

# Transformer Top-Oil Temperature Modeling and Simulation

T. C. B. N. Assunção, J. L. Silvino, and P. Resende

**Abstract**—The winding hot-spot temperature is one of the most critical parameters that affect the useful life of the power transformers. The winding hot-spot temperature can be calculated as function of the top-oil temperature that can be estimated by using the ambient temperature and transformer loading measured data. This paper proposes the estimation of the top-oil temperature by using a method based on Least Squares Support Vector Machines approach. The estimated top-oil temperature is compared with measured data of a power transformer in operation. The results are also compared with methods based on the IEEE Standard C57.91-1995/2000 and Artificial Neural Networks. It is shown that the Least Squares Support Vector Machines approach presents better performance than the methods based in the IEEE Standard C57.91-1995/2000 and artificial neural networks.

**Keywords**—Artificial Neural Networks, Hot-spot Temperature, Least Squares Support Vector, Top-oil Temperature.

## I. INTRODUCTION

POWER transformers are high cost important equipment used in the transmission and distribution of the electric energy. Its right performance is important for the electric systems operation, since the loss of a critical unit can generate great impact in safety, reliability and cost of the electric energy supply. One of the main factors adopted for monitoring transformers operation conditions are its internal temperatures, specially the winding hot-spot temperature (HST) and the top-oil temperature (TOT), which affect the isolation aging and, consequently, the useful life of the equipment. The thermal modeling is considered as one of most important aspects for monitoring of the power transformer operation conditions. Calculated values of TOT and HST can be to provide a diagnostic of the equipment conditions, and to indicate possible abnormalities, reducing the risk of defects, and avoiding the problems generated by the emergency operations. There are several methods used for calculation of the transformer internal temperatures.

According to Jardini [1], the method of the IEEE Standard C57.91-1995/2000 [2] is the more widely used, and it provides reliable results over transformers in operation. In the IEEE Standard C57.91-1995/2000, the thermal behavior of the transformers is represented by means of a first order model. In the G Annex of the IEEE Standard C57.91-1995/2000 [2] the TOT and HST are determined from the characteristic data of the transformer. In addition to this technique, the estimation of HST and TOT can be obtained by means of other methods [3], [4], [5]. For this purpose Artificial Neural Networks (ANN) can be used due to its learning capacity in the modeling complex and nonlinear relations [6]. ANN is submitted to a training process from real cases, and then handling appropriately new supplied data. The most popular ANN configuration is the multi-layer feedforward network that have been applied successfully to solve some difficult and assorted problems including nonlinear system identification and control, financial market analysis, signal modeling, power load forecasting etc. Several ANN structures have been proposed by researchers that can be classified as static (SNN), dynamic temporal processing (TPNN) and recurrent (RNN). Recently, the Support Vector Machine (SVM) has been proposed as a new and promising technique for classification and regression of the linear and nonlinear systems. The Least Squares Support Vector Machines (LS-SVM) is a learning machine proposed in [7] corresponding a modified version of the SVM. Like the SVM, LS-SVM can be used in classification problems and approximation functions. Its formulation is based on a problem of binary classification that can be extended to approximation problems that involve more than two classes [8], [9]. The main characteristic of LS-SVM is smallest computational cost in relation to the SVM, without loss in the quality of the solutions. The LS-SVM training is based on solving a system of linear equations. In this paper, the TOT will be estimated using the ANN and LS-SVM, and also it will be calculated by Annex G of the IEEE Standard C57.91-1995/2000 [2].

## II. FUNDAMENTAL MODEL

The methodology proposed in Annex G of the IEEE Standard C57.91-1995/2000 [2], based on Pierce's researches [10], suggests an alternative method for the calculation of temperatures, and it depends on the transformers project or construction characteristics (winding constant time, oil time constant, oil volume, weight of the tank, core and winding), and also it depends of the ambient temperature (TA) and of the transformer loading. The errors, in the calculation of TOT

Manuscript received September, 2006.

T. C. B. N. Assunção, is with the Department of Electric Engineering, Federal University of São João del-Rei (UFSJ), Brazil (e-mail: bessa@ufsj.edu.br).

J. L. Silvino is with the Department of Electronics Engineering, Federal University of Minas Gerais (UFMG), Brazil (e-mail: silvino@cpdee.ufmg.br).

P. Resende is with the Department of Electronics Engineering, Federal University of Minas Gerais (UFMG), Brazil (e-mail: pr@cpdee.ufmg.br).

and consequently in the calculation of HST, can be caused by the poor data, because in practice, many variables required for the calculation of these temperatures are not measured: the wind speed and direction, the solar radiation, the rain/evaporative cooling, cloud cover, humidity, transformer internal oil flows and the state and the ventilation type and also the formulation of thermal model of the transformer. Another source of error in the estimate of the top-oil temperature comes from erroneous data, for example, the estimate of the top-oil temperature is dependent of TA measurements, and Tylavsky [11] point out that in practice, the average TA is evaluated or merely measured at places that do not correspond to actual operation condition of the transformer. Depending on the place chosen for measuring the ambient temperature, errors up to 10 °C are found. The assumptions considered in this model also may produce errors in the determination of the transformers internal temperatures, and therefore becomes necessary to employ more precise methods. For this reason, the objective of this paper is to use alternative methods (ANN and LS-SVM) for obtaining more precise values of TOT, and consequently of HST. The estimated values are then compared with measured temperature for an actual transformer.

### III. ARTIFICIAL NEURAL NETWORK

ANN has been established as a useful tool for regression problems, mainly for pattern recognitions and function approximations. An important characteristic of the ANN is that is not necessary to obtain a complete knowledge about the relations among the variables involved in the problem.

The static neural network (SNN) is implemented as one nonlinear function of the following form:

$$\hat{y}_k = f_{snn}(x_k) \quad (1)$$

The temporal neural networks are classified in two basic types: non recurrent neural network (TPNN) and recurrent neural network (RNN). The inputs and outputs relationships of TPNN and RNN can be written as nonlinear functions given by (2) and (3), respectively:

$$\hat{y}_{k+1} = f_{rnn}(x_k, x_{k-1}, x_{k-2}, \dots, x_{k-d}) \quad (2)$$

$$\hat{y}_{k+1} = \begin{pmatrix} x_k, x_{k-1}, x_{k-2}, \dots, x_{k-d}, y_k, y_{k-1}, y_{k-2}, \\ \dots, y_{k-q} \end{pmatrix} \quad (3)$$

where  $\hat{y}_k$  = kth training output,  $x_k$  = kth training input vector, d and q are the number of input and output temporal delay lines.

### IV. LEAST SQUARES SUPPORT VECTOR MACHINES

Least Squares Support Vector Machines (LS-SVM) is a method used for solving non-linear classification or modeling

problems and has been applied to classification, function estimation and nonlinear system optimal control problems. The basis of the method is the mapping of all available data points to a feature space, thus transforming the problem into a simple linear problem. LS-SVM expresses the training in terms of solving a linear set of equations.

#### A. Estimation Function

Given a training set of  $N$  points  $\{x_K, y_K\}_{K=1}^N$ , with input data  $x_K \in R^n$ , and output data  $y_K \in R$ , the LS-SVM model for estimation function has the following representation,

$$\hat{y}(x) = \sum_{K=1}^N \alpha_K K(x, x_K) + b \quad (4)$$

where  $\alpha_i$  are positive real constants and  $b$  is a real constant and comprise the solution to the linear system.  $K(.,.)$  is called the kernel function that is used for the realization of an implicit mapping of the input data into a high-dimension feature space. In this paper the Radial Basis Function (RBF) kernel has been chosen since it tends to give good performance under general smoothness assumptions. The RBF function Kernel is given by:

$$K(x, x_K) = \exp\left(-\frac{\|x - x_K\|^2}{2\sigma^2}\right) \quad (5)$$

where  $\sigma$  is a parameter specifying the width of the kernel.

In order to make an LS-SVM model with the RBF Kernel, it is necessary to calculate the  $\gamma$  regularization parameter in the algorithm, determining the trade-off between the fitting error minimization and smoothness of the estimated function, and also to calculate the  $\sigma$  kernel function parameter.

The temporal LS-SVM model is:

$$\hat{y}_{k+1} = f_{lssvm}(x_k, x_{k-1}, x_{k-2}, \dots, x_{k-d}) \quad (6)$$

The recurrent LS-SVM model is:

$$\hat{y}_{k+1} = f_{lssvm}\left(x_k, x_{k-1}, x_{k-2}, \dots, x_{k-d}, \hat{y}_k, \hat{y}_{k-1}, \dots, \hat{y}_{k-q}\right) \quad (7)$$

where the outputs (estimated values) are reinserted in the input vector.

### V. SIMULATION RESULTS

This section presents the estimation results of TOT using ANN, LS-SVM and the IEEE method. It is also presented the effect of the TOT in the calculation of HST.

In order to implement the methods it was used the experimental data set illustrated in Fig. 1, and the transformer data presented in the Table I.

The experimental data illustrated in Fig. 1 corresponds to the measured values for thirty days operation of the transformer.

TABLE I  
CHARACTERISTICS OF THE TRANSFORMER

Nameplate Rating	30/40 MVA
$V_{\text{primary}}/V_{\text{secondary}}$	138/13.8 kV
Iron Losses	17.8 kW
Cooper Losses	244.9 kW
Type of Cooling	ONAN/ONAF

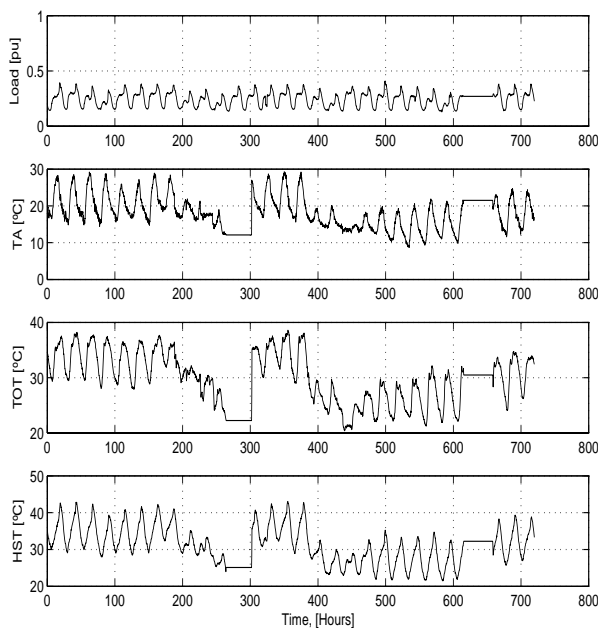


Fig. 1 Experimental Data: Load pu, Ambient Temperature °C, Top-oil Temperature °C, Hot-spot Temperature °C

#### A. TOT Calculation using G Annex of the IEEE Standard

In Fig. 2, it is illustrated estimated values of TOT calculated from the IEEE model with the actual values of TOT and the testing errors (that is defined as the difference between the estimated and actual values of TOT).

The effect of TOT in the calculation of HST is analyzed below. HST will be calculated of two way; firstly will be used the equations of the G Annex and the measured TOT (Fig. 1), and next the equations of the G Annex and the calculated TOT (G Annex).

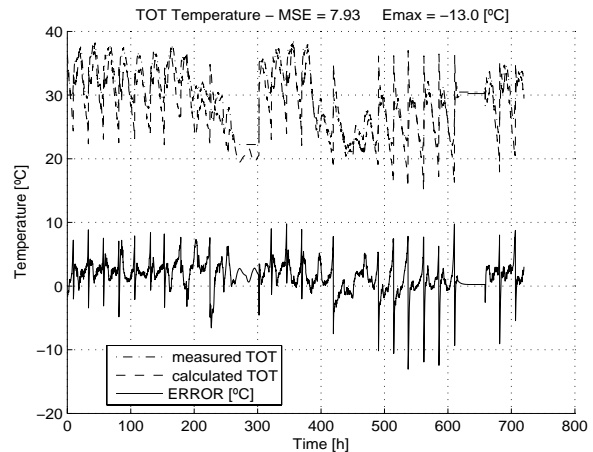


Fig. 2 Actual and estimated values of TOT using IEEE model with prediction error

Fig. 3 shows the estimated HST using the measured value of TOT.

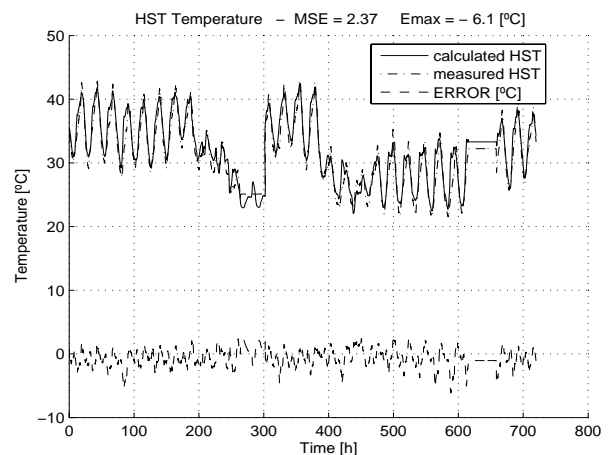


Fig. 3 Actual and estimated HST with prediction error (measured TOT)

Fig. 4 shows the estimated HST obtained from the calculated value of TOT.

Table II presents the obtained results of the MSE (mean square error) and Emax (maximum difference between estimated and measured temperatures in Celsius degrees), showing that a more accurate value of TOT results in better estimate of HST.

#### B. TOT Calculation using ANN

In this section, it is used a two layers feedforward structure for the ANN, using the ambient temperature and the loading as the input while TOT is considered as output. The hyperbolic tangent function is used as activation function for both layers.

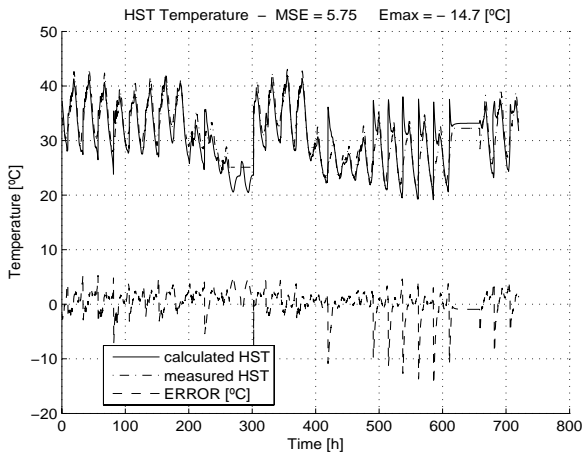


Fig. 4 Actual and estimated HST with prediction error (calculated TOT)

TOT	HST (MSE)	HST (EMAX)
Calculated	5.75	-14.7
Measured	2.37	-6.1

The algorithm used for ANN training is the Levenberg-Marquardt (LM), considering 100 epochs and assuming a MSE goal as 0.001. The Levenberg-Marquardt algorithm was chosen since it takes less CPU time and it is more stable in all the training tasks when compared to other algorithms. The number of hidden nodes ( $n_h$ ) is varied from 2 to 20, choosing the result that provides better training and testing errors. To eliminate the random effects of arbitrary initialization of network weights, ten training process were executed, and therefore the error performance was averaged over ten runs for a given network. The data was normalized into the range of  $[-1, +1]$ . The experimental data were separated in two groups, first 40% data samples for model building/training and the remaining 60% samples for testing. The TOT was estimated using routines in the Neural Toolbox of Matlab [12]. The best results for TOT obtained from the ANN are summarized in Table III.

TOT	MSE	EMAX
SNN ( $n_h = 2$ )	6.62	-6.6
SNN ( $n_h = 3$ )	6.60	6.9
SNN ( $n_h = 4$ )	5.34	6.4
SNN ( $n_h = 5$ )	5.24	6.1
SNN ( $n_h = 6$ )	5.62	-6.5
TPNN ( $d = 1, n_h = 5$ )	4.91	6.5
TPNN ( $d = 2, n_h = 4$ )	4.41	6.3
TPNN ( $d = 3, n_h = 3$ )	3.87	-7.8
RNN ( $d = q = 1, n_h = 5$ )	2.76	-4.7
RNN ( $d = q = 2, n_h = 2$ )	3.58	-5.9
RNN ( $d = q = 3, n_h = 3$ )	3.64	-5.8
RNN ( $d = q = 4, n_h = 4$ )	2.76	-4.7

For SNN it was observed that better training and testing performance with 5 hidden nodes, obtaining MSE = 5.24 and Emax = 6.1 °C. For TPNN it was compared the results by using the numbers of tapped delay lines as  $d = 1, 2, 3$ . It was verified that the MSE error decreases reasonably compared to that obtained by SNN. It was observed better training and testing performance with 3 hidden nodes and  $d = 3$ , obtaining MSE = 3.87 and Emax = -7.8 °C. For RNN the results was also compared by using the numbers of tapped delay lines as  $d = q = 1, 2, 3, 4$ , and it was observed that better training and testing performance was obtained with  $d = q$ . The results indicate that the better training and testing performance was obtained with 5 hidden nodes and  $d = q = 1$ , resulting in MSE = 2.76 and Emax = -4.7 °C.

Table IV presents the results of the implemented ANN, showing that RNN gives better results than SNN and TPNN.

TABLE IV  
COMPARISON BETWEEN THE MSE AND EMAX OF THE BETTER RESULTS FOR THE IMPLEMENTED ANN

ANN	MSE	EMAX
SNN ( $n_h = 5$ )	5.24	6.1
TPNN ( $d = 3, n_h = 3$ )	3.87	-7.8
RNN ( $d = q = 1, n_h = 5$ )	2.76	-4.7

In the Fig. 5 shows the performance of the RNN with 5 hidden nodes,  $d = q = 1$  and the prediction error.

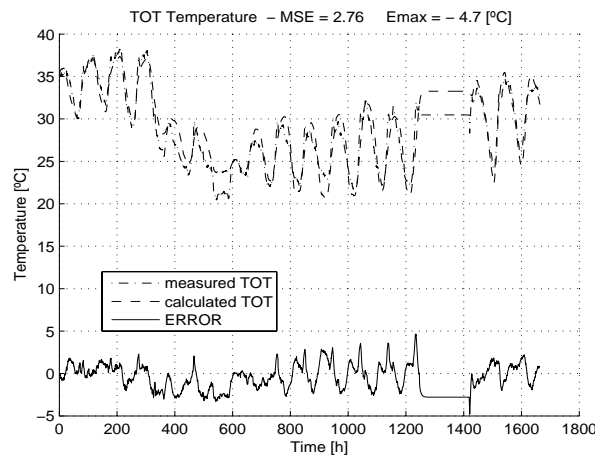


Fig. 5 Actual and estimated TOT with prediction error using RNN

### C. TOT Calculation using LS-SVM

The implementation of LS-SVM is performed by routines of the LS-SVMlab Toolbox version 1.5 [13]. In this toolbox is used an optimization algorithm for tuning the hyperparameters  $\sigma$  and  $\gamma$  of the model with respect to the given performance measure. Using the default values the optimization algorithm was shown efficient but, relatively slow. Then, the design of LS-SVM model of the transformer consists of the following steps:

- The experimental data were separate in two groups, first 40% samples will be used for model building/training and remaining 60% samples will be reserved for testing.

- The regularization parameter  $\gamma$  and the parameter  $\sigma$  specifying the width of the kernel are determined using 96 points (24 hours of operation of the transformer), reducing the computational time and avoiding the overfitting of the network. In the simulations was noticed, that a larger number of points in the determination of the hiperparameters results in overfitting of the network, and besides the optimization algorithm used is slow.

- The LS-SVM model is trained maintaining the hiperparameters  $\gamma$  and  $\sigma$ , determined previously. For training it was used 1152 points, corresponding to 288 hours of the transformer operation. The LS-SVM recurrent is trained as one feedforward network as follows:

$$\hat{y}_{k+1} = f_{ls-svm} \begin{pmatrix} x_k, x_{k-1}, x_{k-2}, \dots, x_{k-d}, y_k \\ y_{k-1}, y_{k-2}, \dots, y_{k-q} \end{pmatrix} \quad (8)$$

To calculate the p-step ahead prediction, it is used:

$$\hat{y}_{k+p} = f_{ls-svm} \begin{pmatrix} x_k, x_{k-1}, x_{k-2}, \dots, x_{k-d} \\ \hat{y}_{k+p-1}, \hat{y}_{k+p-2}, \dots, \hat{y}_{k+p-q} \end{pmatrix} \quad (9)$$

and gradually has to include more previous estimates for the output  $\hat{y}$ , until arrives at the p-th sample prediction  $\hat{y}_{k+p}$ . In fact the LS-SVM is used as a recurrent model to generate the prediction.

- The LS- SVM model can be retrained using the same data set, but with the new estimated outputs shifted through the input vector and old inputs are discarded. The retrained model is simulated using one validation algorithm until small testing error is reached.

The results of the performance of the LS-SVM are summarized in Table V.

LS-SVM	MSE	EMAX
Temporal (d = 1, q = 0)	6.06	8.9
Temporal (d = 2, q = 0)	4.89	-9.7
Recurrent (d = 1, q = 1)	0.07	-1.4
Recurrent (d = 2, q = 2)	0.06	-1.1
Recurrent (d = 3, q = 3)	0.10	1.5
Recurrent (d = 4, q = 4)	0.08	1.3
Recurrent (d = 5, q = 5)	0.63	3.7

It is observed that better testing performance is obtained with d = 2 and q = 2.

Fig. 6 shows the performed by the recurrent LS-SVM.

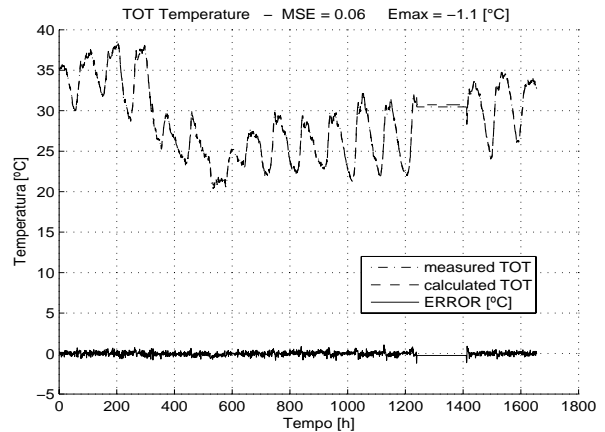


Fig. 6 Actual and estimated TOT with prediction error using recurrent LS-SVM

#### D. Comments on the Results

Better results of the TOT estimation for each implemented method is summarized in Table VI with tabulated the performance values of MSE and Emax.

METHOD	MSE	EMAX
G ANNEX	7.93	-13.0
SNN ( $n_h = 5$ )	5.24	6.1
TPNN (d = 3, $n_h = 3$ )	3.87	-7.8
RNN (d = q = 1, $n_h = 5$ )	2.76	-4.7
Recurrent LS-SVM	0.06	-1.1

It is important to remark that Recurrent LS- SVM outperforms the other four models considering MSE and Emax. It is also observed that the result achieved with the recurrent LS- SVM was done with one only training of the network. In the following, it will be shown in the Fig. 7 the calculated HST with the better result of the TOT estimated by recurrent LS- SVM.

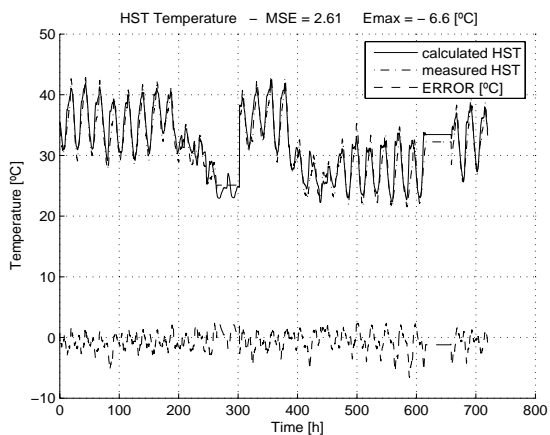


Fig. 7 Actual and estimated HST with prediction error using estimated TOT by recurrent LS-SVM

In the Table VII it is summarized the performance values (MSE and Emax) of HST, using the estimated TOT by recurrent LS-SVM and the calculated TOT by G Annex.

TABLE VII  
CALCULATION OF HST

Calculated TOT	HST (MSE)	HST (EMAX)
G Annex G	5.75	-14.7
Recurrent LS-SVM	2.61	-6.6

With the results shown in the Table VII, it is observed the increase of the MSE, when the calculated TOT is used for estimating HST. Therefore, a more precise estimate of the TOT results in a more precise estimate of the HST. Comparing errors obtained in the estimative of HST, using recurrent LS-SVM with the errors tabulated in table II (MSE = 2.37 and Emax = - 6.1 °C), it is verified that on-line monitoring of TOT can be substituted by Recurrent LS-SVM, in a straightforward manner and maintaining the same accuracy in estimating HST.

## VI. CONCLUSION

The IEEE model, ANN and recurrent LS-SVM are used to estimate TOT of power transformers. Of the five models, the recurrent LS-SVM provided the best performance in terms of the MSE and Emax. The superior results obtained with LS-SVM justify its application in the estimate of TOT. It is also recognized that LS-SVM holds a high generalization capability in relation to multilayer feedforward network such as multilayer perceptron trained com backpropagation or other more efficient variation of this algorithm. This is due to the fact that the LS-SVM network is more robust and efficient in identification of complex dynamic plants [7]. Since the LS-SVM training is equivalent to solving a set of linear equations, the solution of the LS-SVM is always unique and globally optimal [7]. For the implementation of the networks it was used the Neural Toolbox [12] and LS-SVMlab Toolbox version 1.5 [13], both of MATLAB, and it was verified that with the default parameters of the respective algorithms the implementation of the LS-VM model is easier than the ANN model. ANN involves more experience for modeling and training of the network, mainly for the definition of the number of hidden layers. Therefore, recurrent LS-SVM can be used as an important alternative to ANN and IEEE method in the estimate of the TOT. The results obtained in this work point out that Recurrent LS-SVM Model may be used to replace the on-line measuring of TOT, and to constitute a tool for on-line diagnosis of power transformers.

## ACKNOWLEDGMENT

The authors acknowledge Mr. José Luis Pereira Brittes of the Companhia Paulista de Força e Luz (CPFL) – Brazil, for the experimental data set used to test the methodology proposed in this paper.

## REFERENCES

- [1] J. A. Jardini, J. L. P. Brittes, L. C. Magrini, M. A. Bini, and J. Yasuoka, "Power transformer temperature evaluation for overloading conditions,"

- IEEE Transactions on Power Delivery*, vol. 20, no. 1, pp. 179–184, January 2005.
- [2] IEEE, *Guide for Loading Mineral-Oil-Immersed Transformers*, June 2002.
- [3] V. Galdi, L. Ippolito, A. Piccolo, and A. Vaccaro, "Neural diagnostic system for transformer thermal overload protection," *IEE Proceedings Electric Power Applications*, vol. 147, pp. 415–421, September 2000.
- [4] W. H. Tang, K. Spurgeon, Q. H. Wu, and Z. Richardson, "Modeling equivalent thermal dynamics of power using genetic algorithms," in *Proceedings of the IEEE*, 2002, pp. 1396–1400.
- [5] Q. He, J. Si, and D. J. Tylavsky, "Prediction of top-oil temperature for transformers using neural networks," *IEEE Transactions on Power Delivery*, vol. 15, pp. 1205–1211, October 2000.
- [6] K. Narendra and K. Parthasarathy, "Adaptative identification and control of dynamical systems using neural networks," in *Proceedings of the 28<sup>th</sup> IEEE Conference on Decision and Control*, 1990, pp. 1737–1738.
- [7] V. Vapnik, *The Nature of Statistical Learning Theory*, 1995.
- [8] J. A. Suykens and J. Vandewalle, "Multiclass least squares support vector machines," in *International Joint Conference on Neural Networks*, 1999.
- [9] T. V. Gestel, J. K. Suykens, D. Baestaens, A. Lambrechts, G. Lanckriet, B. Vandaele, B. D. Moor, and J. Vandewalle, "Financial time series prediction using least squares support vector machines within the evidence framework," *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 809–821, 2001.
- [10] D. Peterchuck and A. Pahwa, "Sensitivity of transformer's hottest-spot and equivalent aging to selected parameters," *IEEE Transactions on Power Delivery*, vol. 17, no. 4, pp. 996–1001, October 2002.
- [11] D. J. Tylavsky, Q. He, G. A. McCulla, and J. R. Hunt, "Sources of error in substation distribution transformer dynamic thermal modeling," *IEEE Transactions on Power Delivery*, vol. 15, no. 1, pp. 178–185, 2000.
- [12] H. Demuth, M. Beale, and M. Hagan, *Neural Network Toolbox User's Guide for Use with Matlab*.
- [13] K. Pelckmans, J. Suykens, T. V. Gestel, J. D. Brabanter, B. Hamers, B. Moor, and J. Vanderwalle, *LS-SVMlab Toolbox, Version 1.5*, Katholieke Universiteit Leuven, Department of Electrical Engineering - ESAT-SCDSISTA, February 2003.