Available online at www.sciencedirect.com
ScienceDirect
Electric Power Systems Research 77 (2007) 1532–1540
ELECTRIC POWER SYSTEMS RESEARCH
www.elsevier.com/locate/epsr

Design of an artificial neural network for the estimation of the flashover voltage on insulators

V.T. Kontargyri *, A.A. Gialketsi, G.J. Tsekouras, I.F. Gonos, I.A. Stathopulos

National Technical University of Athens, School of Electrical and Computer Engineering, Electric Power Department, 9 Iroon Politechniou St., Zografou, 157 80 Athens, Greece

Received 22 December 2005; received in revised form 31 July 2006; accepted 29 October 2006
Available online 19 December 2006

Abstract

This work attempts to apply an artificial neural network in order to estimate the critical flashover voltage on polluted insulators. The artificial neural network uses as input variables the following characteristics of the insulator: diameter, height, creepage distance, form factor and equivalent salt deposit density, and estimates the critical flashover voltage. The data used to train the network and test its performance is derived from experimental measurements and a mathematical model. Various cases have been studied and their results presented separately. Training and testing sets have been modified for each case.

Keywords: High voltage insulators; Polluted insulators; Critical flashover voltage; Artificial neural network

1. Introduction

The critical flashover voltage of a polluted insulator is a significant parameter for the reliability of power systems. Several approaches have been developed for the estimation of the flashover voltage. The exposure of insulating materials to different environmental conditions is inevitable in all power systems. Although the knowledge of overvoltages caused either by lightning or by switching overvoltages has increased, the pollution of the insulators still remains a problem capable of affecting the reliability of the electric system.

The main types of insulator pollution are marine and industrial, as well as the combination of the two. 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. The presence of electrolytic particles and moisture can form a thin film with high conductivity on the insulating surface. This layer reduces the surface resistance, leading to the flow of a leakage current. The result of this current is the ohmic heating of the surface and the creation of dry bands. Once a dry band is formatted, partial discharges can take place within it and if the voltage and the leakage current reach certain critical values, there can start the flashover phenomenon [1].

There are several techniques used for the reduction of this phenomenon and some of them include a periodical cleaning of the polluted insulators. However, if the washing and maintenance program is not reliably established, the cost increases dramatically.

Experiments concerning the critical flashover voltage U_c are time-consuming and have further obstacles, such as high cost and the need for special equipment. This has resulted in the development of several approaches for the estimation of the flashover voltage on polluted insulators. Most are based on circuit models for the calculation of the analytical mathematical relationship for either dc or ac flashover voltage on polluted insulators.

Artificial neural networks (ANNs) can be used in problems requiring function approximation, modelling, pattern recognition and classification, estimation and prediction, etc. [2]. In the field of high voltage insulators, ANNs can be used to estimate the pollution level [3,4], to predict a flashover [5,6], to analyse surface tracking on polluted insulators [7] and also to estimate the critical flashover voltage on a polluted insulator. This last case will be thoroughly examined later.

* Corresponding author. Tel.: +30 210 7722226; fax: +30 210 7723504.
E-mail address: vkont@central.ntua.gr (V.T. Kontargyri).

0378-7796/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.epsr.2006.10.017

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. Experimental measurements and data collection

The data used for training and testing the ANN was collected from both experiments [8–10] and application of a mathematical model for the calculation of the flashover voltage [11].

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 [8] and according to the IEC norm [12]. In this station, which consists mainly of two chambers: the pollution chamber and the fog chamber, 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. Before suspending the insulators in the pollution chamber they were carefully cleaned so that all traces of grease are removed, by immersing them into trisodium phosphate solution; thereafter the insulator surface was thoroughly rinsed with tap water. Then the insulators were suspended in the pollution chamber. The contaminant used was: 75 g/l kaolin clay, 675 g/l silica flour, NaCl as required, suspended in isopropyl alcohol. The contamination time needed amount to about 30 min. After the insulators were contaminated, they were allowed to dry for about 1 h. The salt deposit density on the insulator surface was used as an index for the pollution severity. During the fogging procedure the insulators were suspended in the fog chamber. The spraying system arrangement was fully in accordance to the IEC norm [12]. The time needed to reach the maximum layer conductivity was approximately 35–40 min. After the maximum value of layer conductivity was reached the test voltage was applied on the insulator in a time not exceeding 5 s, and was maintained until flashover or for 15 min if no flashover occurred. The insulator was then removed from the fog chamber and allowed to dry. It was placed for the second time in the fog chamber until the layer conductivity reaches a maximum value. The voltage was then applied again and the above procedure was carried out again. This procedure was repeated a third time, excepting both previous tests resulted to a flashover. The maximum withstand voltage, that depends upon the contamination and the geometric characteristics of the insulator, was the result of the experimental procedure [8]. Apart from this set of experimental measurements, measurements from similar experiments performed by Zhicheng and Renyu [9] and Sundararajan et al. [10] were also used.

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 Fig. 1, where V_{arc} is the arcing voltage, R_{p} the resistance of the pollution layer and U a stable voltage supply source.

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 [11]:

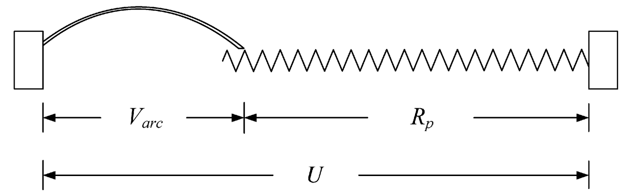

Fig. 1. Equivalent circuit for the evaluation of the flashover voltage.

$$U_{\mathrm{c}} = \frac{A}{n+1} \cdot (L + \pi \cdot n \cdot D_{\mathrm{m}} \cdot F \cdot K) \cdot(\pi \cdot A \cdot D_{\mathrm{m}} \cdot \sigma_{\mathrm{s}})^{-n/(n+1))} \quad (1)$$

where L is the creepage distance of the insulator (in cm), D_{m} the maximum diameter of the insulator disc (in cm) and F is the form factor. The form factor of an insulator is determined from the insulator dimensions. For graphical estimation, the reciprocal value of the insulator circumference ($1/p$) is plotted versus the partial creepage distance ℓ counted from the end of the insulator up to the point reckoned. The form factor is given by the area under this curve and calculated according to the formula [12]:

$$F = \int_0^L \frac{\mathrm{d}\ell}{p(\ell)}$$

The arc constants A and n have been calculated using a genetic algorithm model [13] 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:

$$\sigma_{\mathrm{s}} = (369.05 \cdot C + 0.42) \times 10^{-6} \quad (2)$$

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

$$R = 0.469 \cdot (\pi \cdot A \cdot D_{\mathrm{m}} \cdot \sigma_{\mathrm{s}})^{1/(2\cdot(n+1))} \quad (4)$$

The above mathematical model is a result of experiments in specific insulators types and specific pollutants in their surface. There are many values for the arc constants A and n in the literature [14] as a result, the above mathematical model could be applied with satisfactory accuracy in specific insulator types and pollutants.

3. ANN algorithm

ANNs can utilize the data from a learning set to model a certain problem with great accuracy. This model can then be used to estimate the output variable for given values of the input variables. ANNs try to simulate the learning process of the human brain and therefore require examples in order to be trained, rather than mathematical functions.

An ANN consists of a number of single units, called neurons, bonded with weighted connections. In a successful learning process, the weights are gradually being modified in order to give an output close to the expected. An ANN can have three types of layers: the input layer, one or more hidden layers and the output layer. When creating an ANN, one must first decide how many neurons there will be in each layer [2].

An ANN is usually trained with the error back-propagation 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 acceptable values.

In the present work, an adaptive ANN has been designed in Digital Fortran 6.0 and trained to estimate the critical flashover voltage when given some of the insulator's characteristics. The geometric characteristics of the insulator that have been used as input variables are: the diameter D_m (in cm), the height H (in cm), the creepage distance L (in cm), the form factor F and the equivalent salt deposit density C (in mg/cm^2), while the output variable was the critical flashover voltage U_c (in kV).

The ANN model determines the critical flashover voltage U_c, which has been calculated either by the experimental tests, or by the aforementioned mathematical type. The basic notion is the optimization of the model regarding the number of neurons, the initial values and time parameters of the momentum term and training rate (Sections 3.2–3.4).

Fig. 2 presents the outline of the procedure for the construction of the adaptive artificial neural network.

To summarize, 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.
(c) For each ANN parameter the adaptive back-propagation (a-BP) algorithm is separately executed for the respective range of values in order to identify the regions with satisfactory results.
(d) Then the a-BP algorithm is repeatedly executed, while all parameters are simultaneously adjusted into their respective regions, in order to select the combination that produces the minimum forecast error for the given evaluation set.
(e) Finally, the flashover voltage is estimated for the under study experiments.

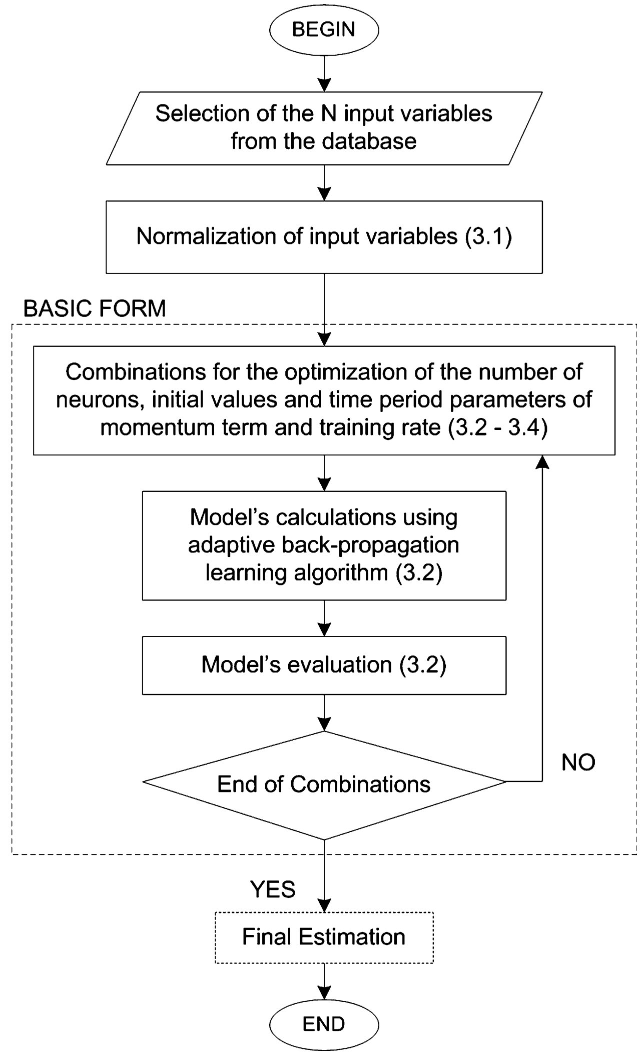

Fig. 2. The flow chart of the procedure to obtain the adaptive artificial neural network.

3.1. Selection and normalization of input variables

The model uses as input variables the geometric characteristics of the insulator and the pollution and it gives as output the critical flashover voltage. The selected set of variables, in vector form, is given by:

$$\vec{x}_i = \begin{pmatrix} x_{i1} & x_{i2} & \dots & x_{in} \end{pmatrix}^{\mathrm{T}} = (x_{ij})^{\mathrm{T}}, \quad j = 1, \dots, n_v \qquad (5)$$

where x_{ij} is the value of the jth selected variable for experiment or case study I and n_v is the total number of input variables. There are m_1 vectors for training the model, m_2 for optimizing-evaluating its parameters and m_3 for estimating the flashover voltage in the experiments under study. The m_2 vectors can be a subset of the model's training set.

In order to avoid saturation phenomena during the training process of the ANN model [2], the input and output variable values are normalized. Through preliminary algorithm executions, normalization is chosen by the maximum and minimum values of each variable, as shown in the following type:

$$x^i_{in,j} = c + (x^i_j - r_{\min,j}) \cdot \frac{b - c}{r_{\max,j} - r_{\min,j}} \qquad (6)$$

where $r_{\max,j}$ and $r_{\min,j}$ are the upper and the lower values of variable x_j for the training set and b, c the respective values of the normalized variable.

3.2. Artificial neural network

Once the connection weights are adjusted by the adaptive back-propagation learning algorithm (a-BP), the ANN can estimate flashover voltage for experiments. The basic steps of the back-propagation algorithm have been described in several textbooks [2]. According to Kolmogorov's theorem [2], 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.

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. In our case, the error function is the root mean square error $\mathrm{RMSE}_{\mathrm{va}}$ for the evaluation set according to:

$$\mathrm{RMSE}_{\mathrm{va}} = \sqrt{\frac{1}{m_2 \cdot q_{\mathrm{out}}} \sum_{i=1}^{m_2} \sum_{k=1}^{q_{\mathrm{out}}} e_k^2(i)} \qquad (7)$$

where q_{out} is the number of neurons of the output layer and $e_k(i)$ is the error of the kth output neuron for the ith pattern of the evaluation set.

If one of the three criteria is true, the main core of the back-propagation algorithm comes to an end. Otherwise, the number of epochs is increased by one, the adaption rules are applied (Section 3.3) and the feed forward and reverse pass calculations are repeated.

- *Validation criteria*: For the evaluation set, the root mean square error ($\mathrm{RMSE}_{\mathrm{va}}$), the mean absolute square error ($\mathrm{MAPE}_{\mathrm{va}}$) and the correlation ($R^2_{\mathrm{va}}$) can be calculated. $\mathrm{MAPE}_{\mathrm{va}}$ is given by:

$$\mathrm{MAPE}_{\mathrm{va}}(k) = 100\% \cdot \frac{\sum_{i=1}^{m_2} |t_k(i) - o_k(i)/t_k(i)|}{m_2} \qquad (8)$$

where $t_k(i)$ and $o_k(i)$ are the real and the estimated value of the kth output neuron for the ith pattern of the evaluation set, respectively.

Correspondingly, the correlation (R^2) is defined as:

$$R^2_{\mathrm{va}}(k) = \frac{\sum_{i=1}^{m_2} (o_k(i) - \bar{t}_k)^2}{\sum_{i=1}^{m_2} (t_k(i) - \bar{t}_k)^2} \qquad (9)$$

where $\bar{t}_k$ is the respective mean value of $t_k(i)$. For the final estimated flashover voltage, Eqs. (7)–(9) can also be applied to the respective patterns.

- *Activation function*: A number of activation functions, also known as transfer functions, can be applied. In the case of hyperbolic tangent, the unknown parameters are h_1 and h_2, as:

$$f(x) = \tanh(h_1 \cdot x + h_2) \qquad (10)$$

Other candidate functions, like unipolar sigmoid and arctangent, also include parameters that have to be similarly determined. Through preliminary algorithm executions, the hyperbolic tangent gives better results in this kind of problem.

3.3. Adaption rules

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

$$\eta(\mathrm{ep}) = \begin{cases} \eta(\mathrm{ep}-1), & \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}) > \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}-1) \\ \eta(\mathrm{ep}-1) \cdot \exp\left(\dfrac{-1}{T_\eta}\right), & \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}) \le \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}-1) \end{cases} \qquad (11)$$

$$a(\mathrm{ep}) = \begin{cases} a(\mathrm{ep}-1), & \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}) \le \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}-1) \\ a(\mathrm{ep}-1) \cdot \exp\left(\dfrac{-1}{T_a}\right), & \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}) > \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}-1) \end{cases} \qquad (12)$$

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 $\mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep})$ 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 $\mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}-1)$ is larger than $\mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep})$, 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 $\mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}) > \mathrm{RMSE}_{\mathrm{tr}}(\mathrm{ep}-1)$, 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.

It should be noted that increasing the momentum term or the learning rate, as proposed in Refs. [2,15], leads to unstable solutions during the convergence process due to the restricted population of our training set.

3.4. Combination analysis

Until this point, there are seven parameters to be selected: the number of neurons, the time parameters, the initial values of the training rate and the momentum term and the parameters of the activation function. The activation functions do not play a substantial role, especially when one of the sigmoid functions is chosen. In this case, the hyperbolic tangent is used with $h_1 = 0.7$ and $h_2 = 0$.

If each mth parameter corresponds to s_i values, then the possible combinations are $s_t = s_1, s_2, \ldots, s_5$, which is a very large number despite the limited training set.

The ANN parameters need to be specified. In order to reduce these combinations, two steps are taken. In the first step, the basic algorithm is executed separately for each parameter's range of values and 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. This combination is used for the flashover voltage in the experiments of interest.

4. Case studies

Several cases have been examined using the ANN described above. These cases will be described in this section.

As mentioned, the data used are derived from two sources: from experiments and from a mathematical model. Different case studies have been performed, each time using different populations (i.e. different sets for the training and testing of the ANN).

4.1. Case 1

In case 1, the data from the mathematical model and a set of the experimental data, containing the maximum and minimum values were used to train the network, while the rest of the experimental data were used to test its performance. The training set consists of 148 patterns/vectors (of which the 140 vectors are derived from the model and 8 vectors are real values) and the network was tested using 20 patterns (experimental data).

With this first test, 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 fewer real values.

4.2. Case 2

In the second case studied, the only data used are the experimental data (28 in total). Twenty patterns were used to train the network and the other 8 to test it.

4.3. Case 3

In the third case, the database also consists of the 28 real patterns. From these patterns, 24 are used to train and 4 were to test the ANN.

In both cases 2 and 3 only real values (experimental data) are used. As a result, the training sets are small and consist of only few vectors, something that was expected to cause problem to the generalization of the ANN. However, judging from the results, the ANN has a good performance when the maximum and minimum values of all variables are included in the training set.

In all three cases there is a small number of vectors/patterns available. That is why the optimization-evaluation set is the same with the training set.

The results given by the ANN are presented in the following section.

5. Results

5.1. Case 1

Table 1 shows the first test that was made in order to decide the number of neurons (N) of the hidden layer. The criterion for that was the minimization of the $RMSE_{tr}$ for the training set. As shown in the second column of Table 1, the minimum $RMSE_{tr}$ appears for 6 neurons ($RMSE_{tr} = 0.112$) and for 17 neurons. However, the final choice is a network with 6 neurons in its hidden layer, because in the case of 17 neurons the ANN becomes unstable. It must be mentioned that the number of neurons altered from 2 to 20 and Table 1 presents the $RMSE_{tr}$ for 2–20 neurons.

The next step was to define the parameters of the momentum (constant term and time parameter) that lead to minimum $RMSE_{tr}$. Fig. 3 shows a 3D plot of the $RMSE_{tr}$ as a function of 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 $T_a = 4000$.

Table 1
Test concerning the architecture of the network

Number of neuron	$RMSE_{tr}$		
	Case 1	Case 2	Case 3
2	0.4883	0.3580	0.3632
3	0.1420	0.3331	0.3589
4	0.1390	0.3421	0.3147
5	0.1196	0.3468	0.3286
6	0.1120	0.3266	0.3263
7	0.1144	0.3337	0.3292
8	0.1152	0.3540	0.3218
9	0.1142	0.3279	0.3475
10	0.1131	0.3343	0.3187
11	0.1129	0.3641	0.3338
12	0.1124	0.3423	0.3384
13	0.1131	0.3954	0.3596
14	0.1126	0.3834	0.3402
15	0.1122	0.3458	0.3474
16	0.1169	0.3920	0.3432
17	0.1116	0.3619	0.3529
18	0.1128	0.3995	0.3498
19	0.1145	0.3743	0.4142
20	0.1123	0.3609	0.4063

The minimum $RMSE_{tr}$ appears for six neurons.

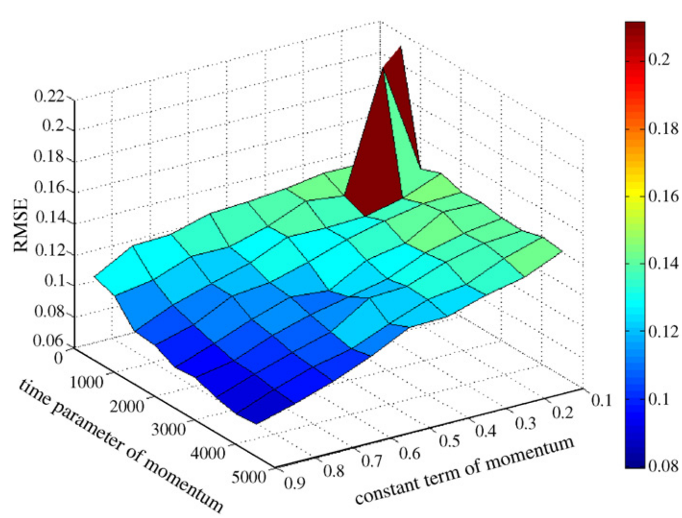

Fig. 3. Test made to find the constant term and time parameter of the momentum that give the minimum $RMSE_{tr}$ for case 1. The hidden layer has six neurons ($N = 6$).

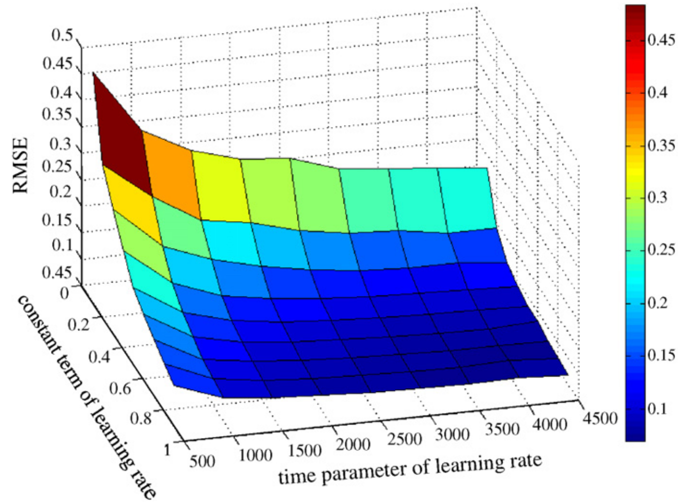

Fig. 4. Test made to find the constant term and time parameter of the learning rate that give the minimum $RMSE_{tr}$ for case 1. $N=6$, $a_0=0.8$, $T_a=4000$.

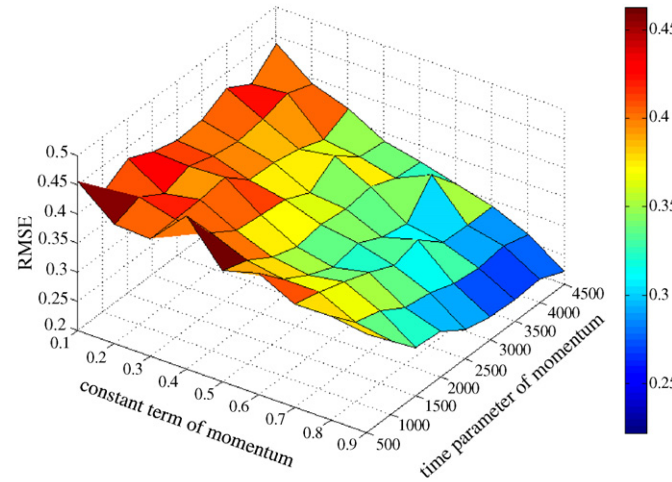

Fig. 6. Test made to find the constant term and time parameter of the momentum that give the minimum $RMSE_{tr}$ for case 2. The hidden layer has six neurons ($N=6$).

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. 4 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 0.070 kV. 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.

As mentioned before, the training accuracy was measured by $RMSE_{tr}$, however the mean absolute error (MAPE) was also calculated for both the training and test data. Table 2 shows the values for the $RMSE_{tr}$, $MAPE_{tr}$ and $MAPE_{test}$ when the optimum values of all the parameters (N, α_0, T_a, η_0 and T_η) have been definitized.

Table 2
Final results for case 1. $N=6$, $a_0=0.8$, $T_a=4000$, $\eta_0=0.9$, $T_n=4500$

Case study	$RMSE_{tr}$ (kV)	$MAPE_{tr}$ (%)	$MAPE_{test}$ (%)
1	0.070	1.300	4.040

It must be mentioned that if, for the estimation of the flashover voltage only the mathematical model was used, then the MAPE would be 4.574%, instead of 4.040%. Therefore, it is obvious that the two methods for the estimation of the flashover voltage give comparable results.

Another way to test the accuracy of the ANN is to see the correlation between the real values of the U_c and the values that come up as output of the ANN, that is the estimated values. The maximum value that the correlation can take is 1. That means that the closer the correlation is to 1, the better the network operates. Fig. 5 presents the correlation between the estimated and the real values for the training set and for the set used to test the network. In both cases the values of correlation are very close to the ideal value 1, even when the data are presented to the network for the testing set.

5.2. Case 2

In case 2, the same tests were repeated in order to find which network architecture and which values in the parameters of the ANN give best results. The third column of Table 1 shows the test that was made in order to decide the number of neurons (N) of the hidden layer. The minimum $RMSE_{tr}$ appears for six neurons in the hidden layer, as in case 1.

Since the architecture of the network has been defined, the following step is the definition of the momentum parameters. Fig. 6

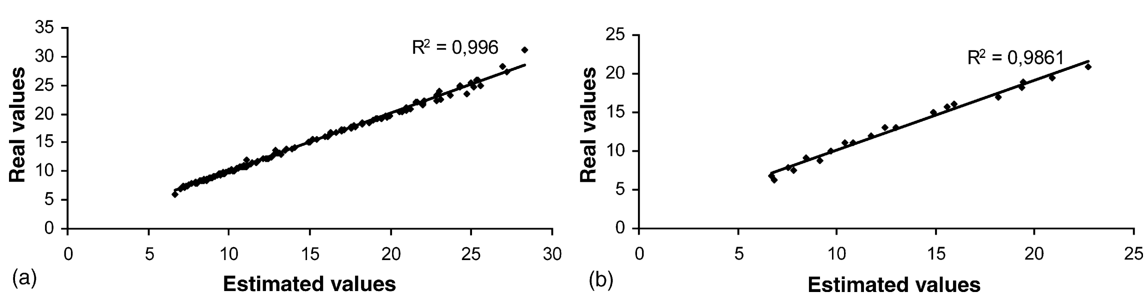

Fig. 5. Correlation between estimated and real values of U_c for case 1 (a) for the training set and (b) for the testing set.

Table 3
Final results for case 2. $N = 6$, $a_0 = 0.9$, $T_a = 4500$, $\eta_0 = 0.8$, $T_n = 4500$

Case study	$RMSE_{tr}$ (kV)	$MAPE_{tr}$ (%)	$MAPE_{test}$ (%)
2	0.178	4.154	4.255

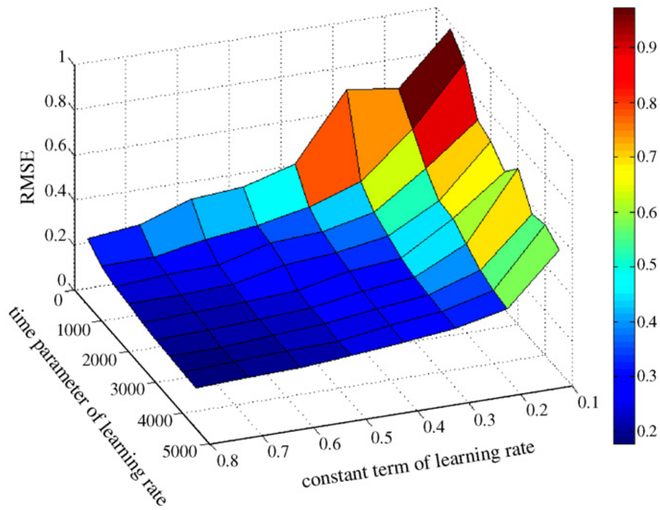

Fig. 7. Test made to find the constant term and time parameter of the learning rate that give the minimum $RMSE_{tr}$ for case 2. $N = 6$, $a_0 = 0.9$, $T_a = 4500$.

shows a 3D plot of the $RMSE_{tr}$ as a function of the momentum. The constant term of momentum (a_0) changes again from 0.1 to 0.9 and the time parameter (T_a) from 500 to 5000. The minimum $RMSE_{tr}$ appears for $a_0 = 0.9$ and $T_a = 4500$.

The determination of the learning rate is the final step in case 2. In this last test 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 $RMSE_{tr}$ as a function of the learning rate. The minimum $RMSE_{tr}$ appears for $\eta_0 = 0.8$ and $T_n = 4500$. The value of the $RMSE_{tr}$ is now 0.178 kV (Table 3).

5.3. Case 3

The same sequence of steps is followed in case 3. First, the number of neurons of the hidden layer is determined, then the

Table 4
Final results for case 3. $N = 4$, $a_0 = 0.9$, $T_a = 4500$, $\eta_0 = 0.7$, $T_\eta = 4500$

Case study	$RMSE_{tr}$ (kV)	$MAPE_{tr}$ (%)	$MAPE_{test}$ (%)
3	0.165	4.004	2.152

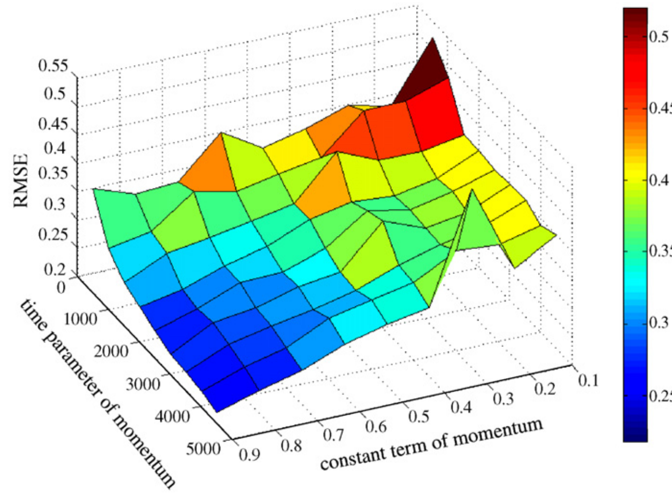

Fig. 8. Test made to find the constant term and time parameter of the momentum that give the minimum $RMSE_{tr}$ for case 3. The hidden layer has four neurons ($N = 4$).

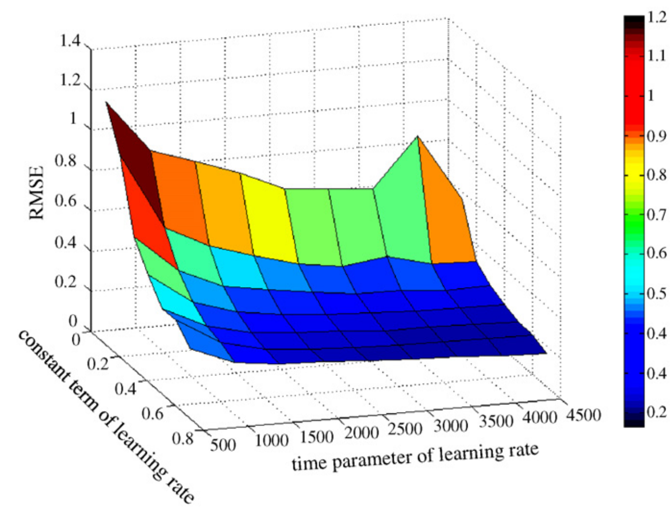

Fig. 9. Test made to find the constant term and time parameter of the learning rate that give the minimum $RMSE_{tr}$ for case 3. $N = 4$, $a_0 = 0.9$, $T_a = 4500$.

parameters of the momentum that give the minimum error and finally the parameters of the learning rate. In the fourth column of Table 1 and in Figs. 8 and 9, the results of those three tests are presented. This time the ANN has four neurons in the hidden layer, the constant term of momentum is $a_0 = 0.9$, the time parameter of momentum is $T_a = 4500$, the constant term of learning rate is $\eta_0 = 0.7$ and the time parameter of learning rate is $T_\eta = 4500$. The value of the $RMSE_{tr}$ is now 0.165 kV (Table 4).

Table 5
Values that were used in the mathematical model for the calculation of the flashover voltage

D_m (cm)	26.8	26.8	25.4	25.4	29.2	27.9	32.1	28.0	25.4	20.0
H (cm)	15.9	15.9	16.5	14.6	15.9	15.6	17.8	17.0	14.5	16.5
L (cm)	33.0	40.6	43.2	31.8	47.0	36.8	54.6	37.0	30.5	40.0
F	0.79	0.86	0.90	0.72	0.92	0.76	0.96	0.80	0.74	1.29

6. Conclusions

In this paper, an ANN has been successfully applied for the estimation of the flashover voltage on polluted insulators. The network was trained to estimate the critical flashover voltage when given some of the insulator's characteristics. This was made even more efficient with an adaptive algorithm, in which the parameters of momentum and learning rate changed during the learning procedure, in order to optimize the training process. The fundamental advantage of this proposed ANN is the ability to find the optimized choice of parameters such as the learning rate, the momentum term and the number of neurons. Meanwhile, it leads to satisfactory results, even when the training set is small, as long as it contains the minimum and maximum presented values of the variables. The results clearly show that the ANN can estimate the flashover voltage efficiently and effectively. The training and test data were obtained from experimental studies and from application of a mathematical model for the estimation of the flashover voltage on polluted insulators. In case 1, in which the data were derived from both experimental measurements and the mathematical model, the results of the ANN were better than those of the mathematical model alone, as shown by the $MAPE_{test}$ of the test data. When only the experimental data were used (case 2), the $MAPE_{test}$ is comparable with the previous $MAPE_{test}$ (case 1), although its small rise can be explained by the fact that the training set had fewer vectors than in case 1. In case 3 the training data were increased and the $MAPE_{test}$ was decreased. In conclusion, it could be said that if the experimental data set was larger, then the results of the ANN would be even better. Also the ANN could be applied in various types of insulators with higher accuracy than the mathematical model.

Table 6
Experimental values

D_m (cm)	H (cm)	L (cm)	F	C (mg/cm^2)	U_c (kV)
25.4	14.6	27.9	0.68	0.13	12.0
25.4	14.6	27.9	0.68	0.16	11.1
25.4	14.6	27.9	0.68	0.23	8.7
25.4	14.6	27.9	0.68	0.28	9.1
25.4	14.6	27.9	0.68	0.34	7.5
25.4	14.6	27.9	0.68	0.37	7.8
25.4	14.6	27.9	0.68	0.49	6.2
25.4	14.6	27.9	0.68	0.52	6.8
25.4	14.6	27.9	0.68	0.55	6.1
25.4	14.6	30.5	0.70	0.02	22.0
25.4	14.6	30.5	0.70	0.05	16.0
25.4	14.6	30.5	0.70	0.10	13.0
25.4	14.6	30.5	0.70	0.16	11.0
25.4	14.6	30.5	0.70	0.22	10.0
25.4	14.6	30.5	0.70	0.30	8.5
25.4	14.6	43.2	0.92	0.02	26.0
25.4	14.6	43.2	0.92	0.05	19.0
25.4	14.6	43.2	0.92	0.10	15.0
25.4	14.6	43.2	0.92	0.16	13.0
25.4	14.6	43.2	0.92	0.22	12.0
25.4	14.6	43.2	0.92	0.30	10.5
22.9	16.6	43.2	1.38	0.02	23.5
22.9	16.6	43.2	1.38	0.03	20.9
22.9	16.6	43.2	1.38	0.04	19.4
22.9	16.6	43.2	1.38	0.05	18.3
22.9	16.6	43.2	1.38	0.06	16.9
22.9	16.6	43.2	1.38	0.10	15.8
22.9	16.6	43.2	1.38	0.20	13.6

Acknowledgements

This project is co-funded by the European Social Fund (75%) and National Resources (25%) Operational Program for Educational and Vocational Training II (EPEAEK II) and particularly by the Program HERAKLEITOS.

Appendix A

In this section, the theoretical and experimental data that were used in this work is presented.

Using the data given in Table 5 and the following values for the equivalent salt deposit density C (in mg/cm^2): $\{0.02, 0.03, 0.04, 0.05, 0.06, 0.13, 0.16, 0.23, 0.28, 0.34, 0.37, 0.49, 0.52, 0.55\}$ and applying Eq. (1), the flashover voltage can be calculated. The experimental data are also given in Table 6.

References

[1] IEC 815, Guide for the selection of insulators in respect of polluted conditions, 1986.

[2] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice Hall, 1994.

[3] A.S. Ahmad, P.S. Ghosh, S.A.K. Aljunid, H. Ahmad, I. Said, H. Hussain, Artificial Neural Network for Contamination Severity Assessment of High Voltage Insulators Under Various Meteorological Conditions, AUPEC, Perth, 2001.

[4] A.S. Ahmad, P.S. Ghosh, S. Shahnawaz Ahmed, S.A.K. Aljunid, Assessment of ESDD on high-voltage insulators using artificial neural network, Elsevier Electr. Power Syst. Res. 72 (2) (2004) 131–136.

[5] P.S. Ghosh, S. Chakravorti, N. Chatterjee, Estimation of time-to-flashover characteristics of contaminated electrolytic surfaces using a neural network, IEEE Trans. Dielectr. Electrical Insulation 2 (6) (1995) 1064–1074.

[6] P. Cline, W. Lannes, G. Richards, Use of pollution monitors with a neural network to predict insulator flashover, Elsevier Electr. Power Syst. Res. 42 (1) (1997) 27–33.

[7] M. Ugur, D.W. Auckland, B.R. Varlow, Z. Emin, Neural networks to analyze surface tracking on solid insulators, IEEE Trans. Dielectr. Electrical Insulation 4 (6) (1997) 763–766.

[8] K. Ikonomou, G. Katsibokis, G. Panos, I.A. Stathopoulos, Cool fog tests on artificially polluted insulators, in: Fifth International Symposium on High Voltage Engineering, Braunschweig, vol. II, 1987, paper 52.13.

[9] G. Zhicheng, Z. Renyu, Calculation of DC and AC flashover voltage of polluted insulators, IEEE Trans. Electrical Insulation 25 (4) (1990) 723–729.

[10] R. Sundararajan, N.R. Sadhureddy, R.S. Gorur, Computer-aided design of porcelain insulators under polluted conditions, IEEE Trans. Dielectr. Electrical Insulation 2 (1) (1995) 21–127.

[11] F.V. Topalis, I.F. Gonos, I.A. Stathopoulos, Dielectric behaviour of polluted porcelain insulators', IEE Proc. Gen. Transm. Distribution 148 (4) (2001) 269–274.

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

[13] I.F. Gonos, F.V. Topalis, I.A. Stathopoulos, Genetic algorithm approach to the modelling of polluted insulators, IEE Proc. Gen. Transm. Distribution 149 (3) (2002) 373–376.

[14] D.C. Chaurasia, Scintillation modelling for insulator strings under polluted conditions, in: 11th International Symposium on High Voltage Engineering, London, UK, 1999.

[15] S. Theodoridis, K. Koutroumbas, Pattern Recognition, first ed., Academic Press, USA, 1999.

Vassiliki T. Kontargyri received her diploma in electrical engineering from National Technical University of Athens in 2002. Since 2002 she is a Ph.D. student and a research associate, at the High Voltage Laboratory of National Technical University of Athens. Her research interests concern high voltage insulators, high voltages, grounding systems and artificial neural networks.

Angeliki A. Gialketsi received her diploma in electrical engineering from National Technical University of Athens in 2005. Since 2005 she is a postgraduate student at National Technical University of Athens. Her research interests concern high voltages, insulators and artificial neural networks.

George J. Tsekouras received the diploma in electrical and computer engineering from the NTUA, Greece, in 1999. He is a Ph.D. candidate at NTUA. His research interests include power system analysis, load and energy Forecasting methods and database design. He is a member of IEEE and member of the Technical Chamber of Greece.

Ioannis F. Gonos received his diploma in electrical engineering and his Ph.D. from the National Technical University of Athens in 1993 and 2002, respectively. He was a teaching assistant at the Greek Naval Academy and the Technological Education Institute of Athens (1996–2001). He is working at the High Voltage Laboratory of NTUA (since 2001). His research interests concern grounding systems, insulators, high voltages, measurements and genetic algorithms. He is a member of IEEE, IEE and CIGRE.

Ioannis A. Stathopulos studied in the faculty of electrical and mechanical engineering of the National Technical University of Athens (1969–1974). He carried out his doctor thesis at the Technical University of Munich (1974–1978). He become teaching assistant at the Technical University of Munich (1974–1978), production engineer in the company “Vianox–Franke” (1979–1980), teaching assistant at the National Technical University of Athens (1979–1983) and thereafter lecturer (1983–1987), assistant professor (1987–1991), associate professor (1991–1995) and professor (since 1995) in the High Voltage Laboratory of the NTUA. He is the author of 8 books and more than 100 papers in scientific journals and conferences proceedings. He is lead assessor of the Hellenic Accreditation Council.