Research Journal of Recent Sciences ISSN 2277-2502
Vol. **8(4),** 20-30, October **(2019)** *Res. J. Recent Sci.*

Estimation of soil electrical resistivity using MLP, RBF, ANFIS and SVM approaches

Sangué Oraléou Djandja[1], AdekunléAkim Salami[1,2*], Kpomonè Komla Apaloo-Bara[1,2] and Koffi-Sa Bedja[1,2]
[1]Electrical Engineering Departement, Ecole Nationale Supérieure d'Ingénieurs (ENSI), University of Lome, Togo
[2]LAboratoirede Recherche en Sciences de l'Ingénieur (LARSI), Ecole Nationale Supérieure d'Ingénieurs (ENSI), University of Lome, Togo
akim_salami@yahoo.fr

Available online at: www.isca.in, www.isca.me
Received 8th February 2019, revised 22nd August 2019, accepted 15th September 2019

Abstract

The knowledge of soil electrical resistivity proves essential for a better earthing in order to ensure the protection of telecommunications and electrical energy networks. This study aims to estimate the value of the electrical resistivity of a site's soil from soil humidity and ambient temperature. The data used were measured at sites in the city of Lomé and its surroundings. We developed models using Artificial Neural Network (precisely Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF)), Adaptive Neuro-Fuzzy Inference System (ANFIS)and Support Vector Machine (SVM). The MAPE (Mean Absolute Percentage Error) errors obtained are 0.0011761% for the MLP model, 0.0719309% for the RBF model, 0.00105% for the ANFIS model and 2.89466% for the SVM model. We can say that the results are satisfactory for all models but the ANFIS model is better, given these performances compared to other models. The latter is then retained for the prediction of soil electrical resistivity.

Keywords: Resistivity, MLP, RBF, ANFIS, SVM, MAPE.

Introduction

Nowadays, electrical energy appears as the protagonist factor for the socio-economic development of a nation. The use of this energy exposes people and property to risks. The leaders then give themselves as a duty to protect power grids in order to ensure use without fear of electricity by the population. One of the quality factors of this protection is the earthing system, which is very important for the good exploitation of telecommunication and electrical energy networks. It is an important element of the electrical installations and the lightning and default current supply systems' protection[1]. Thus, a properly designed earthing system is able to dissipate large currents to the earth safely, whatever the default. A ground resistance kept at low levels throughout the year, makes effective an earthing system[1,2]. The type of soil must be seriously considered in the design of such a system because of its important electrical resistivity or its particularly corrosive environment. The soil electrical resistivity therefore appears as an important parameter of the ground for the electrical grounding. However, this varies according to several variables in a random manner, namely the chemical composition of the soil, its particle size, its water content, the nature of the environment and the soil temperature ... it follows that the earthing system is not characterized by a single value of the soil electrical resistivity; hence the need to monitor the values of this resistivity.

The purpose of this study is to provide a model that will be used to estimate soil electrical resistivity based on soil humidity and ambient temperature. We have chosen to explorate MLP, RBF, ANFIS, and SVM approaches, in order to retain the best approach of these four (04) to be a useful tool for predicting soil electrical resistivity values.

Many studies have shown the strong dependence between the measurement of electrical resistivity of the soil and various intrinsic physical and chemical variables of soils. Thus, an Artificial Neural Network (ANN) model was developed to estimate soil resistance, using measurements of resistivity and precipitation data accumulated for previous days[3]. J.P. Lee et al. published the results of a study titled "Earth parameter and equivalent resistivity estimation using ANN". The chosen model of their study provided effectiveness with studies cases[4]. Judging the key instabilities affecting the variation of soil resistance (soil composition, water content, temperature, mass electrodes and electrode spacing), a Generalized Neural Network Regression (GRNN) was developed to predict the Athens soil resistance[5]. Another study aims to provide an ANN model for estimating the variation in soil resistance throughout the year, using measurements of soil resistivity, temperature and period of time[6]. The work of John Tarilanyo Afa and C.M. Anaele showed that seasonal variation and soil type affect the performance of grounding systems[7]. We therefore remember that the problem around electrical resistivity is topical.

The approaches chosen for our study have been modeled in various fields and have given good results. Marcin Grabarczyk and Piotr Furmanski have developed an ANN model with three hidden layers for estimating the thermal conductivity of granular media[8]. Two neuro-fuzzy models were developed for the estimation of the MPPT based on knowledge of the short-circuit current and the open circuit voltage and the results were satisfactory[9]. Zaki Abda and al. modeled extreme flow rates by artificial neural networks and fuzzy inference systems; The results obtained in the Algerian coastal basins are very encouraging and better than those obtained by traditional statistical models[10]. Another study that predicted solar radiation by day of the year, was made by Kasra Mohammadi and al, using ANFIS; the results were very satisfactory with a correlation coefficient of 0.9908[11].

These literature reviews show that several factors influence soil electrical resistivity and that the approaches cited are approximators that have shown their performance in many areas and therefore these approaches will be applied to model the soil electrical resistivity, based on measurements. The data used for the implementation of these models come from measurements made on sites in the city of Lomé and in the surrounding area.

The main interest is that this study provides a parsimonious model (with only two inputs) and that in need of a prediction, only one of its inputs will be measured (soil humidity); the other (temperature), is often given by the Meteorological Directorate.

Approaches used: A brief presentation of the approaches chosen for our work should be made to explain their operating principle.

Artificial neural network (ANN): it is defined as a reasoning model based on the human brain. A neural network is in the form of a mathematical model composed of neurons connected to each other by weights and operating in parallel[12]. They belong to the category of "black box" models. Mc Culloch and Pitts[13] are the first to show that simple formal neural networks can realize complex logical, arithmetic and symbolic functions.

In principle, artificial neural networks can be applied to perform many tasks, such as pattern recognition or classification problems[14]. In our current investigation, we used their ability to approximate and interpolate functions.

A neural network output also depends, among other parameters, on the learning procedure. The learning step is based on the retro propagation of the error. Its output expression is given by relation (1):

$$O_k = \sum_{j=1}^{q} W_{kj} b_j(x) - \theta_k \quad (1)$$

Where: $1 < k < m$; m is the number of nodes, O_k is the output of the k^{th} node of the output layer, W_{kj} is the connection between the j^{th} neuron of hidden layer and k^{th} neuron of output layer, $b_j(x)$ is the output of the j^{th} neuron of the hidden layer, Θ_k is the bias of the k^{th} neuron of output layer.

Architecture showing the arrangement of hidden and output layers of a two (02) layers ANN is shown in Figure-1.

The difference between the MLP model and the RBF model is at the hidden layer where the MLP has a sigmoidal activation function while RBF has a Gaussian activation function.

For MLP model, the output is given by relation (2):

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i h_i \quad (2)$$

Where: y is the predicted value with the neural network, n is the number of hidden layers, β_0 is the bias, β_i is the weighted coefficients, h_i is the result of the non-linear transformation of the i^{th} hidden unit.

For the RBF model, during the learning process, each neuron in the hidden layer performs a nonlinear transformation. The output of a RBF neuron with Gaussian non-linearity is expressed by relation (3):

$$b_j(x) = \exp\left[-\frac{\sum_{i=1}^{n} (x_i - \mu_j)^2}{2\sigma_j^2}\right] \quad (3)$$

Where: μ_j and σ_j are respectively the center and the width (standard deviation) of the Gaussian function of the j^{th} neuron of the hidden layer, X_i are input variables of the neuron, q is the number of neurons in the hidden layer ($1<j<q$).

Adaptive Neuro-Fuzzy Inference System (ANFIS): It is also an ANN but is based on the fuzzy inference system of Takagi-Sugeno. This system has been used at first by Jang et al.[15] who had used an MLP network.

Let us remember that ANFIS is an association of neural networks and fuzzy logic; this in order to achieve a good reasoning in quality and quantity[16].

There are two types of fuzzy systems, so there are also two types of neuroflous networks that are Tagaki-Sugeno neuroflous networks and Mamdani neuroflous networks. In this work, we used Tagaki-Sugeno's neuroflous networks with reference to their universal approximation properties and the fact that they no longer require a defuzzification module as in case of Mamdani fuzzy system.

It is a network of fuzzy systems of the Sugeno type endowed with the learning capabilities of neurons. For simplicity, we consider that the system has two inputs a and b, one output f and each input is represented by two fuzzy sets. When we consider a first order Sugeno model, we have the following rules:

Rule 1: *If* a *is* W1 *and* b *is* Z1 *Then* $f1 = w1 \times a + z1 \times b + c1$ (4)

Rule 2: *If* a *is* W2 *and* b *is* Z2 *Then* $f2 = w2 \times a + z2 \times b + c2$ (5)

Where Wi and Zi are fuzzy sets, wi, zi and ci are the consequent parameters that are determined.

ANFIS architecture comes down to five layers as shown in Figure-2.

Learning consists of correcting the parameters (premises and consequents) of the network to generalize a transfer function between the inputs and the output of the network.

This consists of a set of known "input / output" pairs (record of data). The deployment of learning algorithms on this database allows to build a function of approximation of the output (desired output) from new input vectors. ANFIS uses back propagation of error learning to determine input of parameter membership functions; and the least mean square method for determining outcome parameters.

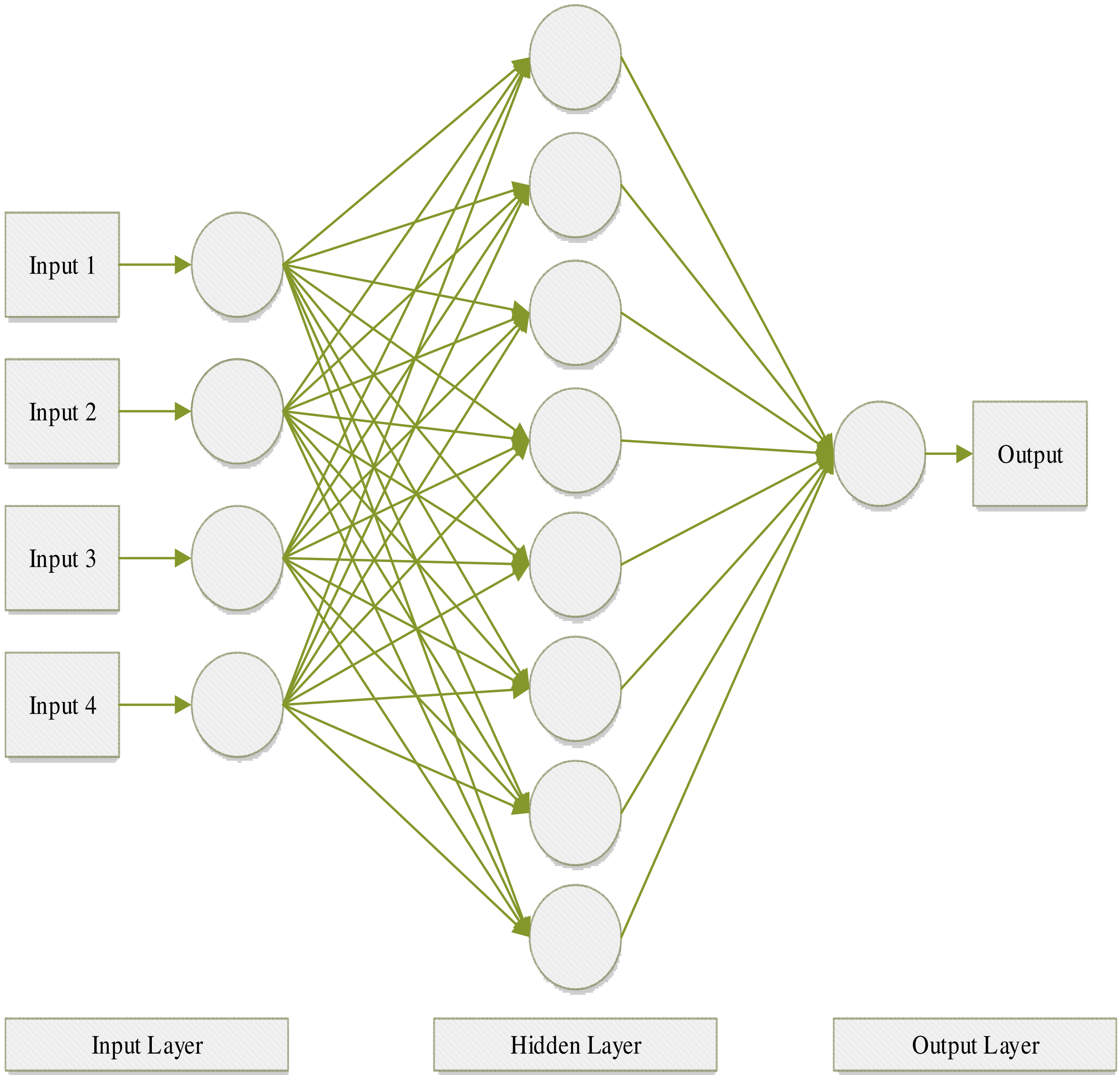

Figure-1: Structure of two layers ANN[10].

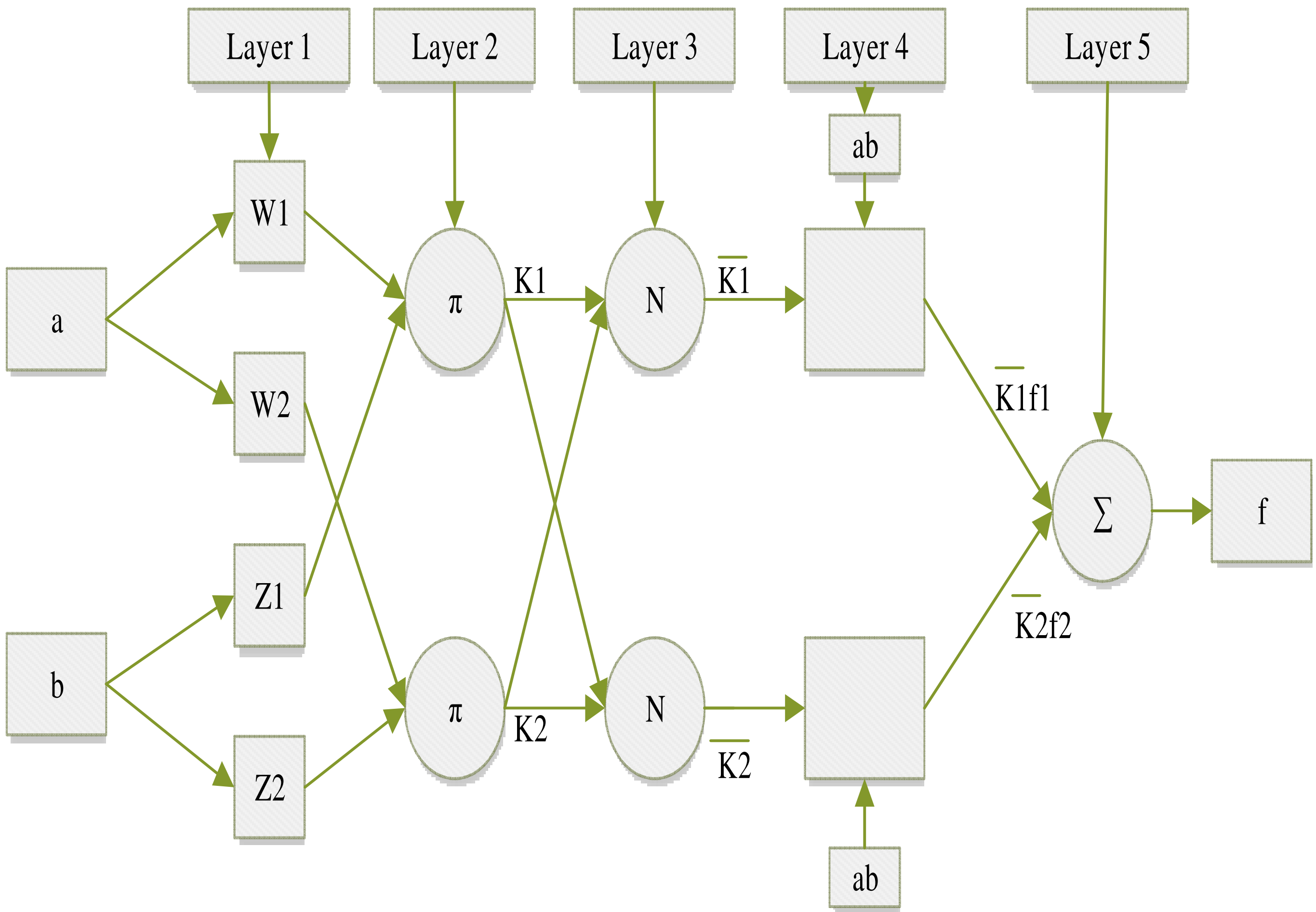

Figure-2: General architecture of ANFIS[10].

Support Vectors Machines (SVM): They belong to the group of statistical learning algorithms that were introduced in 1995 by Vladimir VAPNIK[17]. Developed at first glance for the resolution of classification problems, they have found applications in other areas, including that of regression. SVM involve several mathematical notions, including the theory of generalization, optimization theory and learning methods based on the "kernel" functions.

This technique tries to linearly separate the positive examples from the negative ones, while ensuring that the margin between the nearest positive and negative is maximal. Each example must be represented by a dimension vector "n".

An SVM solving algorithm identifies among the learning examples which are the support vectors and constructs the boundary (or decision function) with a linear combination of this selection. Solving this problem is equivalent to solving a quadratic program under box constraints.

Our work is based on Platt J.'s algorithm, Sequential Minimal Optimization (SMO). John Platt[18] proposed a new algorithm for SVM training that he called Sequential Minimal Optimization (SMO). This algorithm decomposes the problem into subproblems but chooses to solve the smallest possible subproblem at each step of optimizing the objective function.

The main idea of this algorithm is to decompose the problem to the extreme by optimizing only two points at each iteration[19]. The advantage of this is that optimizing a bivariate equation is a problem that has an analytical solution.

It is this algorithm that has been the subject of our estimation of the soil electrical resistivity using the Gaussian kernel, widely used in practice, which is evaluated according to relation (6).

$$K(x,\mathrm{z}) = \exp\left(-\frac{\|x-z\|^2}{2\sigma^2}\right) \quad (6)$$

where σ is a positive real that represents the bandwidth of the kernel.

Performance criteria: Four indicators are taken into account in the evaluation of the performances of the various models: Mean

Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), Relative Root Mean Square Error (RRMSE) and the correlation coefficient (R^2). The correlation coefficient, should be close to 1, to reflect a strong correlation between the predicted and observed values.

$$MAPE = \frac{1}{N}\sum_{k=1}^{N}\left|\frac{\rho_{\mathrm{k,e}} - \rho_{\mathrm{k,m}}}{\rho_{\mathrm{k,m}}}\right| \times 100 \qquad (7)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\rho_{\mathrm{k,e}} - \rho_{\mathrm{k,m}})^2} \qquad (8)$$

$$RRMSE = \frac{\sqrt{\frac{1}{N}\sum_{k=1}^{N}(\rho_{\mathrm{k,m}} - \rho_{\mathrm{k,e}})^2}}{\frac{1}{N}\sum_{k=1}^{N}\rho_{\mathrm{k,m}}} \qquad (9)$$

$$R^2 = \frac{\sum_{k=1}^{N}(\rho_{\mathrm{k,e}} - \rho_{\mathrm{e},avg}) \times (\rho_{\mathrm{k,m}} - \rho_{\mathrm{m},avg})}{\sqrt{\left[\sum_{k=1}^{N}(\rho_{\mathrm{k,e}} - \rho_{\mathrm{e},avg})^2\right] \times \left[\sum_{k=1}^{N}(\rho_{\mathrm{k,m}} - \rho_{\mathrm{m},avg})^2\right]}} \qquad (10)$$

In Equations (7), (8), (9), and (10), N is the number of measured values; $\rho_{\mathrm{k,e}}$ are the estimated values; $\rho_{\mathrm{k,m}}$ are measured values; $\rho_{\mathrm{e},avg}$ is the estimated mean value; and $\rho_{\mathrm{m},avg}$ is the measured mean value.

Methodology

The resistivity data used are based on measurements taken at nine (09) sites in the City of Lomé, using the Wenner method. The penetration depth of the electrodes is one meter (1m). At each measured resistivity value, a value of the soil humidity and a value of the ambient temperature, measured at the same location, are associated. We present in Table-1, Mean and standard deviation (Std) of each type of data for each site. Figure-3,4,5 show the distribution of ambient temperature, soil humidity and soil resistivity for each site.

The input vectors of the models are soil humidity and ambient temperature; the output being a vector containing the predicted values of the soil electrical resistivity.

For the prediction process, the available database has been separated into two (02) subsets, to adjust the model parameters and obtain optimal performance. A set of 80% of the data is used for the learning phase of the model and the remaining 20% was used for the validation phase.

All approaches are explored by raising the performance criteria selected in order to retain the one that will provide the best performance.

For the MLP and RBF models, given the update of the synaptic weights at each run, they were run 20 times each during the learning phase and the validation phase. We then note the average performance criteria. The main implementation parameters for these models are given in Table-2.

Table-1: Mean and standard deviation for each site.

sites	Ambient Temperature	Soil Humidity	Soil Resistivity			
	Mean	Std.	Mean	Std	Mean	Std.
1	28,87	1,67	30,06	1,26	172,87	19,70
2	28,87	1,67	24,75	1,15	89,89	17,97
3	28,87	1,67	27,47	1,35	132,35	21,15
4	28,63	1,67	31,06	0,91	188,46	14,22
5	28,63	1,67	26,83	0,84	122,32	13,18
6	28,63	1,67	30,35	1,19	177,30	18,63
7	28,78	1,60	27,95	1,41	139,80	21,96
8	28,78	1,60	25,61	1,13	103,25	17,69
9	28,78	1,60	31,31	1,29	192,42	20,10

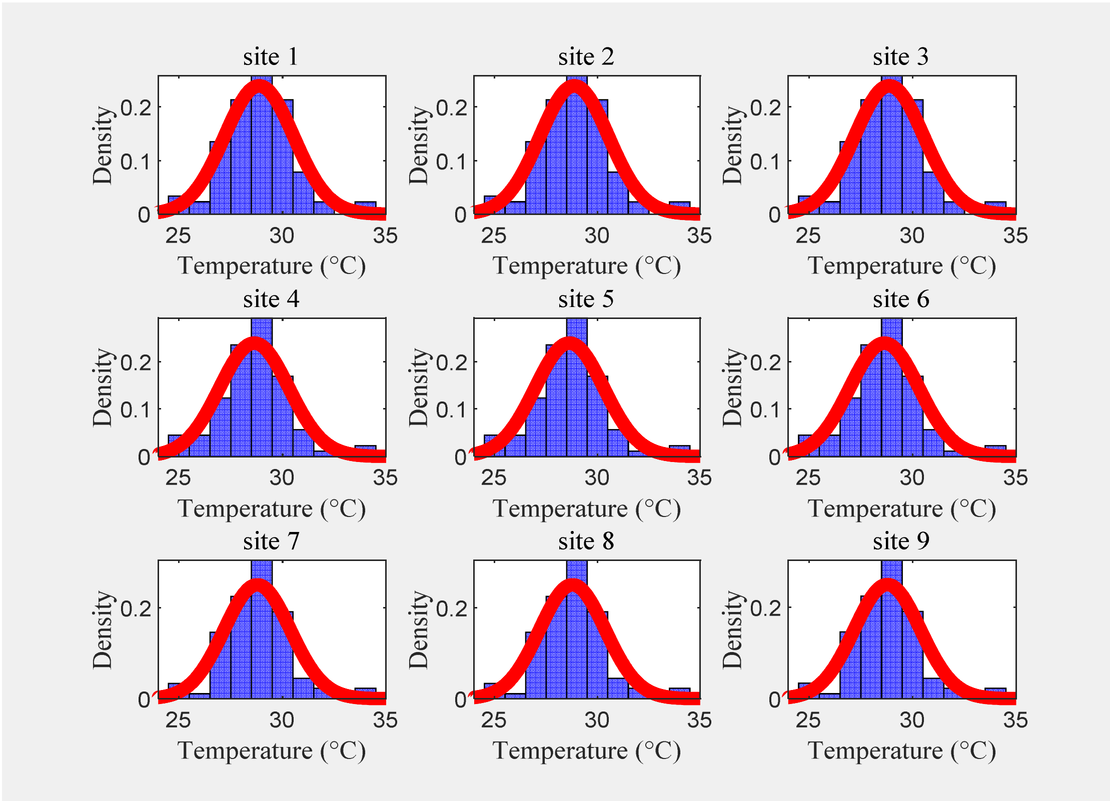

Figure-3: Ambient temperature distribution.

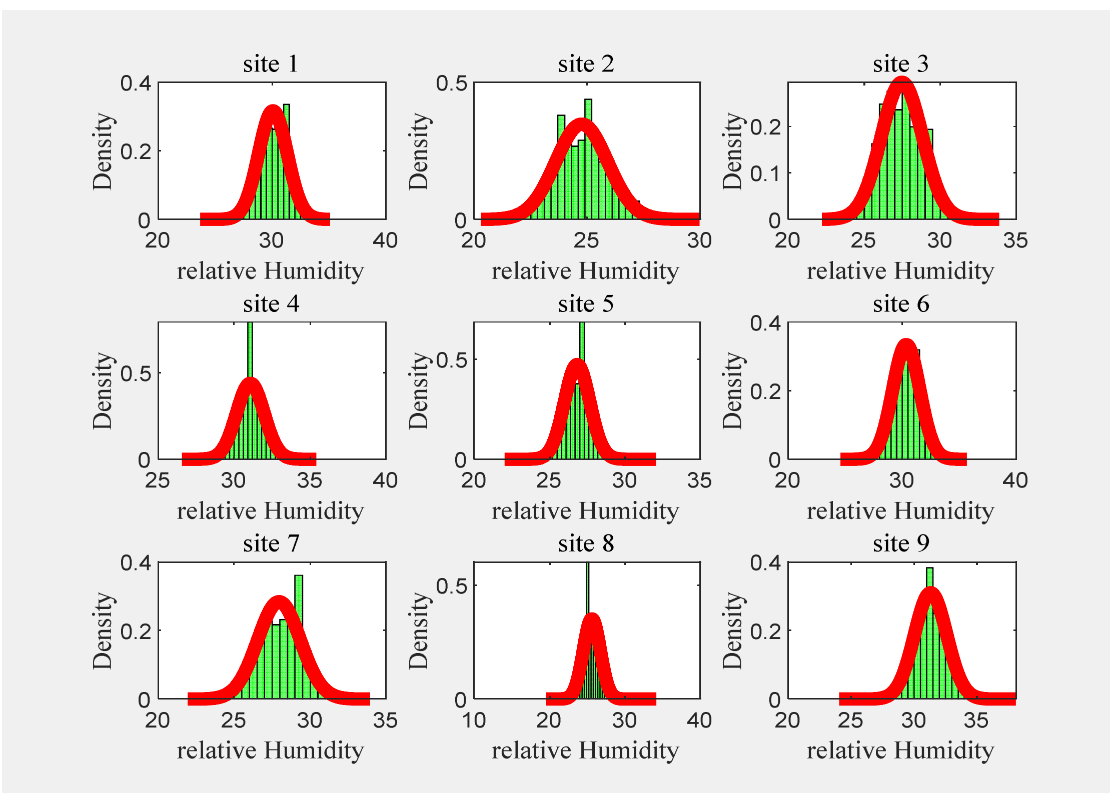

Figure-4: Soil humidity distribution.

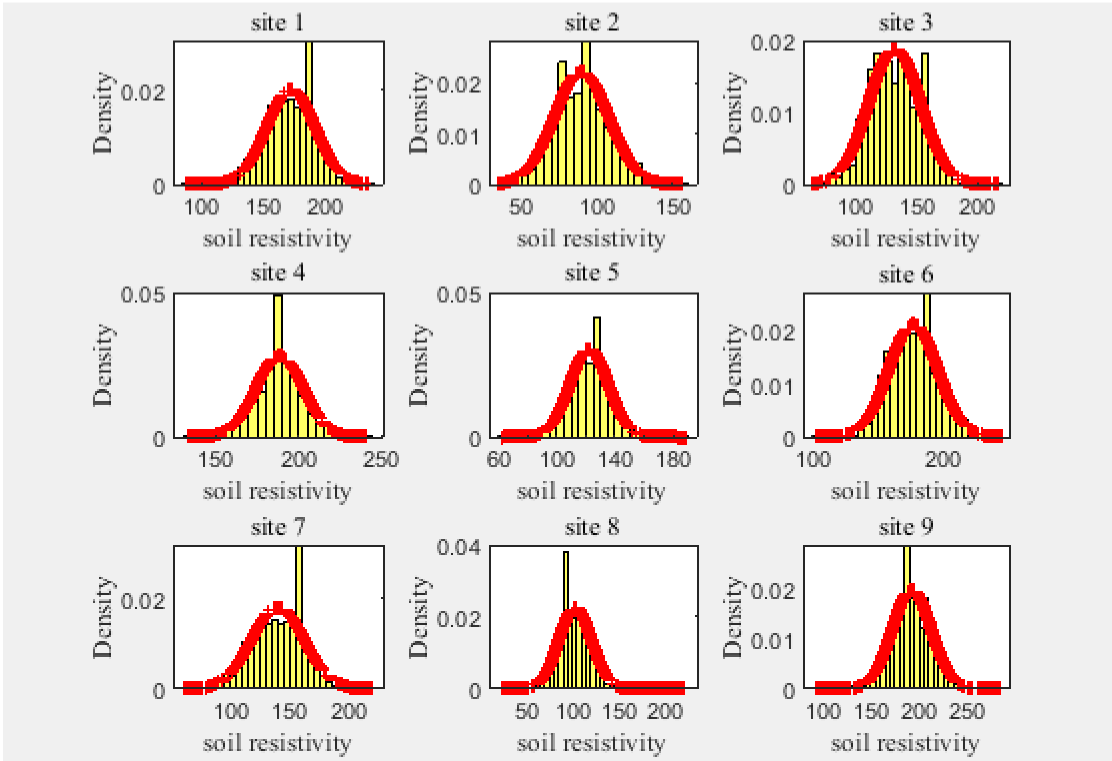

Figure-5: Soil electrical Resistivity distribution.

Table-2: Main parameters of MLP and RBF models for resistivity modeling.

Model	MLP	RBF
Number of layers	2	2
Number of hidden layers	1	1
Activation function of the hidden layer's neurons	Sigmoid	Gaussian
Activation function of the output layer's neurons	Simple linear	Simple linear
Learning Algorithm	Retro propagation of the error	Retro propagation of the error
Algorithm for updating synaptic weights	Levenberg-Marquardt	Levenberg-Marquardt

For the ANFIS model, the inference is Linear-Sugeno type, the membership functions are of Gaussian type and the learning algorithm is that of the retro propagation of the error.

For SVM, we used Platt John's Sequential Minimal Optimization (SMO) algorithm.

Results and discussion

Since for MLP and RBF it is a question of varying the number of the hidden layer's neurons, the results obtained for errors are recorded in Tables-3 and 4.

Table-3: Error Values for the MLP Model.

Hidden layer's neurons	RMSE	RRMSE (%)	MAPE (%)	R^2 (%)
1	23,9199	12,3540	15,9609	52,3242
5	0,0044	0,0011	0,0012	100,0000
10	0,0137	0,0038	0,0043	100,0000
20	0,0181	0,0008	0,0016	100,0000
30	0,0987	0,0017	0,0033	99,9992
40	0,9018	0,0243	0,0412	99,9326
60	0,8895	0,0245	0,0537	99,9430
70	1,2635	0,0738	0,0854	99,8899
80	1,7184	0,0581	0,1284	99,7997
90	1,3187	0,0446	0,1229	99,8725
100	2,4301	0,0524	0,1128	99,6069

Table-4: Error Values for the RBF Model.

Hidden layer's neurons	RMSE	RRMSE (%)	MAPE (%)	R^2 (%)
1	0,9153	0,6008	0,4992	99,9659
5	1,9305	1,2673	0,6593	99,7547
10	3,8589	2,5331	1,2134	98,7926
20	1,3488	0,8854	0,0982	99,8493
30	1,1075	0,7270	0,0719	99,8984
40	3,6266	2,3806	0,1430	98,8915
60	2,1321	1,3996	0,0770	99,6191
70	5,0215	3,2963	0,2167	97,8821
80	4,8484	3,1826	0,1927	98,0248
90	8,6857	5,7016	0,1922	94,0213
100	6,9293	4,5486	0,1534	96,0787

Considering the MAPE error as the main criterion of performance, it appears that for the MLP model, the smallest error (MAPE=0.0012) is obtained with 5 hidden layer's neurons and for the RBF model the smallest error (MAPE = 0.07) is obtained with 30 hidden layer's neurons.

Figure-6 illustrates the variation of the MAPE error as a function of the number of neurons under the hidden layer for each of the two models.

Thus, after exploration with the ANFIS and SVM models, all the results are grouped together in Table-5.

Figures-7, 8, 9, 10 show the correlation between the measured value and the predicted value of the MLP, RBF, ANFIS and SVM models for the validation phase.

It emerges from these validations that the effectiveness of the MLP, RBF and ANFIS models in predicting resistivity exceeds that of SVM. This performance reflects the strength and accuracy of ANFIS and ANN models' outputs through rules that allow them to make the right decisions to calculate outputs.

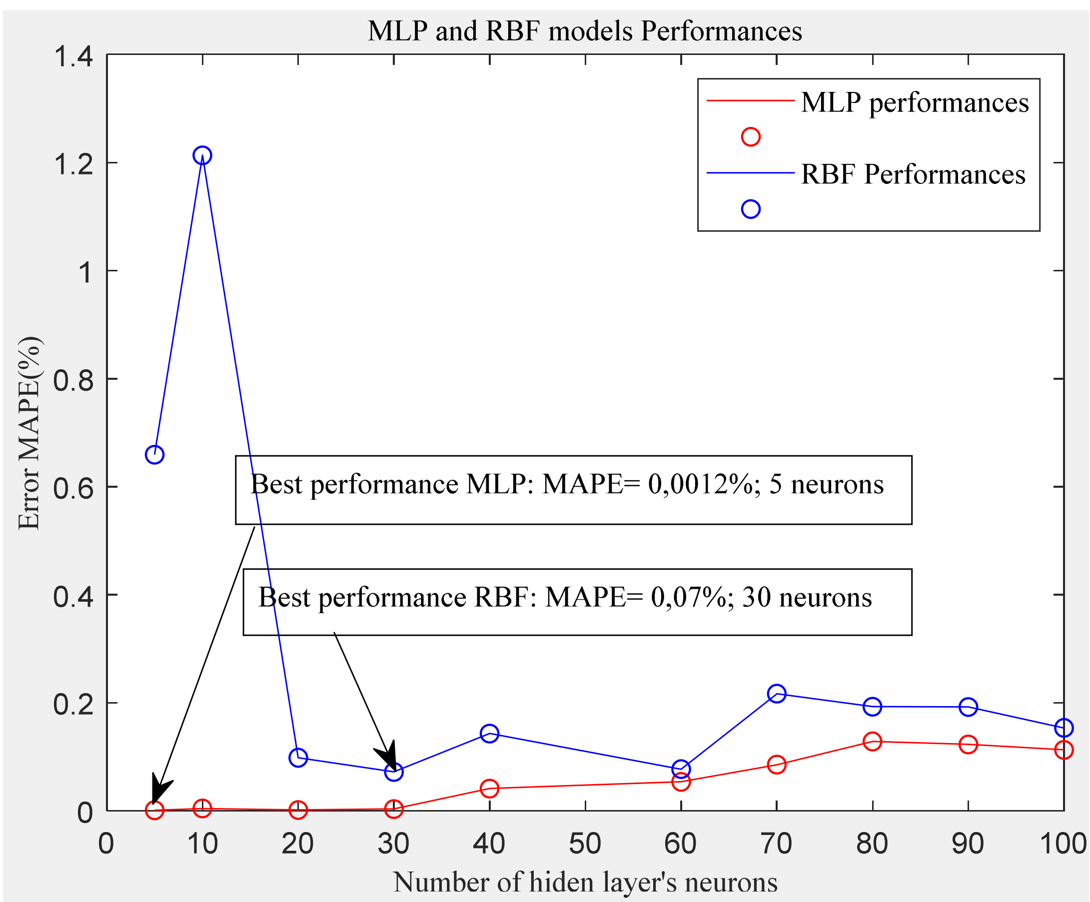

Figure-6: MAPE error according to the number of hidden layer's neurons for MLP and RBF.

Table-5: Validation errors of the MLP, RBF, ANFIS and SVM models.

Model	RMSE	RRMSE (%)	MAPE (%)	R^2 (%)
MLP(with 5 hidden neurons)	0,0044	0,0011	0,0012	100,0000
RBF (with 30 hidden neurons)	1,1075	0,7270	0,0719	99,8984
ANFIS	0,0175	0,0115	0,0011	100,0000
SVM	6,9487	4,5614	2,8947	98,3824

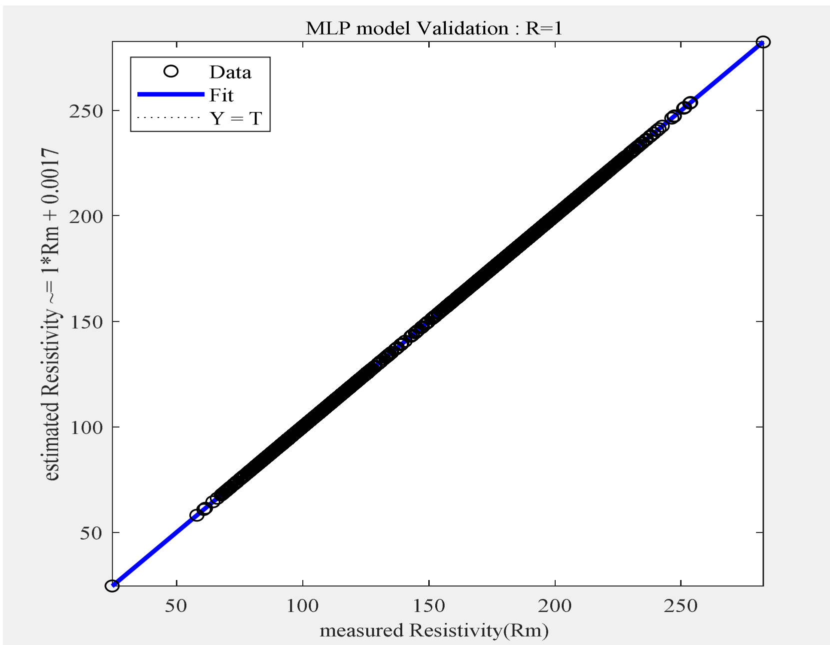

Figure-7: MLP validation.

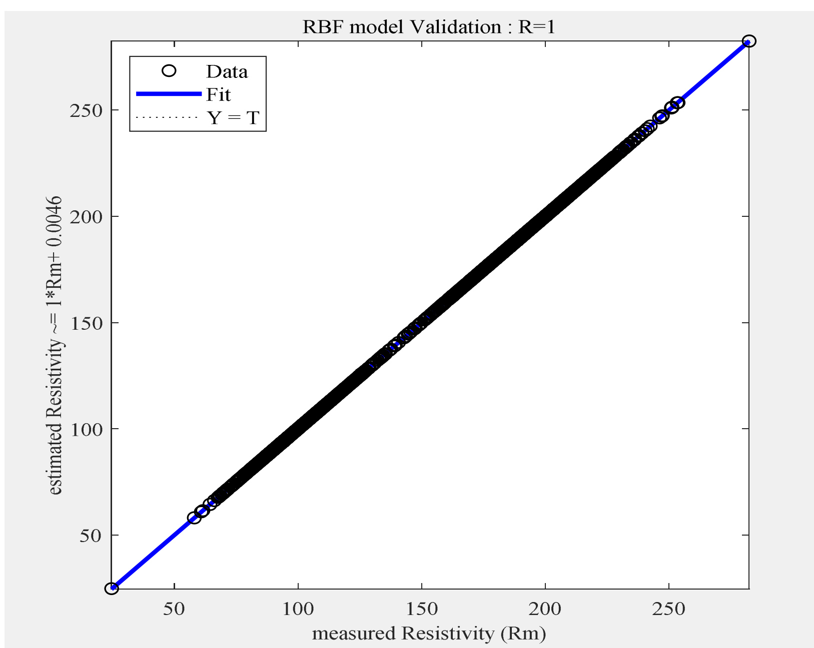

Figure-8: RBF validation

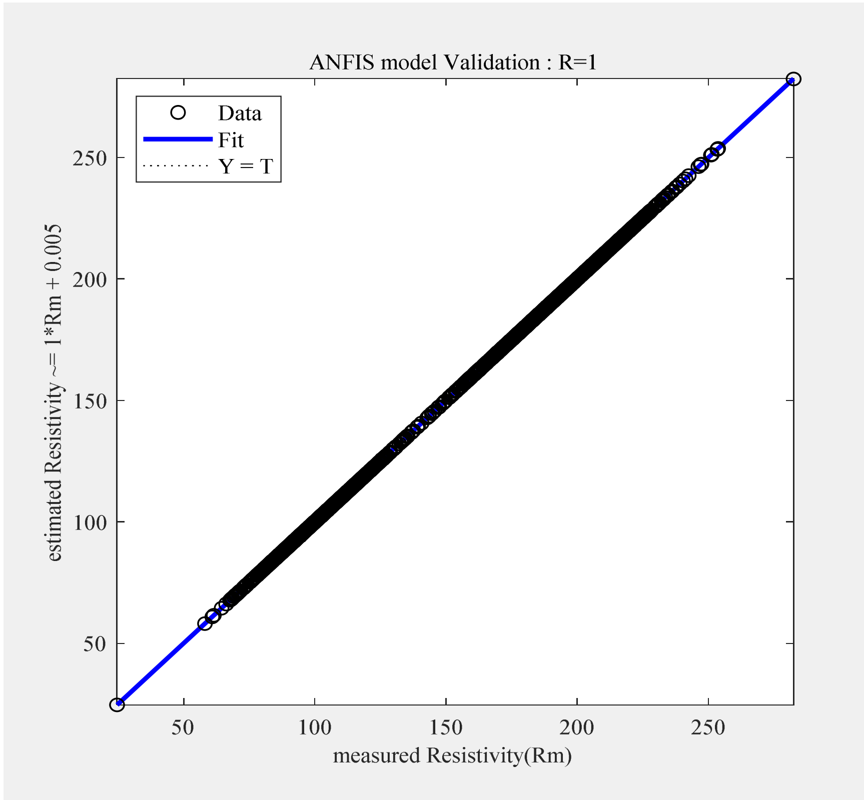

Figure-9: ANFIS validation.

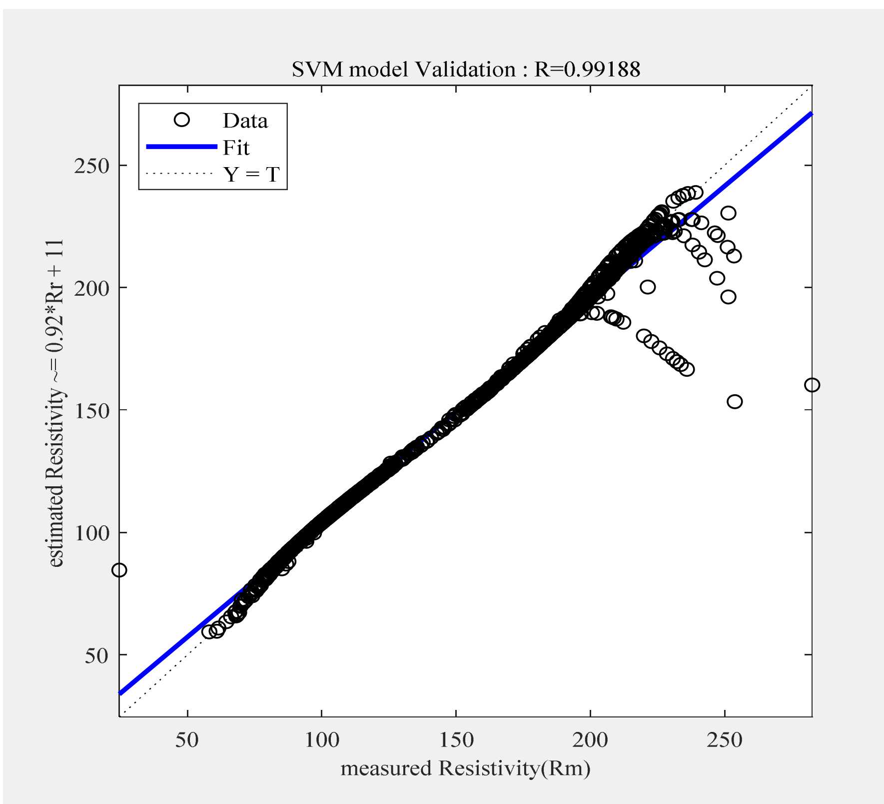

Figure-10: SVM Validation.

Conclusion

The objective of our study was to provide a model that can be used to predict the electrical resistivity of the soil. Thus, after a bibliographic review, we made the choice to explore the approaches by MLP, RBF, ANFIS and SVM.

The results obtained in this study indicate that ANFIS, RBF and MLP are slightly better than SVM. It should also be noted that if the number of hidden layer's neurons must be considered in order to make a choice for an ANN model, the choice must be made on the MLP model which has a small number of neurons.

In general, it should be noted that if a single choice has to be made, it will have to relate to the ANFIS model because it has an advantage over the ANN models, which is the fact that there is no update of synaptic weights and therefore with the error obtained, it is more reassuring.

References

1. Choi J.H. and Lee B.H. (2012). An analysis of conventional grounding impedance based on the impulsive current distribution of a horizontal electrode. *Electric Power Systems Research*, 85, 30-37. http://dx.doi.org/10.1016/j.epsr.2011.07.005
2. Rashad Mohammedeen Kamel, Aymen Chaouachi and Ken Nagasaka (2011). Comparison the Performances of Three Earthing Systems for Micro-Grid Protection during the Grid Connected Mode. *Smart Grid and Renewable Energy*, 2(3). http://dx.doi.org/10.4236/sgre.2011.23024
3. Androvitsaneas V.P., Gonos I.F. and Stathopulos I.A. (2014). Artificial neural network methodology for the estimation of ground enhancing compounds resistance. *IET Science, Measurement & Technology*, 8(6), 552-570. http://dx.doi.org/10.1049/iet-smt.2013.0292
4. Lee J.P., Ji P.S., Lim J.Y., Kim S.S., Ozdemir A. and Singh C. (2005). Earth parameter and equivalent resistivity estimation using ANN. *In IEEE Power Engineering Society General Meeting*, 2597-2602. IEEE. http://dx.doi.org/10.1109/PES.2005.1489485
5. Anbazhagan S. (2015). Athens Seasonal Variation of Grounding Prediction Using Neural Networks. *ICTACT Journal On Soft Computing*, 06(1). http://dx.doi.org/10.21917/ijsc.2015.0154
6. Asimakopoulou F.E., Tsekouras G.J., Gonos I.F., Moronis A.X. and Stathopulos I.A. (2010). An artificial neural network for estimating the ground resistance. In International Conference on Grounding and Earthing & 4th International Conference on Lightning Physics and Effects.
7. Afa J.T. and Anaele C.M. (2010). Seasonal variation of soil resistivity and soil temperature in Bayelsa State. *Am. J. Eng. Appl. Sci*, 3(4), 704-709. http://dx.doi.org/10.3844/ajeassp.2010.704.709
8. Marcin Grabarczyk and Piotr Furmanski (2013). Predicting the effective thermal conductivity of dry granular media using artificial neural networks. *Journal of Power Technologies*, 93(2), 59-66.
9. Mehdaoui A., Chaker A., Zerikat M. and Messikh L. (2009). Development of two neuro-fuzzy models for thecontinuation of the MPPT of the photovoltaic modules UDTS-50 Application to the Adrar site. *Revue des Energies Renouvelables*, 12(2), 257-268.
10. Zaki Abda, Mohamed Chittih and Bilel Zerouali (2015). Modeling extreme flows by artificial neural networks and neuro-fuzzy inference systems (application to Algiers coastal basins). *International Conference on African Large Basin Hydrology River Hammamet*, Tunisia, 26-30th October.
11. Mohammadi K., Shamshirband S., Tong C.W., Alam K.A. and Petković D. (2015). Potential of adaptive neuro-fuzzy system for prediction of daily global solar radiation by day of the year. *Energy conversion and management*, 93, 406-413. http://dx.doi.org/10.1016/j.enconman.2015.01.021
12. Schmitt A., Le Blanc B., Corsini M.M., Lafond C. and Bruzek J. (2001). Artificial neural networks. A promising data processing tool for anthropology. Bulletins and memoirs of the Anthropology Society of Paris, 13(1-2).
13. Warren S. McCulloch and Walter Pitts (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115–133.
14. Gronarz T., Habermehl M. and Kneer R. (2016). Modeling of particle radiation interaction in solid fuel combustion with artificial neural networks. *Journal of Power Technologies*, 96(3), 206-211.
15. Roger Jang J. S. and Sun C. T. (1993). Functional equivalence between radial basis function networks and fuzzy inference systems. *IEEE Trans. on Neural Networks*, 4(1), 156–159.
16. Alby S. and Shivakumar B.L. (2018). A prediction model for type 2 diabetes using adaptive neuro-fuzzy interfacesystem. Biomedical Research, S69-S74. http://dx.doi.org/10.4066/biomedicalresearch.29-17-254
17. Corinna Cortes and Vladimir Vapnik (1995). Support-Vector Networks. Machine Learning, 20(3), 273-297. John Platt (1998).
18. Platt J. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines. https://www.microsoft. com/en-us/research/publication/sequential-minimal-optimization-a-fast-algorithm-for-training-support-vector-machines/
19. Osuna E., Freund R. and Girosi F. (1997). Improved Training Algorithm for Support Vector Machines. *Proceedings of the 1997 IEEE Workshop on Neural Network for Signal Processing*. http://dx.doi.org/10.1109/NNSP.1997.622408