Comparison between artificial neural networks algorithms for the estimation of the flashover voltage on insulators

V.T. KONTARGYRI G.J. TSEKOURAS A.A. GIALKETSI P.A. KONTAXIS

School of Electrical and Computer Engineering
National Technical University of Athens
9, Iroon Politechniou Str., GR-157 80, Zografou Campus, Athens
GREECE
vkont@central.ntua.gr, tsekouras_george_j@yahoo.gr, agialketsi@gmail.com, pkont@mail.ntua.gr

Abstract: - This work attempts to apply Artificial Neural Networks in order to estimate the critical flashover voltage on polluted insulators. First, an ANN was constructed in MATLAB and has been trained with several MATLAB training functions. Then, an ANN was constructed in FORTRAN using an adaptive algorithm, in which the parameters of momentum and learning rate changed during the learning procedure, in order to optimize the training process. In each case the Artificial Neural Network uses as input variables the following characteristics of the insulator: the diameter, the height, the creepage distance, the form factor and the equivalent salt deposit density and estimates the critical flashover voltage.

Key-Words: - High voltage insulators, polluted insulators, critical flashover voltage, artificial neural networks

1 Introduction

The critical flashover voltage of a polluted insulator is a significant parameter for the condition of power systems. The reliability of a power system is dependent on environmental and weather conditions, which cause flashovers on polluted insulators, leading to system outages. Therefore several approaches have been developed for the estimation of the flashover voltage on polluted insulators.

The main types of insulator pollution are marine and industrial, as well as the combination of both types. The coexistence of both pollution (marine and/or industrial) and moisture (as dew, fog or drizzle rain) is an unfavorable condition for the operation of insulators. Heavy atmospheric pollution creates an electrolytic layer on the surface of the insulator. When combined with fog or rain, a leakage current flows along the conducting layer. Additionally, surface pollution and non-uniform potential distribution along the insulator surface cause glow discharges or quasi-stable arcs to appear. When the applied voltage exceeds a critical value, these discharges or quasi-stable arcs elongate through successive root formation over the polluted insulator surface until the flashover causes the complete bridging. Therefore, it is important to monitor the insulator's condition so as to ensure that the maintenance takes place in due time.

For this purpose, several researches have been done in which mathematical and physical models are used [1, 2], experiments have been conducted [3, 4] or simulation programmes have been developed [5, 6, 7]. New technologies for the qualitative control of the insulators, such as ANNs and fuzzy logic, are developed. Neural Network algorithms have been successful in estimating the equivalent salt deposit density, by using information regarding temperature, humidity, pressure, rainfall and wind speed as input data, with the intention of establishing an effective maintenance policy.

In the field of high voltage insulators, ANNs can be used to estimate the pollution level [8, 9], to predict a flashover [10, 11], to analyse surface tracking on polluted insulators [12] and also to estimate the critical flashover voltage on a polluted insulator.

This work attempts to utilize the available experimental data and the results of a theoretical approach, in order to construct and train an ANN that can estimate the critical flashover voltage on polluted insulators, using as inputs some characteristics of the insulator.

2 Data collection

Data concerning cap and pin type insulators was used for the training and testing of the ANN. Specifically, the following geometric characteristics were used as input variables: the maximum diameter D_m (in cm) of the insulator, its height H (in cm), the creepage distance on it L (in cm), its form factor F and the layer conductivity σ_s (in μS), while the output variable was the critical flashover voltage U_c (in kV).

The experiments were carried out in an insulator test station, installed in the High Voltage Laboratory of Public Power Corporation's Testing, Research and Standards Center in Athens [13] and according to the IEC standard 507:1991 [14]. In this station tests have been performed on artificially polluted insulators, in order to determine the critical flashover voltage. The pollution was simulated according to the solid layer-cool fog method. Apart from this set of experimental measurements, other measurements were used too, from experiments performed by Zhicheng et al [15] and Sundararajan et al [16].

The mathematical model for the evaluation of the flashover process of a polluted insulator consists of a partial arc spanning over a dry zone and the resistance of the pollution layer in series, as shown in Figure 1, where V_{arc} is the arcing voltage, R_p the resistance of the pollution layer and U a stable voltage supply source.

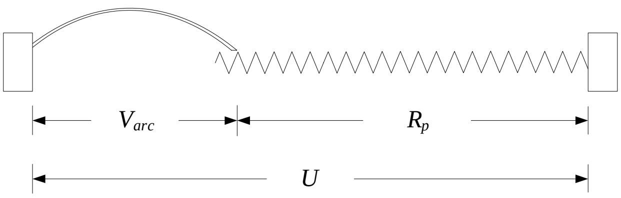

Fig. 1: Equivalent circuit for the evaluation of the flashover voltage.

The critical voltage U_c (in V), which is the applied voltage across the insulator when the partial arc is developed into a complete flashover, is given by the following formula [17]:

$$U_c = \frac{A}{n+1} \cdot \left(L + \pi \cdot n \cdot D_m \cdot F \cdot K\right) \cdot \left(\pi \cdot A \cdot D_m \cdot \sigma_s\right)^{-\left(\frac{n}{n+1}\right)} \quad (1)$$

where L is the creepage distance of the insulator (in cm), D_m is the maximum diameter of the insulator disc (in cm) and F is the form factor. The arc constants A and n have been calculated using a genetic algorithm model [18] and their values are A=124.8 and n=0.409. The surface conductivity σ_s (in Ω^{-1}) is given by the following type:

$$\sigma_s = \left(369.05 \cdot C + 0.42\right) \cdot 10^{-6} \quad (2)$$

where C is the equivalent salt deposit density in mg/cm^2.

The coefficient of the pollution layer resistance K in case of cap-and-pin insulators is given by

$$K = 1 + \frac{n+1}{2 \cdot \pi \cdot F \cdot n} \cdot \ln\left(\frac{L}{2 \cdot \pi \cdot R \cdot F}\right) \quad (3)$$

where R is the radius of the arc foot (in cm) and is given by

$$R = 0.469 \cdot \left(\pi \cdot A \cdot D_m \cdot \sigma_s\right)^{\frac{1}{2 \cdot (n+1)}} \quad (4)$$

3 ANN algorithm in MATLAB

An ANN is usually trained with the error backpropagation algorithm, in which the occurring errors of the output layer return in the input layer to modify the weights. This procedure is repeated until the occurring errors reach accepted values.

In the present work an adaptive ANN has been designed in MATLAB and trained to estimate the critical flashover voltage, when the geometric characteristics of an insulator mentioned above are given. The total number of vectors, which include the input and output variables, was 118. The 80% of these 118 input-output patterns was decided to be used to train the network, while the rest 20% was used to test the function of the network. That means that the training set consisted of 94 vectors and the testing set consisted of 24 vectors.

The ANN was designed through an algorithm that used functions of MATLAB. The data used to train and test the network was set as matrices in different files, which were then called by the program. By changing some parameters in the code of this program, tests could be made with different training methods, in order to see which function gave the best results. The four training methods that were tested were: traingd, traingda, traingdx and trainlm and are all variations of the basic error backpropagation algorithm [19]. With traingd the network is trained according to the gradient descent backpropagation, with traingda the network is also trained according to the gradient descent backpropagation with adaptive learning rate. The function traingdx combines adaptive learning rate with momentum training. An adaptive learning rate is a learning rate that is adjusted according to an algorithm during training to minimize training time. Finally, trainlm uses the Levenberg-Marquardt backpropagation [19].

For each one of those training methods a set of scenarios was taken, in which the parameter that was changing was the number of epochs. One epoch is the presentation of the set of training (input and target) vectors to the network and the calculation of new weights and biases. So, for each training method, there was a set of 10 scenarios, for a change at the number of epochs from 500 to 5000 with a step of 500. In each scenario there was an inner change of the number of neurons (from 2 to 25), in order to find the best architecture for the network, i.e. the number of neurons in the hidden layer that gives the minimum root mean squared error (*RMSE*) that is defined by the following type:

$$RMSE = \sqrt{\frac{1}{m_2}\sum_{i=1}^{m_2} e_k^2(i)} \tag{5}$$

where m_2 is the number of the testing vectors and e_k the absolute difference between the real and the estimated flashover voltage for the testing set.

3.1 Results

For each training function, a 3D graph of the *RMSE* versus the number of neurons and the epochs has been made, in order to see which are the "areas" that minimize the error. Those 3D graphs are presented in Fig. 2 – Fig. 5.

As it is shown by the four figures below, the best results (minimum *RMSE*) are given by traingdx (gradient descent backpropagation with adaptive learning rate and momentum) for 6-12 neurons and approximately 1500 epochs. For over 13 neurons the network becomes unstable, giving very big errors. This was to be expected, as – according to an empiric rule – the number of neurons in the hidden layer should not be greater than the twofold of the input variables. The next step is to define for which value of the learning rate, this network gives minimum *RMSE*. To obtain this, the number of epochs was kept constant, the number of neurons changed from 6 to 12 (because the best results appeared in this area) and using as training method the traingdx, the learning rate altered from 0.1 to 0.9 with a step of 0.05. The minimum *RMSE* appeared for learning rate 0.3, 12 neurons and its value was 0.31. The momentum constant during the whole procedure had the default value 0.9, as defined by MATLAB.

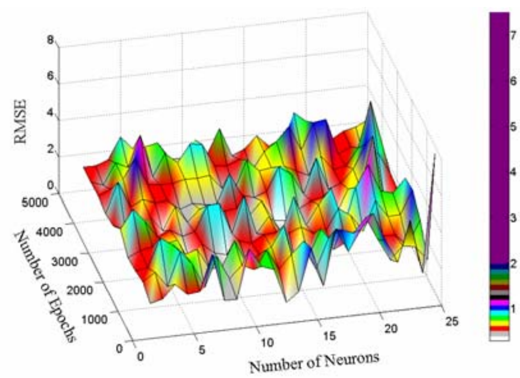

Fig. 2: *RMSE* versus number of neurons and number of epochs for traingd.

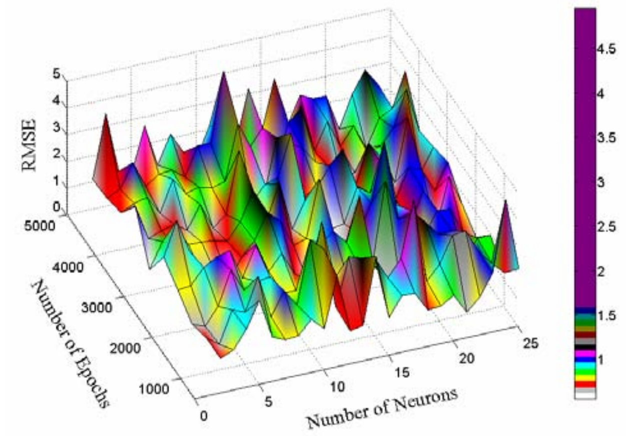

Fig. 3: *RMSE* versus number of neurons and number of epochs for traingda.

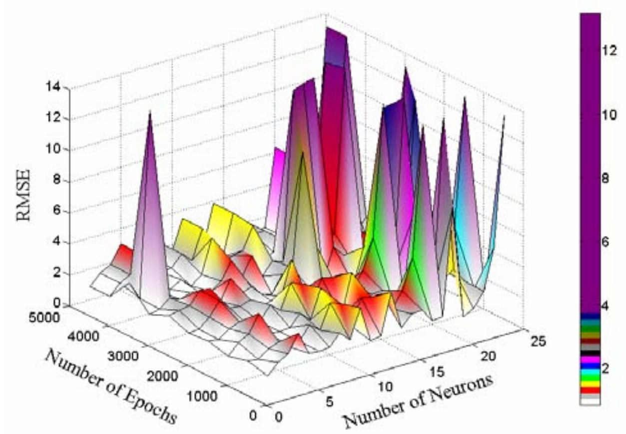

Fig. 4: *RMSE* versus number of neurons and number of epochs for traingdx.

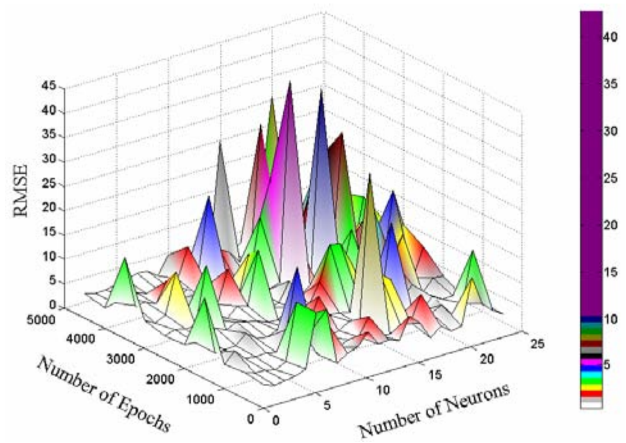

Fig. 5: *RMSE* versus number of neurons and number of epochs for trainlm.

The correlation between real and estimated values of U_c, for learning rate equal to 0.3 is $R^2 = 0.95305$.

ISBN: 978-960-6766-56-5

ISSN: 1790-5109

4 ANN algorithm in FORTRAN

An adaptive ANN has been designed in Digital Fortran and trained to estimate the critical flashover voltage, when given some of the insulator's characteristics.

The main steps of the proposed estimation model are the following:

a. The *N* input variables are selected from the respective database. In this case five parameters (D_m, H, L, F, C) are used as input variables.
b. All variables are properly normalized, in order to avoid saturation phenomena during the training process of the ANN model [20].
c. For each ANN parameter the adaptive back propagation (*a-BP*) algorithm is separately executed for the respective range of values in order the regions with satisfactory results to be identified. According to Kolmogorov's theorem [20] an ANN can solve a problem, using a single hidden layer, if the last one has the proper number of neurons. Under these circumstances one hidden layer is used, however the number of neurons has to be properly selected.
d. Then the *a-BP* algorithm is repeatedly executed, while all parameters are simultaneously adjusted into their respective regions, so as the combination that produces the minimum forecast error for the given evaluation set, is selected.
e. Finally, the flashover voltage is estimated for the under study experiments.

Three points need to be noted:

•*Stopping criteria*: The feed forward and reverse pass calculations are repeated per epoch (one epoch is the presentation of the set of training, input and target, vectors to the network and the calculation of new weights and biases) until the weights are stabilized, or until the respective error function is not further minimized or the maximum number of epochs is reached. If one of the three criteria comes true, the main core of back propagation algorithm comes to an end. Otherwise, the number of epochs is increased by one, the adaption rules are applied and the feed forward and reverse pass calculations are repeated.

•*Validation criteria*: For the evaluation set the root mean square error ($RMSE_{va}$), the mean absolute square error ($MAPE_{va}$) and the correlation (R_{va}^2) can be calculated.

•*Activation function:* The hyperbolic tangent gives better results in this kind of problem. In the case of hyperbolic tangent the unknown parameters are h_1 and h_2, as:

$$f(x) = \tanh(h_1 \cdot x + h_2) \tag{6}$$

In order to converge rapidly, both the training rate and the momentum term are adaptively changed as:

$$\eta(ep) = \begin{cases} \eta(ep-1), & RMSE_{tr}(ep) > RMSE_{tr}(ep-1) \\ \eta(ep-1)\cdot \exp(-1/T_\eta), & RMSE_{tr}(ep) \le RMSE_{tr}(ep-1) \end{cases}$$

$$a(ep) = \begin{cases} a(ep-1), & RMSE_{tr}(ep) \le RMSE_{tr}(ep-1) \\ a(ep-1)\cdot \exp(-1/T_a), & RMSE_{tr}(ep) > RMSE_{tr}(ep-1) \end{cases}$$

where T_η, $\eta_0 = \eta(0)$, T_a, $a_0 = a(0)$ are respectively the time parameters and the initial values of both the training rate and the momentum term and $RMSE_{tr}(ep)$ is the root mean square error for the training set after the end of the ep epoch.

In fact the ANN adapts its parameters according to the error's progress. If $RMSE_{tr}(ep-1)$ is larger than $RMSE_{tr}(ep)$, which means that weights are updated in the correct direction, then it is desired to maintain this direction in the next epoch. This is achieved by decreasing the learning rate and keeping the momentum term constant in the next epoch. Otherwise, if $RMSE_{tr}(ep) > RMSE_{tr}(ep-1)$, which means that the weights are shifted to the opposite direction, it is reasonable to reduce the influence of this direction in the next epoch by decreasing the momentum term and keeping the learning rate constant.

The ANN parameters need to be specified. In order to reduce the combinations, two steps are realized. In the first step, the basic algorithm is executed separately for each parameter's range of values. The program registers the regions where satisfactory results for the current parameter are achieved. In the second step, the main process is repeated for the reduced number of combinations, in which all parameters can take any value of their respective region, as determined in the first step. When this procedure is completed the combination that presents the minimum error in the forecast of the evaluation set is selected.

4.1 Results

The training set consisted of 148 patterns / vectors (of which the 140 vectors had derived from the model and 8 vectors were real values) and the network was tested using 20 patterns (experimental data). The goal is to reduce the number of experiments needed for the operation of the ANN. Using the results produced by the mathematical

model, the ANN can be tested with even less real values.
The first test is the decision of the number of neurons (N). The criterion was the minimization of the $RMSE_{tr}$ for the training set. The minimum $RMSE_{tr}$ appears for six neurons ($RMSE_{tr}$=0.112).
The next step was to define the parameters of the momentum (constant term and time parameter) that lead to minimum $RMSE_{tr}$. Fig. 6 shows a 3D plot of the $RMSE_{tr}$ as a function of the momentum. The constant term of momentum (a_0) changes from 0.1 to 0.9 and the time parameter (T_a) from 500 to 5000. The minimum $RMSE_{tr}$ appears for $a_0 = 0.8$ and $T_a = 4000$.

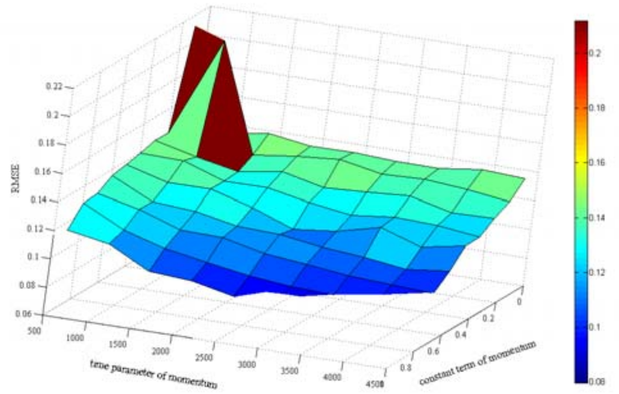

Figure 6: Determination of the constant term and time parameter of the momentum.

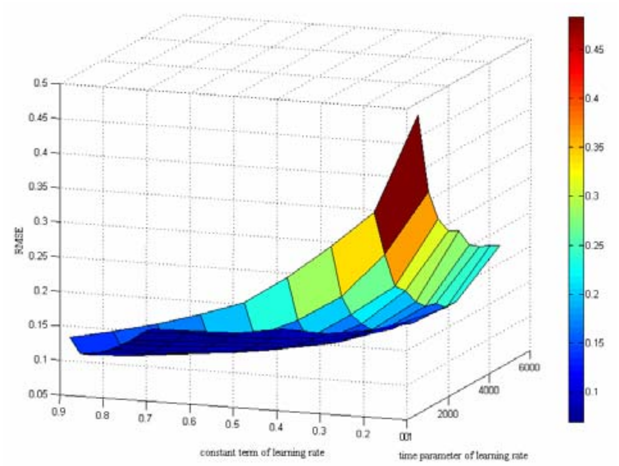

Figure 7: Determination of the constant term and time parameter of the learning rate.

Finally, the parameters of the learning rate (initial value and time parameter) should be determined. For this reason, a third test took place, in which the constant term of the learning rate (η_0) changed from 0.1 to 0.9 and the time parameter (T_n) of the learning rate from 500 to 5000. Fig. 7 shows a 3D plot of the *RMSE* as a function of the learning rate. It is obvious that the minimum $RMSE_{tr}$ appears for $\eta_0 = 0.9$ and $T_n = 4500$. The value of the $RMSE_{tr}$ is now 0.070kV. This error is smaller than the one the network gave before the optimization of the momentum and the learning rate and it shows that the ANN is now capable of estimating the value of the critical flashover voltage very accurately.
The correlation between real and estimated values of U_c is $R^2 = 0.98610$. It must be mentioned that the ideal value for the correlation is 1, so 0.98610 is not only an acceptable value, but also a very good one.

5 Conclusions

In this paper ANNs have been successfully applied for the estimation of the flashover voltage on polluted insulators. The network was trained to estimate the critical flashover voltage when some of the insulator's characteristics are given. The ANN that was designed in FORTRAN gives better results than the ANN that was designed in MATLAB, using fixed functions for the construction of it.
A disadvantage of the ANN in MATLAB is that no one can influence on the learning rate variation upon the time. Also, the momentum is kept constant during the learning process. In the ANN that was constructed in FORTRAN all these parameters could be adjusted to give the best results.

References:

[1] F.A.M. Rizk: "Mathematical models for pollution flashover2, Electra, No. 78, October 1981, pp. 71-103
[2] Z. Aydogmus and M. Cebeci: "A new flashover dynamic model of polluted HV insulators", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 11, No. 4, August 2004, pp. 577-584
[3] C.S. Engelbrecht, R. Hartings, H. Tunell, B. Engström, H. Janssen and R. Hennings: "Pollution tests for coastal conditions on an 800kV composite bushing", IEEE Transactions on Power Delivery, Vol. 18, No. 3, July 2003, pp. 953-959
[4] R. Boudissa, A. Haddad, Z. Sahli, A. Mekhaldi and R. Baersch: "Performance of outdoor insulators under non-uniform pollution conditions", XIVth International Symposium on High Voltage Engineering, China, August 2005, D-51
[5] J.L. Rasolonjanahary, L. Krähenbühl and A. Nicolas: "Computation of electric fields and potential on polluted insulaors using a boundary

element method", IEEE Transactions on Magnetics, Vol. 28, No. 2, March 1992, pp. 1473-1476

[6] C.H. de Tourreil and P.J. Lambeth: "Aging of composite insulators: Simulation by electrical tests", IEEE Transactions on Power Delivery, Vol. 5, No. 3, July 1990, pp. 1558-1567

[7] Y. Cheng, Ch. Li, Ch. Niu and F. Zhang: "Porcelain insulators detection by two dimensions electric field on high voltage transmission lines", XVth International Symposium on High Voltage Engineering, Slovenia, August 2007, T4-495

[8] Ahmad S. Ahmad, P.S. Ghosh, Syed Abdul Kader Aljunid, Hussein Ahmad, Ismail Said, Halil Hussain: "Artificial Neural Network for Contamination Severity Assessment of High Voltage Insulators Under Various Meteorological Conditions", 23-26 September 2001, AUPEC, Perth.

[9] Ahmad S. Ahmad, P.S. Ghosh, S. Shahnawaz Ahmed, Syed Abdul Kader Aljunid: "Assessment of ESDD on high-voltage insulators using artificial neural network", 2004, ELSEVIER Electric Power Systems Research, Vol. 72, Issue 2, pp. 131-136.

[10] P.S. Ghosh, S. Chakravorti and N. Chatterjee: "Estimation of Time-to-flashover Characteristics of Contaminated Electrolytic Surfaces using a Neural Network", December 1995, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 2, No. 6, pp. 1064-1074.

[11] Paul Cline, Will Lannes, Gill Richards: "Use of pollution monitors with a neural network to predict insulator flashover", 1997, ELSEVIER Electric Power Systems Research, Vol. 42, Issue 1, pp. 27-33.

[12] M. Ugur, D.W. Auckland, B.R. Varlow and Z. Emin: "Neural Networks to Analyze Surface Tracking on Solid Insulators", December 1997, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 4, No. 6, pp. 763-766.

[13] K. Ikonomou, G. Katsibokis, G. Panos and I.A. Stathopoulos: "Cool fog tests on artificially polluted insulators", 1987, 5th International Symposium on High Voltage Engineering, Braunschweig, Vol. II, paper 52.13.

[14] IEC 507, "Artificial pollution tests on high-voltage insulators to be used on a.c. systems", 1991.

[15] G. Zhicheng and Z. Renyu: "Calculation of DC and AC flashover voltage of polluted insulators", 1990, IEEE Transactions on Electrical Insulation, Vol. 25, No. 4, pp. 723-729.

[16] R. Sundararajan, N.R. Sadhureddy and R.S Gorur: "Computer-aided design of porcelain insulators under polluted conditions", 1995, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 2, No. 1, pp. 121-127.

[17] F.V. Topalis, I.F. Gonos and I.A Stathopoulos: "Dielectric behavior of polluted porcelain insulators", July 2001, IEE Proceedings Generation Transmission and Distribution, Vol. 148, No. 4, pp. 373-376.

[18] I.F. Gonos, F.V. Topalis and I.A. Stathopoulos: "Genetic algorithm approach to the modelling of polluted insulators", 2002, IEE Proceedings Generation Transmission and Distribution, Vol. 149, No. 3, pp. 373-376.

[19] Matlab Help, Version 6.5.

[20] S. Haykin: "Neural networks: A comprehensive foundation", Prentice Hall, 1994.