VOL. 12, NO. 8, APRIL 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

AN INTEGRATION OF COMPUTATIONAL INTELLIGENCE TECHNIQUES IN ENGINEERING DATA PROCESSING FOR IMPROVING FORECAST ACCURACY USING CUBIC-SPLINE INTERPOLATION AND ANN MODEL

Saifullizam Puteh[1] and Rosnawati Buhari[2]
[1]Faculty of Technical and Vocational Education, Universiti Tun Hussein Onn Malaysia
[2]Faculty of Civil and Environmental Engineering, Universiti Tun Hussein Onn Malaysia
E-Mail: saifull@uthm.edu.my

ABSTRACT

In engineering operations and maintenance, the system failures and wrong decisions cause critical effects; and subsequently large economic losses. Therefore, designing a model that is able to predict the future trends of an engineering operation system has become an important issue. This paper investigation shows to solve prediction problems that have a limited amount of data. In order to demonstrate the efficiency of the proposed integration of computational intelligent techniques, Dissolved Gas Analysis (DGA) was chosen as a case study. DG Aisthest and ard technique used in power transformer condition monitoring and fault diagnosis. Many computational intelligences and statistical techniques have been proposed to develop a forecasting model to predict the future condition of a transformer in transmission system using DG Analysis. Dissolved gasses (e.g., hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4) and ethane (C_2H_6)) in transformer oil are used to detect the type of electrical faults. However, because the dissolved gas's data, collected from the oil-insulated transformer, require expensive laboratory analysis costs, only limited data has been received. The limited number of data points in time-series data collection is a cause of significant accuracy problems for analysis and prediction results. In this paper, integrations of computational intelligence techniques using cubic-spline interpolation (CSI) and artificial neural network (ANN) are proposed to improve training data set, in order to obtain better results from intelligent prediction models. The cubic-spline interpolation technique is applied to enhance the limited data of dissolved gases, by fitting smoothly to the limited data points and thus generate new and sufficient data. This generated data is used as training data to re-train a focused time delay neural network (FTDNN) model, to predict dissolved gases in oil-insulated transformers. Experiments have shown that even using only 10 data points (generated by the CSI technique) can lead to a significantly improved accuracy of forecasting dissolved results.

Keywords: computational intelligence, soft computing techniques, neural network, cubic spline interpolation and dissolved gas analysis.

INTRODUCTION

In asset management, such as a power transmission network, the oil-immersed power transformer is an essential part of balancing operational costs, performance and risk management [19]. In the continuous supply of electricity to users, power transformer operations can cause incipient faults under thermal and electrical stresses. This paper will focus on a literature study of Dissolved Gas Analysis (DGA) as the most important analysis for the incipient fault diagnosis of a power transformer. Various DGA methods, including Key Gas, Roger's Ratio, Doernenburg ratio, Logarithmic Monograph, IEC Ratio, and Duval Triangle are utilized [11]. Many organizations establish procedures and standard's criteria for DGA, such as IEEE Standard C57.113 -2010 [16] and IEC60599-2007 [2]. The study in [5] stated that failures in power transformers are proven to be significantly related to dissolved gas content in power transformer, such as hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4) and ethane(C_2H_6).

These dissolved gases become a critical task for the discovery of the transformer's condition status; because, if incipient failures of oil-immersed power transformers are found early, then transformer failures can be prevented [18]. A study in [2] stated that sufficient fault gas data, with reliable historical information, is required to generate a prediction model. However, DGA techniques are expensive, require specialized laboratory equipment, and tests are scheduled on a monthly or event yearly basis, thus causing limited data to be collected.

Due to the complexity of oil-immersed transformer faults [4], the integration of computational intelligence techniques is a scientific technique that increases the ability of a decision-making process and helps understandable system behaviour to be more in uncertain situations[8][13[14].The aim of our research is to integrate a CSI technique and an ANN-dissolved gas forecast model. The main focus is to investigate a CSI's technique to generate sufficient data to achieve better accuracy for a Focused Time Delay Neural Network (FTDNN) model to predict dissolved gases. This integration approach will enhance oil-immerse transformer fault diagnosis; with future forecast values for each dissolved gas.

The paper is divided in to six sections. Section 2 presents a summary of related works, regarding computational intelligence technique applications that are used to forecast dissolved gases in transformer oil. In Section 3, the forecasting based on ANN is presented. In Section 4 CSI technique for constructed and generated data is explained. Meanwhile, the proposed technique for dissolved gas forecasts using FTDNN is presented in Section 5. Details of our experiments are also presented in

Section 5. Finally, Section 6outlines several conclusions from this study.

RELATED WORK

Applying computational intelligence techniques to develop learning models that represent the condition of a power transformer have been studied by many researchers. The study reported in [18], explored the performance of autoregressive moving average (ARMA), artificial neural network (ANN), adaptive neuro-fuzzy inference systems (ANFIS), support vector machine (SVM), and genetic programming, for forecasting long-term observation monthly discharge time series. It is important to review the accuracy ability of all computational models used in DGA forecasting using a limited number of data points. The research reported in [8] used support vector machine (SVM) and compared its performance to the grey model (GM) for forecasting dissolved gas content in oil power transformers. The experimental results reported in [21] indicated that better forecasting accuracy could be achieved using applied particle swarm optimization, based on the SVM method.

A fundamental analysis of DGA to interpret fault types based on the historical information of dissolved gas data was reported in [16]. The stability reading performance of gas sensors was un sat is factory; because sometimes, it provides in accurate information [12]. An alternative solution to this problem, using the previous data forecasting system, can provide future data for the DGA process. Combining DGA and the developed intellective evaluating method is essential for the maintenance state of the power transformer [5].

The application of computational intelligent techniques in a transformer diagnosis system converts the conventional DGA technique into intelligent diagnosis methods to analyse the power transformer dissolved gas's data. ANN is commonly used in nonlinear time series forecasting [1][15] and could be extended to specific cases of DGA forecasting. Furthermore, [12] reported that many researchers use ANFIS in DGA data prediction. However, a study in [10] stated that ANN suffers from several weaknesses; namely, a requirement for a relatively large training dataset and over-fitting. Therefore, investigation of ANN and ANFIS is needed, because other forecasting exploration techniques are less popular due to a limited number of data points. Consequently, ANN and ANFIS still need support from other computational intelligent techniques and more investigation into forecasting exploration with a limited number of data points is required.

FORECASTING BASED ON ANN

Since the 1990s, ANN systems (which are based on the understanding of the brain and nervous systems) were gradually used in prediction systems [17]. ANN is most commonly used in nonlinear systems to model and simulate. One of the most useful and interesting factors of ANN is the forecasting time series [1]. It can be applied to forecasting dissolved gas content in oil power transformers, to convert historical data into a time series. In order to improve ANN performance, several integrated algorithms were proposed to predict the future by optimizing values. The best time series model was fitted to the dataset [21]. The objective was to find a suitable data structure to forecast future consumption with greater accuracy.

In theory, an inter-connection of neurons in ANN has many connections, known as multi-layer perceptron (MLP). Figure-1 shows a three layer feed forward model. In [1], were typical of MLP most popular in forecasting purpose. The previous logged observations are in the input nodes, while the output is used for forecasting purposes. The model is expressed using the following formula:

$$y_t = \alpha_0 + \sum_{j=1}^{n} \alpha_j f\left(\sum_{i+l}^{m} \beta_{ij} y_{t-i} + \beta_{0j} \right) + \varepsilon_t \qquad (1)$$

Where;

m - Number of input nodes

n - Number of hidden nodes

f - $f(x) = \dfrac{1}{1+e^{-x}} \{\alpha j, j = 0,1...,n\}$

β_{ij} - Weight from input to hidden nodes

$\alpha_0 and \beta_{0j}$ - Weight of arcs, value always 1

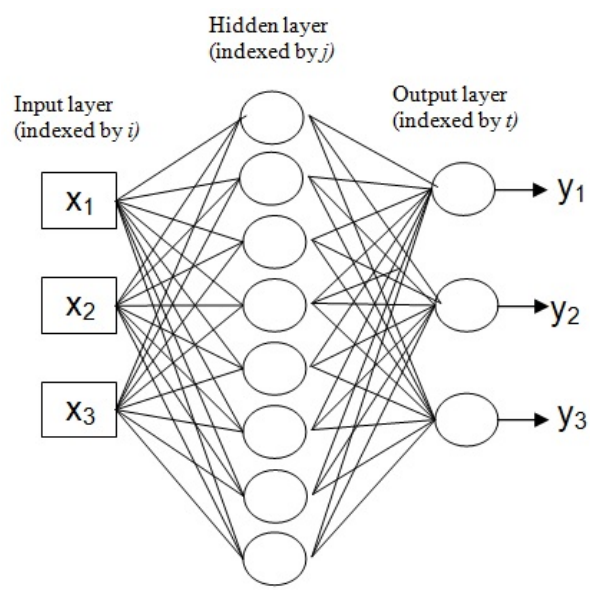

Figure-1. An MLP network.

To train the model using a large amount of data, many ANN systems (based on forecasters) have been proposed. Facing worldwide demand; especially from power transmission networks, management and businesses, the decisions target requires these methodologies to produce accurate prediction results using a limited training data set. Referring to [18], ANN suffers from several weaknesses, such as requiring a large amount

of training data, over-fitting, slow coverage velocity and relapsing in to local extremum easily. A study in [22], discussed the problem of over-fitting and several generalization learning techniques in the ANN literature. They introduced a Genetic Algorithm-based regulation (GARNET) for short-term load forecasting. Other studies have trained MLP networks and applied Earlying-Stopping, Bayesian Regulation, Adaptive-Regularization and GARNET using one year of data (150 input-output apirs of week day data were extracted from February to September) from a US utility. The findings from this experiment indicate that the model over-fits the training set; test errors increased while training errors decreased during the learning process. Therefore, improving the generalization performance of ANN within the forecaster system with a limited amount of training data becomes a critical problem.

This paper used a Focused Time Delay Neural Network (FTDNN). FTDNN is the specific structure of a neural network, which is used to forecast time series data [9]. Furthermore, FTDNN consists of a feed for ward network, with at aped delay line at the input (www.mathwork.com). Other popular structures of neural network model series data include Distributed Time Delay Neural Network (DTDNN) and Nonlinear Autoregressive Network with Exogenous Inputs (NARX). However, mostre as on that FTDNN does not require dynamic back propagation to compute network gradient. This is because that apped delay line appears only at the input of the network, and contains feedback loops or adjustable parameters. Another advantage is that this network trains faster than other dynamic networks (www.mathwork.com).

CUBIC-SPLINE CONSTRUCTION AND GENERATION

Data distribution of forecasting time series should measure starting from initial points of data to the end data point. Sometimes it is difficult to display data without a trend line and describe patter with limited data points. Data fitting by parametric curve can be used in wider fields, such as pattern recognition, image processing, statistical data analysis, reverse engineering and many other industrial applications [3].Because ANN, ANFIS and ARIMA suffer from over-fitting and under-fitting under a limited amount of data, looking for a reliable algorithm to reconstruct and generate adequate data for an insufficient number for training forecaster system, has become important. In [17] states that Cubic Spline Interpolant (CSI) can obtain a much smoother curve; because it is continuous up to the second-order derivative at the given knots. CSI requires less memory and is faster than high order polynomial. These advantages are suitable for CSI to be selected as are construction and generation algorithm for a forecast system that has problems with limited data points.

Cubic spline construction

The theory of classical univariate variation spline states that there is a relationship between the degree and smoothness of a spline. The greater approximation of a spline depends on a higher degree of spline. The experimental results shown in [6], indicate that for cubic and higher degree splines, a much better shape depends continuously on the position of the data. The integral to minimize principle for spline must be designed with weight that depends on the local interval length.

The CSI construction process, using adjacent data points with a third order polynomial, is shown in Figure-2. The cubic spline experiment has the following form:

$$f(x)=\begin{cases} p_1(x) & x^{[1]} \le x < x^{[2]} \\ p_2(x) & x^{[2]} \le x < x^{[3]} \\ \vdots & \vdots \\ p_{i-1} & x^{[i-1]} \le x < x^{i} \end{cases}$$

Where $\left(x^{[1]}, y^{[1]}\right), \left(x^{[2]}, y^{[2]}\right), \ldots, \left(x^{[i]}, y^{[i]}\right)$ and if $x^{[1]} < x^{[2]} < \ldots < x^{[i]}$, constructed a cubic spline by interpolating a cubic polynomial p_k between each pair of consecutive points $\left(x^{[k]}, y^{[k]}\right)$ and $\left(x^{[k+1]}, y^{[k+1]}\right)$ according to the following process:

Each polynomial passes through its respective end points:

$$p_k\left(x^{[k]}\right) = y^{[k]} \text{ and } p_k\left(x^{[k+1]}\right) = y^{[k+1]} \tag{2}$$

First derivatives match at interior points

$$\frac{d}{dx} p_k\left(x^{[k+1]}\right) = \frac{d}{dx} p_{k+1}\left(x^{[k+1]}\right) \tag{3}$$

Second derivatives match at interior points:

$$\frac{d}{dx^2} p_k\left(x^{[k+1]}\right) = \frac{d}{dx^2} p_{k+1}\left(x^{[k+1]}\right) \tag{4}$$

Second derivatives vanish at the end points:

$$\frac{d^2}{dx^2} p_1\left(x^{[1]}\right) = 0 \text{ and } \frac{d^2}{dx^2} p_{i+1}\left(x^{[i]}\right) = 0 \tag{5}$$

The above processes specify a system of line is equations that can be solved for the cubic spline, with a smooth transition between splines at each node.

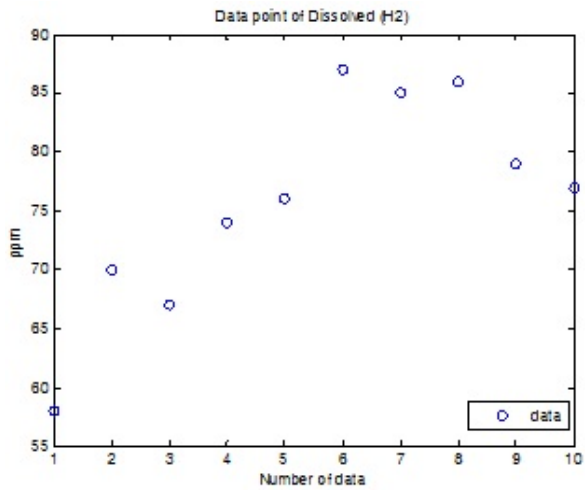

Figure-2. Data points of dissolved gas (H2).

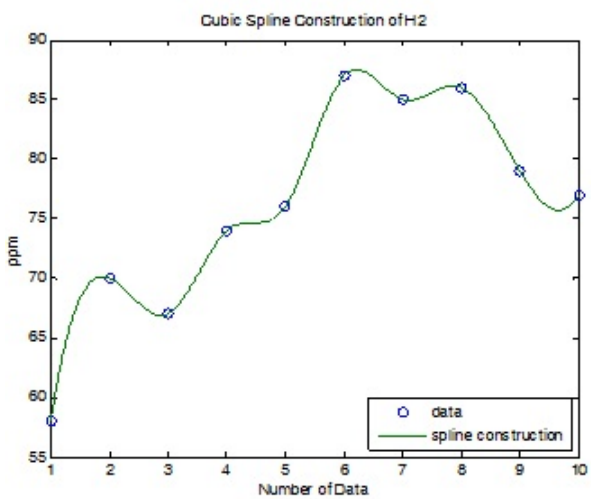

Figure-3. Cubic spline construction based on data points of dissolved gas (H2).

With the above condition, the equation system is completely defined with a smooth transition between splines at nodes [7]. The construction, using nine cubic polynomials for data points in Figure-2, is shown in Figure-3. Each segment, between all data points, is connected with third-order polynomials. Of the many mathematical and engineering computer software used to program cubic spline equations, MATLAB was chosen in this work; because most of the equations above are simple to program using MATLAB, and it is popular among researchers in engineering and technology areas.

CSI generator

In this paper, an experiment was designed for ANN and ANFIS training to examine forecast remodel performance. A forecaster system with a limited number of data points usually faces inaccurate results. In order to train ANN and ANFIS, a cubic spline was applied to constructed data and an algorithm generator was used to generated at a points based on the original data from Table-1 [16]. Table-2 shows the cubic-spline generator results for 10 data points of H2 from Table-1.

Table-1. The dissolved gas's content in power transformer oil.

Number of samples	H2	CH4	C2H2	C2H4	C2H6	Remark
1	58	130.6	5	300.3	39	Training data
2	70	139.5	5.1	311.5	39.1	Training data
3	67	133.4	5	307.9	39.7	Training data
4	74	142.2	5.3	324.9	41.8	Training data
5	76	140	5.2	317	40.8	Training data
6	87	138	5.2	317.2	40.7	Checking data
7	85	131.7	5.2	305.3	41.4	Checking data
8	86	133.8	5.3	300.8	38.6	Checking data
9	79	129.3	5.2	280.8	36.6	Checking data
10	77	138.4	5.4	308	38.9	Checking data

The completion code of the spline algorithm is presented in Figure-4. Through this technique, the application for using a smooth curve to interpolate data shows that a successfully generated limited number of data to the insufficient number of data for training model of forecasting. In the application using smooth curve to interpolate data, the results show that CSI successfully generated a limited number of data to the insufficient number of data for training model of forecasting.

1. Import data H2 from file of data
2. Define data H2 in variable Y
3. Define variable X as number of data
4. Increase value of X to any number and save into variable t
5. Performed spline function using command 'spline (X,Y,t)'
6. Plot graph (X,Y) and new spline graph.

Figure-4. Pseudocode of CSI generator

IMPLEMENTATION AND EXPERIMENTAL RESULTS

Build an FTDNN for prediction

FTDNN is the most straight forward dynamic network and is structured as a feed forward network with at apped delay line at the input. The generated data contained in Table-1 used CSI and formatted new data as inputs and targets of the network system for training. The procedure for forecasting dissolved gases, using nonlinear FTDNN, is described below:

Step 1: Initialize the FTD network system. Create the FTDNN network using a feed forward network

with additional input of the tapped delay line vector (the second input).

Step 2: Taped delay line, with delays from 1 to 8, and use five neurons in the hidden layer:

Step3: Arrange the network inputs and targets for training and load the tapped delay line with the eight initial values of the time series (contained in the variable p_i). The input of the network is the same as the target. This means the network has a one-step-ahead prediction.

Step 4: The training proceeds to set the training function.

Step 5: Validate with independent data.

Observation and investigation

In order to improve the performance of the predictive fault diagnostic system for power transmission networks, optimization (operating and learning from data) and design of the experiments to investigate the performance of learning model should be accomplished. The different machine-learning and algorithm shave their own advantages and disadvantages in different data sets. Therefore, the observation and investigation process of behaviours from the performance of forecast models has become a critical success factor to obtain a solution. This task to design experiments, applies different approaches to provide complementary information from different points of view and processes to modify improving the generalization performance in the forecaster system under limited numbers of training data. In this work, popular error measures of Mean Absolute Percentage Error (MAPE) were used and are defined as:

$$MAPE(\%) = \frac{1}{T}\sum_{t=1}^{T}\frac{|y_t - \hat{y}_t|}{y_t} X100 \qquad (6)$$

Where, y_t is the actual observation at time t, $\hat{y}_t$ is the predicted value, and T is the number of predictions.

RESULTS

From the experiment, using FTTDN with 10 data points, we found a poor percentage error of accuracy. Further experiments were implemented to investigate how much data was required for the learning model to improve forecasting accuracy. Cubic spline interpolation was applied to the generated 10 point set of the original data to the insufficient points of dissolved gas's data. This process could help to provide enough data points to derive a forecast system that would produce accurate predictions.

Table-2. Data output using cubic spline.

Original data point with 10 data	10 to 20	10 to 40	Infinity
58	58.0000	58.0000	limitation of data generation depend on computer memory
		64.0059	
	67.8525	67.7328	
		69.6218	
70	70.1051	70.1141	
		69.6510	
	68.5281	68.6736	
		67.6231	
67	66.8916	66.9408	
		67.0676	
	68.5949	68.2308	
		70.0428	
74	72.4919	72.0063	
		73.6237	
	74.5474	74.4482	
		74.6186	
76	74.7298	74.6519	
		75.0716	
	77.2577	76.3987	
		78.8092	
87	83.0695	81.7553	
		84.5980	
	87.2658	86.6984	
		87.5089	
85	86.7783	87.2098	
		86.3279	
	85.0345	85.3923	
		84.9264	
86	85.4174	85.1157	
		85.6212	
	86.1243	86.0593	
		86.0460	
79	86.1243	85.2723	
		83.8471	
	80.5395	82.0247	
		80.0597	
77	76.9752	78.2064	
		76.7193	
	77.0000	75.8529	
		77.0000	

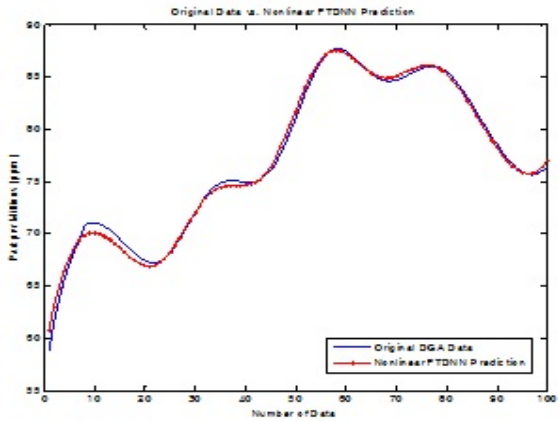

Figure- 5. FTDNN prediction with 100 (CSI generator) data points

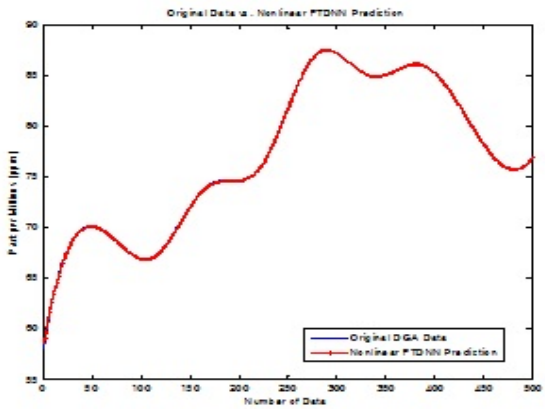

Figure-6. FTDNN prediction with 500 (CSI generator) data points

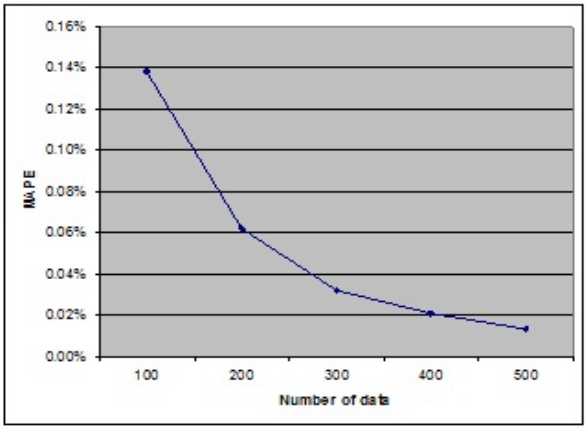

Figure-7.FTDNN error prediction

Table-3. Performance measure (MAPE) of FTDNN.

	Number of data points					
FTDNN Prediction error (%)	100	200	300	500	700	1000
	0.138	0.062	0.032	0.013	0.007	0.003

Figure-7 shows FTDNN performance of prediction for dissolved gases.

The summarized results (see Table-3) show that the prediction performances of FTDNN were very accurate when the number of data points increased. Figure-5 and Figure-6 show the different performances of FTDNN between 100 and 500 data points.

CONCLUSIONS

Sometimes we make the wrong decisions using limited information. This paper presented an integration between CSI technique and FTDNN to enhance the prediction of dissolved gases in power transformer oil. The CSI is used to reconstruct and generate data to an insufficient amount of data for training FTDNN for generating high accuracy prediction. With this capability, dissolved gases from transformer oil can be generate more data from an original limited *set to* a certain amount of data points. The experimental results showed that the proposed integration approach was able to predict only 10 data points of dissolved gases, with a prediction performance based on a MAPE of less than 0.5%.

ACKNOWLEDGEMENTS

The first author would like to acknowledge the support of the Office for Research, Innovation, Commercialization and Consultancy Management (ORICC), Universiti Tun Hussein Onn Malaysia (www.uthm.edu.my)

REFERENCES

[1] S.F. G. Azadeh and S. Sohrabkhani2007. Forecasting electrical consumption by integration of Neural Network, time series and ANOVA. Applied Mathematics and Computation, Volume 186. Issue 2. pp. 1753-1761.

[2] R. Chang. 2006. Applying nonlinear generalized autoregressive conditional heteroscedasticity to compensate ANFIS outputs tuned by adaptive support vector regression. Fuzzy Sets and Systems, Volume 157. Issue 13. pp. 1832-1850.

[3] G. Fenghua. 2009. Data Fitting by Quadratic Splines. International Conference on Measuring Technology and Mechatronics Automation.

[4] Y. C. Huang. 2003. Condition assessment of power transformers using genetic-based neural network. IEE Proc. Sci.Meas.Technol. pp. 19-24.

[5] J. Li, X. Chen and C. Wu. 2009. Power Transformer State Assessment Based on Grey Target Theory. Measuring Technology and Mechatronics Automation, ICMTMA '09. International Conference on Volume 2. pp. 664 – 667

[6] J. E. Lavery. 2006. Shape-preserving univariate cubic and high-degree L_1 splines with function-value-based and multistep minimization principles. Computer Aided Geometric Design 26. pp. 1-16.

[7] K. T. Klasson. 2008. Construction. Advances in Engineering Software, Volume 39. Issue 5. pp. 422-429.

[8] Langensiepen, A. Lotfi and S. Puteh.2014. Activities recognition and worker profiling in the intelligent office environment using a fuzzy finite state machine. Fuzzy Systems (FUZZ-IEEE), IEEE International Conference. pp.873-880

[9] M. A. Sailesh. 2006. (Lecture Note Style) Stok Trend prediction using recurrent neural networks. Dept. of E&E Com. Eng., Indian Institute of Technology. Kharagpur.

[10] A. Mohamed, A. Y. Abdelaziz and A. S. Mostafa.2005. A neural network-based scheme for diagnosis of power transformer. Elect Power Syst Res. 75(1): 29-39.

[11] B. Nemeth, S. Laboncz and I. Kiss. 2009. Condition monitoring of power transformers using DGA and Fuzzy logic; Electrical. pp. 373 - 376.

[12] O. Taylan and B. Karagözoğlu. 2009. An adaptive neuro-fuzzy model for prediction of student's academic performance. Computers & Industrial Engineering, Volume 57. Issue 3. pp. 732-741.

[13] C. Langensiepen, A. Lotfi and S. Puteh. 2012. Fuzzy ambient intelligence for intelligent office environments. IEEE International Conference on Fuzzy Systems.

[14] C. Langensiepen, A. Lotfi and S. Puteh 2013. Activities Recognition in Intelligent Office Environment, Proc. of Smart Offices and Other Workplaces workshop.

[15] P.P. Balestrassi, E. Popova, A. P. Paiva and J.W. M. Lima.2009.Design of experiments on neural network's training for nonlinear time series forecasting. Neurocomputing, Volume 72. Issues 4-6. pp. 1160-1178.

[16] S. W. Fei and Y. Sun. 2008. Forecasting dissolved gases content in power transformer oil based on support vector machine with genetic algorithm. Electric Power Systems Research, Volume 78. Issue 3. pp. 507-514.

[17] S. Q. Wang and H. S. Zhao. 2009. Cubic-spline reconstruction of irregular seismic data using linear time shift. Int. Conference on Wavelet Analysis and Pattern Recognition, Boading.

[18] S. W. Fei, M. J. Wang, Y. B. Miao, J. Tu and C. L. Liu. 2009. Particle Swarm Optimization Based Support Vector Machine for Forecasting Dissolved Gases Content in Power Transformer Oil. Energy Conversion and Management 50. pp. 1604 – 1609.

[19] W. H Tang, Lu and Q.H. Wu. 2008. A Bayesian Network Approach to Power System Asset Management for Transformer Dissolved Gas Analysis. DRPT 2008, Nanjing China.

[20] W. C. Wang, K. W. Chau, C. T. Cheng and L. Qiu. 2009.A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. Journal of Hydrology, Volume 374. Issues 3-4. pp. 294-306.

[21] M. H. Wang and C. P. Hung. 2003. Novel Gray Model for a prediction of trend of dissolved gases in oil-filled power apparatus. Elect Power System Res. 67(1):53-8.

[22] S.H. Zeke, H. W. Chan, A.B. Ngan, A.K. Rad and D.N. Kasabov. 2006. Short-term ANN load forecasting from limited data using generalization learning strategies. Neurocomputing, Volume 70. Issues 1-3.pp. 409-419.