Abdelhalim MAHDJOUBI[1], Boubakeur ZEGNINI[1], Mohammed BELKHEIRI[1]

Amar TELIDJI university of Laghouat, P.O box 37G, Ghardaïa road, Laghouat 03000, Algeria (1)

A new approach using Least Squares Support Vector Machines (LS-SVM) to predict Furan in power transformers

Abstract: *LS-SVM present recently more efficiency in different industrial applications like medicine, engineering and power systems. This paper describes a methodology that was developed for the prediction of Furan in power transformers. The methodology uses as input variables such as the dissolved gases (CO and CO_2). The approach presents the advantage that can reduce the time vs. laboratory tests. The validity of the approach was examined by testing several power transformers. LS-SVM gives a good estimation of results which are validated by experimental tests.*

Streszczenie. *W artykule opisano metodologię prognozowania obecności Furanu w transformatorach energetycznych. Na wejściu podawane są takie parametry jak ilość rozpuszczonych gazów CO i CO_2. Metoda opiera się na wykorzystaniu algorytmów LS-SVM. (**Nowa metoda prognozowania obecności Furanu w transformatorach energetycznych wykorzystująca algorytm LS-SVM**)*

Key words: LS SVM, Furan, modeling, power transformer insulator.
Słowa kluczowe: LS SVM, Furan, modelowanie, moc transformatora izolatorem.

doi:10.12915/pe.2014.02.37

Introduction

Power transformers undergo after many years of utilization different problems; among them the ageing of insulation system. We can find several materials used in insulation system of transformer such as: oil and paper. Testing methods have being developed in recent years. The dissolved gases analysis (DAG), furan derivative and degree of polymerization are the most techniques in practice. These techniques are used to predict the sate of insulators such as oil or paper. The result is minimizing the risk of damage of power transformer and takes more efficiency in network stability.

Power transformer in service is subject to certain electrical, thermal, mechanical and environmental stresses that severely affect its insulation integrity. This reduces the transformer's reliability in service and its life cycle utilization [1].

Predicting the end of life of a transformer is an essential parameter in transformer asset management activities. Moreover, accurate decisions with respect to the end of life save considerable money in the long run and protect the power system against expensive transformer outages. The degradation of solid insulation (paper), can be considered the primary reason for a transformer to reach the end of its life [2]

The gases that are typically found in the transformer insulating oil are Nitrogen (N_2), Oxygen (O_2), Hydrogen (H_2), Carbon dioxide (CO_2), Carbon monoxide (CO), Methane (CH_4), Ethane (C_2H_6), Ethylene (C_2H_4), and Acetylene (C_2H_2) [3]. An investigations made by researchers [4] conclude that the CO and CO_2 accumulation is an essential criterion to see the efficiency of power transformers.

The CO_2/CO ratio is generally considered as an indicator for paper involvement in a transformer fault. The ratio $CO_2/CO < 3$ indicates excessive paper degradation. These can be confirmed by complementing DGA analysis with HPLC (High performance liquid chromatography). HPLC is a well-known analytical technique for the purification, identification, separation and quantification of organic compounds from mixtures. HPLC is the most commonly used method for detection of five furan derivatives, 5-hydroxymethyl-2-furfural (5-HMF), 2-furfurylalcohol (2-FOL), 2-furfural (2-FAL), 2-acetylfuran (2-ACF) and 5-methyl-2-furfural (5-MEF), dissolved in transformer oil [5].

To assess the stability of the furan compounds under the experimental conditions, 2-furfuryl alcohol, 2-furaldehyde, 2-acetylfuran and 5-methylfurfural were added to oil at a level of about 20mg/l, heated in a sealed ampoule and analyzed periodically for furans. The oil was prepared to contain about 1% of oxygen and 10mg/l water [6].

Material and methods

1. Material

We have used in this study one hundred and sixteen transformers mentioned in reference [7]. Those transformers depend to Electricity and gas society of Laghouat City (SONELGAZ). The training and testing data used in the model are collected from the transformers maintenance records with voltage ratings of 60/31.5 kV and 31.5/11 kV and power in range of 20-40 MVA. Some of data are shown in table 1.

Table 1. Simples of the power transformers data used in study

	Tr1	Tr5	Tr12	Tr21	Tr32	Tr49	Tr72	Tr96	Tr106	Tr116
CO(ppm)	96	477	150	172	141	106	389	98	224	1062
CO2(ppm)	1383	6747	3670	1578	3491	1019	1936	877	1368	3885
Furan(ppm)	1.44	27.34	13.88	6.96	14.17	0.56	2.45	0.09	4.35	15.80

2. Methods

Different methods used to identify the quality and goodness of insulation system in power transformers. We can divide these methods in two categories: experimental and analytical with simulation.

- Experimental like: thermal lifetime of transformer insulation which estimates the ambient temperature using the monthly average and the monthly solar clearness index [2]. Duval Triangle method for oil insulated transformers which were developed by Michel Duval of Hydro Quebec's institute of Research. This method is based by using three gases correspond to the increasing levels of energy necessary to generate gases in transformers in service [8]. The experience of an electric power utility in the use of furan-in-oil analysis to detect thermal degradation of cellulose materials in oil filled power transformers [9]. Dissolved gas analysis (DGA); this method was made to diagnosis technology of power insulation condition and it nondestructive, inexpensive and effective. In conclusion the DGA is an efficient tool for diagnosing failure condition in oil-filled electrical equipment [4, 8, 10, 11, 21]. Pulsed amperometric detection method (PAD) to measure furanics in transformer oils in real time; oils were examined by pre-extraction or direct suspension in aqueous measurement solution or by solubilisation and direct PAD measurement in organic solvents [12]. Molecular weight study of cellulose insulation with gel permeation chromatography (GPC); this method is based on molecular weights of derivative cellulose samples and were determined using waters associates chromatography [13]. Strength of cellulose insulation paper with accelerated ageing experiments based on thermal ageing which can be into long-term for low temperatures or into short-term for high temperatures [13, 14].

- Analytical with simulation like: Monte Carlo approach [2]. Artificial neural networks [1, 3, 7]. Support vector machines (SVM) [8, 10, 11, 15-19, 22-24]. Particle swarm optimization (PSO) [9]. Fuzzy logic [4, 20]. We have chosen least square support vectors machines (LS-SVM) like simulation method.

Least Square Support Vector Machines

LS-SVM is applied like an objective method of regression and it is introduced as follows [22-24]:

Consider a given training set

$\{xi, yi\}\phi\in R^2$, $i=1,2,\ldots\ldots N$

With input data x_i, and output data y_i

Regression model can be constructed by using non-linear mapping function $\phi(x_i)$ as shown in (1)

(1) $$y = w^T \phi(x) + b$$

where *w* is the weight vector and *b* is the bias term. As in SVM, it is necessary to minimize a cost function *C* containing a penalized regression error, as follows:

(2) $$\min C(w,e) = \frac{1}{2} w^T w + \frac{1}{2}\gamma \sum_{i=1}^{N} e_i^2$$

Subject to equality constraints

(3) $$y = w^T \phi(x_i) + b + e_i\ , i=1, 2.....,N$$

The first part of this cost function is a weight decay which is used to regularize weight sizes and penalize large weights. Due to this regularization, the weights converge to similar value. The second part of (2) is the regression error for all training data. The parameter *C*, which has to be optimized by the user, gives the relative weight of this part as compared to the first part. The restriction supplied by (3) gives the definition of the regression error. To solve this optimization problem Lagrange function is constructed as

(4) $$L(w,b,e,\alpha) = \frac{1}{2}\|w\|^2 + \gamma \sum_{i=1}^{N} e_i^2 - \sum_{i=1}^{N} \alpha_i \left\{ w^T \phi(x_i) + b + e_i - y_i \right\}$$

where α_i are Lagrange multipliers

The solution of (4) can be obtained by partially differentiating with respect to *w, b,* e_i and α_i which gives:

(5) $$w = \sum_{i=1}^{N} \alpha_i \phi(x_i) = \sum_{i=1}^{N} \gamma e_i \phi(x_i)$$

where a positive definite Kernel is used as follows:

(6) $$K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$$

An important result of this approach is that the weights *(w)* can be written as linear combinations of the Lagrange multipliers with the corresponding data training (x_i). Putting the result of (5) into (1), the following result is obtained:

(7) $$y = \sum_{i=1}^{N} \alpha_i \phi(x_i)^T \phi(x) + b$$

For a point y_i to be evaluated it is:

(8) $$y_i = \sum_{i=1}^{N} \alpha_i \phi(x_i)^T \phi(x_j) + b$$

The vector follows from solving a set of linear equations

(9) $$A\begin{bmatrix}\alpha \\ b\end{bmatrix} = \begin{bmatrix} y \\ 0\end{bmatrix}$$

where A is a square matrix given by

(10) $$A = \begin{bmatrix} K + \frac{1}{\gamma} & 1_N \\ 1_N^T & 0 \end{bmatrix}$$

Hence, the solution is given by:

(11) $$\begin{bmatrix}\alpha \\ b\end{bmatrix} = A^{-1}\begin{bmatrix} y \\ 0\end{bmatrix}$$

From (10) to (11), we can see that usually all Lagrange multipliers (the support vectors) are nonzero, which means that all training objects contribute to the solution. In contrast with standard SVM the LS-SVM solution is usually not sparse. However, a sparse solution can be easily achieved via pruning or reduction techniques.

Depending on the number of training data set either direct solvers can be used or an iterative solver such as conjugate gradients methods, in both cases with numerically reliable methods. In applications involving nonlinear regression it is enough to change the inner product of $(\phi(x_i), \phi(x_j))$ (7) by a kernel function and the ij^{th} element of matrix *K* equals to (5). This leads to the following nonlinear regression function:

(12) $$y = \sum_{i=1}^{N} \alpha_i K(x_i, x) + b$$

For a point x_j to be evaluated it is:

(13) $$y_j = \sum_{i}^{N} \alpha_i K(x_i, x_j) + b$$

For our study based on LS-SVM, RBF kernel (radial basis function), is used because it is simple and give good results. RBF is defined by:

(14) $$K(x_i, x_j) = \exp\left(- \frac{\|x_i - x_j\|^2}{\sigma_{sv}^{\ 2}} \right)$$

where σ_{sv}^2 the squared variance of the Gaussian function, which should be optimized by the user, to obtain the support vector. It should be stressed that it is very important to do a careful model selection of the tuning parameters, in combination with the regularization constant γ, that's in order to achieve a good generalization model. Averaging the validation errors over the test data gives a prediction of the generalization error. The RMSE is given as:

(15) $$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(Y_{test,k} - Y_{predict,k}\right)^2}$$

where N is the number of data patterns in the data set. $Y_{predict,k}$ Indicates the predicted value (furan in this study) and $Y_{test,k}$ is the testing value of datum point k. The R^2 correlation parameter is calculated by:

$$R^2 = 1 - \sum_{k=1}^{N}\left[\frac{\left(Y_{test,k} - Y_{predict,k}\right)}{Y_{test,k}}\right]^2 \tag{16}$$

Mean absolute and max absolute errors are given by (17) and (18) respectively:

$$MEAE = \frac{1}{N}\sum_{k=1}^{N} abs\left(Y_{test,k} - Y_{predict,k}\right) \tag{17}$$

$$MAAE = \max\left(abs\left(Y_{test,k} - Y_{predict,k}\right)\right)_{k=1}^{N} \tag{18}$$

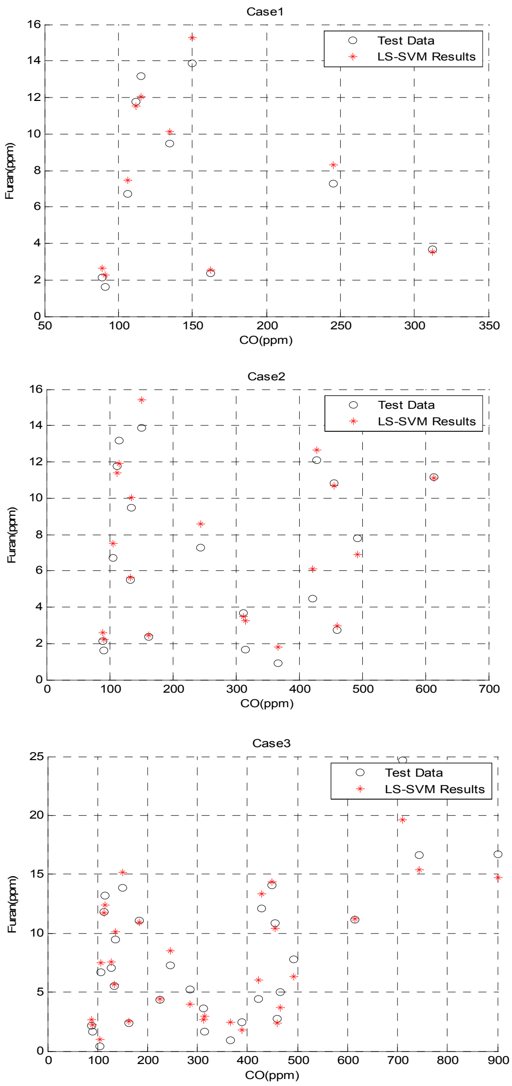

Fig 1. Comparison of testing data and estimated values for CO contents

Results and discussion

All data about one hundred and sixteen power transformers are collected and used by extracting some data to be used in testing as following in table2

Table 2. Overview of training and testing data

	Training data	Testing data
Case1	106	10
Case2	96	20
Case3	86	30

In order to get a good optimisation of results, we are proceeding to varying the tow parameters γ and σ^2. Finally we got the optima values γ =5.6569 and σ^2 =0.17678. The results are shown in Fig.1, 2, 3. In Fig.1, we had done the variation of furan content according to CO values. In Fig.2, we can see the variation of furan according to CO_2 values; in addition we have plotted the model of LS-SVM in Fig.3.

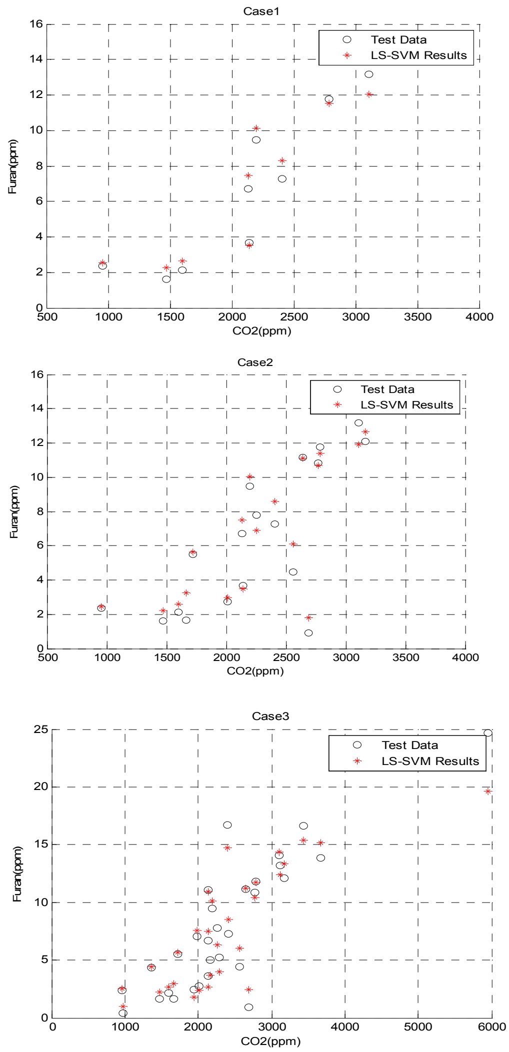

Fig 2. Comparison of testing data and estimated values for CO_2 contents

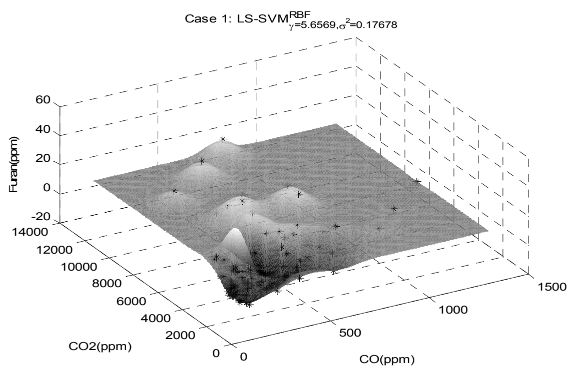

Fig 3. Model of LS-SVM results with optimized values

Fig.1 and 2 show a good results confirmed by the testing values. The trained model is then assumed reliable. The search is tuned to a finer search in the region where the predicted RMSE is lowest possible. The minimum

RMSE value indicates the optimum LS-SVM parameters Fig. 3. Using three cases is in order to see the robustness of LS-SVM method. Some of statistical parameters were shown in table 3.

Table3 Statistical model validation parameters

	R^2 Training	R^2 Test	RMSE Training	RMSE Test	MEAE	MAAE
Case1	0.9991	0.9695	0.2413	0.7832	0.6714	1.4075
Case2	0.9992	0.9639	0.2478	0.8534	0.6674	1.6404
Case3	0.9992	0.9461	0.2494	1.3216	0.9389	5.0325

According to the table 3, we can see clearly that LS-SVM method offers a good accuracy. In our study this accuracy of testing data changes from 96.95% to 94.61% according to the treated cases. The results are less Satisfactory with the increase of testing data set. The accuracy done by LS-SVM model is also better than for example given in [1] which used ANN method.

Conclusion

In this paper, the LS-SVM model is developed in order to determine the relationship between furan content as function of gases (CO and CO_2). The comparisons and the results presented in this investigation clearly reveal the capability of the proposed LS-SVM modelling strategy in predicting furan values of different power transformers by giving detailed information of input insulation system. Predicting furan by this method enhance the efficiency of simulation goodness which complete obviously experimental tests. By collecting more data we can apply this method to show the effect of several parameters and give to LS-SVM more robustness. Finally, experiments and tests are time consuming and increase the cost of the system. To overcome such a problem researcher groups in high voltage engineering proposed some mathematical models based either on physical modeling using electrical equivalent models or on mathematical regressions using artificial intelligent approximates and give more efficiency to simulation methods.

REFERENCES

[1] Refat Atef Ghunem, Ayman H. El- Hag, Khaled Assaleh , prediction of furan content in transformer oil using Artificial Neural Networks (ANN), International symposium on electrical insulation (ISEI), IEEE 2010, pp:1-4.

[2] Ahmed E.B. Abu-Elanien, ,M.M.A. Salama, A Monte Carlo approach for calculating the thermal lifetime of transformer insulation, Electrical Power and Energy Systems 43(2012), pp: 481–487.

[3] Deepika Bhalla , Raj Kumar Bansal, Hari Om Gupta, Function analysis based rule extraction from artificial neural networks for transformer incipient fault diagnosis, Electrical Power and Energy Systems 43 (2012), pp:1196–1203.

[4] Hasmat Malik, Tarkeshwar, R.K. Jarial, Make Use of DGA to Carry Out the Transformer Oil-Immersed Paper Deterioration Condition Estimation with Fuzzy-Logic, Procedia Engineering 30 (2012), pp: 569 – 576.

[5] Dijana Vrsaljko, Veronika Haramija, Andela Hadzi-Skerlev, Determination of phenol, m-cresol and o-cresol in transformer oil by HPLC method, Electric Power Systems Research 93 (2012), pp: 24– 31.

[6] A.M.Emsley,X.Xiao, R.J.Heywood and M.Ali, Degradation of cellulosic insulation in power transformers. Part 2: Formation of furan products in insulating oil, IEE Proc.-Sci. Meas. Techno, Vol 147, No. 3. May 2000 pp:110-114.

[7] Kheira Djeridane, "Contribution à la recherché des traceurs spécifiques à la dégradation du papier dans les transformateurs de puissance" , Master dissertation , Amar TELIDJI university of Laghouat- June 2012.

[8] M. Duval, A review of faults detectable by gas-in-oil analysis in transformers, IEEE Electrical Insulation Mag., 18 (2002), 8-17

[9] W. McDermid, D.H. Grant, Use of Furan-in-Oil Analysis to Determine the Condition of Oil Filled Power Transformers, 2008 International Conference on Condition Monitoring and Diagnosis, Beijing, China, April 21-24, 2008 – IEEE, pp: 479 – 481.

[10] Xiao-Yun Sun, Dong-Hui Liu, Jian-Peng Bian, The Study Of Fault Diagnosis Model Of DGA For Oil-Immersed Transformer Based On Svm Active Learning And K-L Feature Extracting, Proceedings of the Seventh International Conference on Machine Learning and Cybernetics,Kunming,12-15 July 2008 – IEEE, pp: 1510 – 1514.

[11] XIE Hong-ling, LI Nan, LU Fang-cheng, XIE Qing, Application of LS-SVM by GA for Dissolved Gas Concentration Forecasting in PowerTransformer Oil, Power and Energy Engineering Conf., 2009. APPEEC 2009, IEEE 2009, pp: 1-4.

[12] Timothy Bosworth, Steven Setford, Richard Heywood, Selwayan Saini, Pulsed amperometric detection of furan compounds in transformer oil, Analytica Chimica Acta 450 (2001) , pp: 253–261.

[13] D. J. T. Hill, T. T. Le, M. Darvenizab, T. Sahab, A study of degradation of cellulosic insulation materials in a power transformer, part 1. Molecular weight study of cellulose insulation, Polymer Degradation and Stability *48* (1995), pp: 79-81.

[14] D. J. T. Hill, T. T. Le, M. Darvenizab, T. Saha, A study of degradation of cellulosic insulation materials in a power transformer. Part 2: tensile strength of cellulose insulation paper,Polymer Degradation and Stability 49 (1995), pp: *329-435.*

[15] Xiaoyun Sun, Guoqing An, Ping Fu, Fault Diagnosis Model of Power Transformer Based on an Improved Binary Tree and the Choice of the Optimum Parameters of Multi-class SVM, Intelligent Computing and Intelligent Systems, 2009. ICIS 2009, IEEE International Conference on 2009, vol.4, pp.567-

[16] Xiaoyun Sun, Jianpeng Bian, Donghui Liu, Zhenquan Li, The Study of Transformer Fault Diagnosis Based on Means Kernel Clustering and SVM Multi-class Object Simplified Structure, Proceedings of the 7th World Congress on Intelligent Control and Automation June 25 - 27, 2008, Chongqing, China, IEEE 2009, pp:5158-5161.

[17] M. Elsamahy, M. Babiy *,* An Intelligent Approach using SVM to Enhance Turn-to-Turn Fault Detection in Power Transformers, IEEE-2012 pp:255-260.

[18] Li Yanqing , Huang Huaping, Li Ningyuan, Xie Qing, Lu Fangcheng, The Application of the IGA in Transformer Fault Diagnosis Based on LS-SVM, Power and Energy Engineering Conference (APPEEC), 2010 Asia-Pacific , IEEE 2010, pp:1-4.

[19] Tsair-Fwu Lee1,2, Ming-Yuan Cho1, Chin-Shiuh Shieh1, and Fu-Min Fang, Particle Swarm Optimization-Based SVM Application: Power Transformers Incipient Fault Syndrome Diagnosis, 2006 International Conference on Hybrid Information Technology (ICHIT'06), IEEE 2006, pp: 468 – 472.

[20] A. Abu-Siada*,* Sin P. Lai*,* Syed M. Islam*,* A Novel Fuzzy-Logic Approach for Furan Estimation in Transformer Oil, IEEE Tans. on power delivery, vol. 27, no. 2, april 2012, pp:469-474.

[21] Miloš Božić, Miloš Stojanović, Zoran Stajić, and Đukan Vukić, "Power Transformer Fault Diagnosis based on Dissolved Gas Analysis with Logistic Regression," PRZEGLĄD ELEKTROTECHNICZNY, 6 (2013), 83-87

[22] J.A.K Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, Least squares support vector machines, World scientific publishing, Singapore, 2002.

[23] B. Zegnini, A.H Mahdjoubi, M. Belkheiri, A Least Squares Support Vector Machines (LS-SVM) Approach for Predicting Critical Flashover Voltage of Polluted Insulators, 2011 IEEE pp:403-406.

[24] Muhsin Tunay Gencoglu, Murat Uyar, Prediction of flashover voltage of insulators using least squares support vector machines, Expert Systems with Applications 36 (2009), pp: 10789–10798.

***Authors**: Abdelhalim Mahdjoubi, Laboratory of studies and development of the Semiconducting and dielectric materials , Amar TELIDJI university of Laghouat, P.O box 37G, Ghardaïa road, Laghouat 03000, Algeria, E-mail: mahd6875@gmail.com; prof. dr Boubakeur Zegnini, Laboratory of studies and development of the Semiconducting and dielectric materials , Amar TELIDJI university of Laghouat, E-mail: bakzegni@yahoo.fr; dr Mohammed Belkheiri, Laboratory of telecommunications, signals and systems, Amar TELIDJI university of Laghouat,E-mail: m.belkheiri@mail.lagh-univ.dz.*