related with the application of computational intelligence to power systems.