Online Prediction of DGA Results for Intelligent Condition Monitoring of Power Transformers

G Nandagopan
Department of Electrical & Electronics
TKM College of Engineering
Kollam, Kerala, India,
nandagopang29@gmail.com

Sabeena Beevi K
Department of Electrical & Electronics
TKM College of Engineering
Kollam, Kerala, India,
ORCID: 0000-0002-5129-9386

Kunju Lekshmi A S
Department of Electrical & Electronics
TKM College of Engineering
Kollam, Kerala, India,
kunjupanoor@gmail.com

Vishnupriya K J
Department of Electrical & Electronics
TKM College of Engineering
Kollam, Kerala, India,
vishnupriyakj2017@gmail.com

Deepa S Kumar
Power Networks Demonstration Center
University of Strathclyde
Glasgow, Scotland
deepa.kumar@ieee.org

Abhijith S
Department of Electrical & Electronics
TKM College of Engineering
Kollam, Kerala, India,
abhijith2271718@gmail.com

Rishika K K
Department of Electrical & Electronics
TKM College of Engineering
Kollam, Kerala, India,
rishikakk5@gmail.com

Abstract—**Transformers form a major part of a power system in transmission as well as distribution of power. Considering the criticality, finance, and time involved in repair, periodic condition monitoring and maintenance of transformers is the key to ensure electrical safety as well as stable operation of the large interconnected power system. Dissolved Gas Analysis (DGA) is an established tool used to determine the incipient faults within the transformer by analyzing the concentration of different gases in the transformer oil and giving early warnings/diagnoses. Currently, transformers worldwide utilise online sensors to monitor dissolved gases and moisture content in oil. The online DGA sensor uses a small amount of oil from transformer to perform real-time DGA analysis and gives the ppm content of dissolved gases for further course of action. Considering the large quantity of assets and the huge amount of data produced, it is imperative to develop a tool to aid the operators in assimilating the available data for diagnosis and proactive decision making. The present study improvises AI techniques to predict future dissolved gas concentrations using real time DGA data collected from the transmission utility of the country. The prediction helps to forecast the trend of development of incipient faults in the transformer. The complete project scope is to develop a highly reliable diagnostic tool to emulate the decision-making ability of a human expert in transformer DGA analysis to enhance transformer life. In the present paper, models based on Auto-regressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), and Vector Auto Regression (VAR) are implemented to predict DGA data of three in-service transformers. DGA data is forecast for up to 8 monthly samples in the future, and the accuracy of results is compared with each other. The LSTM-VAR combined model is seen to provide the best results among them.**

Keywords— Transformer, fault diagnosis, Dissolved gas analysis (DGA), long short-term memory (LSTM), time series forecasting.

I. Introduction

Oil-filled transformer failures happen due to various reasons, including dielectric breakdown, internal/external short circuit faults, lightning and switching over-voltages, loose joints and hot spots, bushing and breather failure, etc. Utilities are reporting growing failures of the aging fleet of transformers and reactors, and rise in unwarranted and unexpected failures of even newly commissioned equipment. This increasing trend is observed in figure 1, which shows the number of transformer failures over a period between 2014 and 2018. According to the report of the standing committee of experts on the failure of 220 kV and above voltage class equipment published by the government of India, 21 failure cases are said to have occurred between September 2015 and December 2016 [1], whereas 78 equipment failures are reported between April 1, 2018, and March 31, 2019 [2].

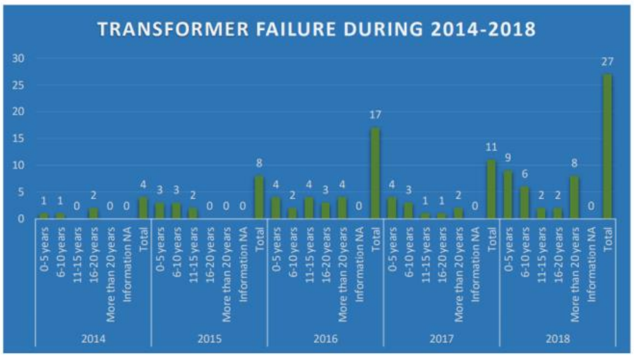

Figure 1 Transformer failure cases during the years 2014-2018

The Dissolved Gas Analysis (DGA) is a tool used for determining the faults in a transformer by analyzing the concentration of different gases dissolved in the transformer oil [3]. The conventional DGA methods such as the Key Gas method and Duval triangle methods interpret fault gas concentrations or gas ratios depending on the experience of practical experts rather than quantitative evidence. A smart real time monitoring tool capable of alerting the operator, through early signs of incipient faults or impending degradation in transformer, the most critical equipment in the transmission network, is required. Although conventional methods fail to provide an accurate analysis, especially at borders, the traditional approaches provide relevant input features and derived methodologies as a base for real time autonomous intelligent techniques. The proposed real time tool is built upon the established DGA analysis technology in combination with the multi-fold development of artificial intelligence for proactive condition monitoring and to limit the damage to the transformer as minimum as possible while avoiding sudden catastrophic failures.

Recent developments in statistical machine learning has brought in various machine learning models for DGA interpretation and subsequent fault diagnosis [4]. DGA interpretation using a deep neural network is proposed in [5]. The Duval triangle method is carried out with a simulated dataset to supply sufficient data. The deep neural network is shown to outperform standard classification algorithms like the k-nearest neighbor. [6] states how fuzzy logic models can enhance DGA interpretation.

The interpretation of DGA results for fault diagnosis employs methods that are subjective to the experts carrying out the observation. This creates some amount of risk for adverse events in the transformer. The availability of a reliable prediction of the DGA data for future instances can minimize this risk. The DGA data prediction is an application of time series forecasting. A model formed by combining artificial neural network (ANN) and autoregressive integrated moving average (ARIMA) is proposed in [7] to obtain time-series forecasts of data such as sunspot data and exchange rates. ARIMA gives forecast results as a linear function of the past values of the variable to be forecasted. ANN captures more information from the series and can predict non-linearities with good accuracy. The combined model has the advantages of both these approaches, and it is observed to be the better approach among the three. An application of LSTM model in classification is discussed in [8]. In [9], LSTM networks are used for time series forecasting, and they even give reduced error in the forecast when compared with ARIMA. It is stated that there is no significant impact of the number of epochs on the prediction accuracy in the case of LSTM.

From the existing studies regarding deep belief networks in DGA interpretation and fault detection, a deep recurrent belief network is proposed in [10] to carry out trend prediction of transformer DGA data, using time series theory. In [11], a multidimensional time series approach is followed to link the DGA values to the operating state of the transformer. Existing single time series forecasting methods like the backpropagation (BP) neural network and radial basis functions allow prediction by adjusting network weights and thresholds. A combined model of wavelet neural network (WNN), support vector machine (SVM) and graph neural network (GNN) are proposed to have an improved prediction result. [12] proposes long short term memory (LSTM) neural network for trend prediction of DGA data, which is said to have a higher accuracy of prediction when compared to the gray model, BP network model, and SVM model.

Although a number of models have been proposed already for predicting the type of fault that can occur in transformers using DGA data, there have not been many attempts at forecasting the gas concentration values in DGA data, which can be used to estimate the incipient fault types. This paper proposes different machine learning models for forecasting DGA data of transformers and predicting the trend of development of incipient faults of the transformer to avoid serious events and minimize the loss. This will enable the operators to apply DGA interpretation techniques to these predicted values, which would give the indication of any incipient fault in the system.

The DGA dataset for three in-service transformers was obtained from Power Grid Corporation of India Ltd (PGCIL). Models of Auto Regressive Integrated Moving Average (ARIMA), Vector Auto Regression (VAR), Long Short-Term Memory (LSTM) algorithms were implemented systematically, and the results of prediction are obtained with the lowest values of error.

The paper is organised in such a way that the first section is about DGA, followed by the description of the models used in the paper. The methodologies followed for each of the models, the results and the future scope of work in this area are discussed thereafter.

II. DISSOLVED GAS ANALYSIS

Faults occurring in transformers cause decomposition of oil/cellulose insulation. This produces several gases, like H2, CH4, C2H2, C2H4, C2H6, CO, and CO2, which get dissolved in the oil inside the transformer. The quantity of these gases is measured normally by gas chromatography. It was recognized early that specific gases, or gas ratios, could be associated with specific fault types. So analysing the concentrations of various gases can enable detection of the fault and the type of it that has occurred in the transformer. Dissolved gas analysis (DGA) is the commonly used tool for detecting incipient faults in power transformers by establishing a relation between the amount of certain gases dissolved in transformer oil and a corresponding malfunction. Some of the faults that can be identified by DGA are partial discharge, thermal faults, arcing, and fault involving insulation.

Gas	IEEE C.57.104	IEC 60599
Hydrogen	100	60-150
Methane	120	40-110
Ethane	65	50-90
Ethylene	50	60-280
Acetylene	1	3-50
Carbon monoxide	350	540-900
Carbon dioxide	2500	5100-13000

Table 1 Acceptable limits of fault gases in ppm

The table 1 shows the acceptable limits of gases in parts per million (ppm) in transformers according to two standard specifications.

The conventional DGA methods like Key Gas, Duval Triangle, Duval Pentagon, etc. are widely used to detect the faults in the transformers. However, these methods are heavily dependent on expert opinions for fault diagnosis. This has led to many Artificial Intelligence (AI)-based methods being developed in order to overcome the drawbacks and improve fault detection based on ANN, SVM, fuzzy systems, and expert systems, among others. But there has not been enough work done using machine learning models for forecasting gas concentration values. This paper attempts to overcome this deficiency by proposing such a model after implementing and evaluating different approaches.

III. MODEL DESCRIPTION

A. ARIMA

ARIMA is the abbreviation for Auto-regressive Integrated Moving Average. It is a time series prediction algorithm based on the concept that information contained in historical values of time series can be utilised for predicting the future**.** This class of models shows a given time series according to its past values, i.e. its own lags and the lagged forecast errors. The ARIMA model is characterised by three terms ‘p’, ‘d’ and ‘q’.The ‘p’ represents the order of the Auto-regressive term, ‘d’ represents the order of the differencing term and ‘q’ denotes the order of the moving average term, where these parameters are substituted with integer values to indicate the specific ARIMA model to be used quickly.

The Autoregression term ‘AR’ uses the dependence of any given observation on a specific number of lagged observations i.e. it predicts future values based on linear combination of past values based on equation (1).

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \varepsilon_1 \qquad (1)$$

In above eq. (4), Y_t is a function of the lags of forecasted values, α is a constant value or the intercept term, β_1 is the coefficient of first lag value, Y_{t-1} is the first lag value of the time series data and ε_1 is the error term. The Integrated term ‘I’ denotes the difference between the current observation and the preceding time step to make the time series stationary.

The Moving Average ‘MA’ term uses the dependency of an observation on a residual error of a moving average model applied to lagged observations based on equation 2.

$$Y_t = \alpha + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \cdots + \phi_p \varepsilon_{t-q} \qquad (2)$$

The overall mathematical equation of Arima model is represented by equation 3.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \phi_1 \varepsilon_{t-1} + \varepsilon_{t-2} + \cdots + \phi_p \varepsilon_{t-q} \qquad (3)$$

B. VECTOR AUTO-REGRESSION

Vector auto-regression (VAR) is a model used to utilize the relationship between multiple quantities over time to predict future values of the quantities. VAR models enhance the univariate autoregressive model, such as the ARIMA model, by allowing for multivariate time series analysis. Each variable in the VAR model is modeled as a linear combination of its own previous values and those of other variables in the system. It is used when there are multiple time series that influence each other. These series are modeled into a system of equations with one equation for each time series.

C. RECURRENT NEURAL NETWORK

A recurrent neural network (RNN) is a class of artificial neural networks in which a network is constructed along a temporal sequence using connections between nodes. RNN has a memory, for storing all information about what has been calculated. RNNs use their memory to process input sequences of variable lengths. Long short-term memory is a popular RNN algorithm that can be used for time series forecasting.

D. LSTM

The Long Short-Term Memory network, or LSTM network, is a RNN trained using backpropagation through time, overcoming the vanishing gradient problem. It can be used to create large recurrent networks which can be used to address complex sequential problems in machine learning and obtain accurate results. LSTM networks consists of memory blocks ,instead of neurons, that are connected through the layers.

Figure 2 represents the architecture of the LSTM neural network. It is composed of several memory blocks, or cells, as represented by the rectangles in the figure. The information that is passed from one cell to the next is two states, called cell state (x) and hidden state (h). the memory blocks remember or forget information using three mechanisms, called gates.

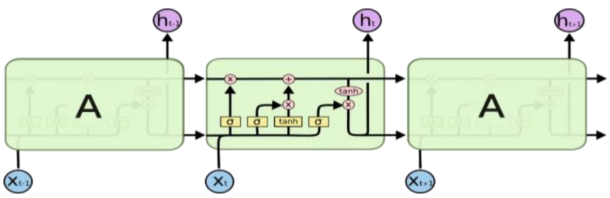

Figure 2 LSTM Architecture

IV. METHODOLOGY

The DGA data of 3 in-service power transformers recorded over a period of 9 years, from 31-05-2011 to 20-05-2020 is available for the analysis. The dataset includes several parameters, from where only the relevant features, which are values of gas concentrations, winding temperature, and load were extracted for this purpose. The dataset was cleaned to filter out duplicate rows. The number of samples in each transformer dataset was limited in the range of 35 and the data points were at unevenly spaced with time. A resampling process is applied to tackle the class imbalance.

Up-sampling is a procedure where synthetically generated data points are injected into the dataset by which the frequency of data points available in a dataset is increased. It is carried out by interpolating values in between the existing data points to get monthly data points. Linear interpolation was done to preserve the characteristics of the original dataset, as a non-linear interpolation like polynomial interpolation would bring in new crests and troughs in the existing dataset values. Through up-sampling, the dataset was made to consist of 109 monthly samples of time series data. The various machine learning models used in the project are discussed hereafter.

A. ARIMA Model

The ARIMA model gives univariate prediction utilizing the past values of C_2H_2, C_2H_4, C_2H_6, CH_4, H_2, CO_2 and CO gases. ARIMA (4,2,0) was used for this purpose. The data was split into train and test data in a ratio of 80:20. From the results, we concluded that the proposed ARIMA model provides a good accuracy in predicting the immediate future values of gases. The ARIMA model is a univariate time series model; we cannot solely depend on this model because gases such as CO and CO_2 evolved during incipient fault depends on other parameters such as load, winding temperature etc. Also, the prediction often tends to saturate into a linear regression. Thus, a multivariate time series model could be a more reliable approach.

B. LSTM Model

For applying the LSTM model, the data was transformed into one that consists of lag values of the features as inputs, and the gas concentration in the next time step as output. Before training the LSTM model we normalized the datasets using *MinMax* Scaler and converted the datasets to a supervised learning dataset. The *MinMax* scaler is a feature transformation tool that converts all values in the dataset to those within a specified range, such as between 0 and 1 in this project. Train and test data were assigned in a 70:30 ratio. We designed a single layer LSTM model in which the network contains a visible layer of 50 LSTM units. A single unit dense layer follows this. *RMSprop* optimizer was used. The training was for 50 epochs, and a batch size of 1 is used. Finally, we generated the predictions using this LSTM model for the test input data to visualize the model's performance.

The proposed LSTM model can be used to obtain single-step prediction only because it is a supervised learning approach. Two approaches were used to obtain predictions for multiple time steps - an encoder-decoder model and a combined LSTM-VAR model.

C. LSTM Encoder-Decoder Model

The Encoder-Decoder LSTM is a recurrent neural network that is developed for sequence-to-sequence problem cases. This architecture comprises of two models. One model is for reading the input sequence and encoding it to a vector of a fixed length. The second one decodes the vector of fixed length and outputs the predicted sequence. This model is named Encoder-Decoder LSTM because of this special architecture mentioned above, and the model is specifically designed for sequence-to-sequence problems. This model consists of two parts: encoder and decoder. Firstly, the input sequence is made available to the network as a single encoded character at a time. Here we have used two LSTM layers having 25 neurons to implement the model. The output of the model is a fixed-size vector that carries the internal information of the input sequence. The decoder needs to transform this internal representation into the correct output sequence. These LSTM layers can also be used to construct the decoder model. This model takes information from the fixed sized output of the encoder model. A 100-unit dense layer followed by a single unit dense layer facilitates the output for the network. These weights can also be used to output each time step of the output sequence with the help of a Time Distributed wrapper. A repeat vector layer is used for fitting the encoder and decoder models together. The last 8 DGA samples were predicted using this model.

D. LSTM-VAR Model

For each time step in the future, the inputs are chosen as the previous value of gas concentration obtained from the LSTM model, and winding temperature and load values of the previous time instance obtained from the VAR model of lag order 5, as shown in figure 3. This is continued until the prediction of the required number of time steps. Thus, winding temperature at time instance 't' is taken as a linear function of values of winding temperature and load at previous time instances.

The model is trained using the first 100 samples and is used to predict the last 8 samples in the test set. The results are validated with actual values and error is calculated.

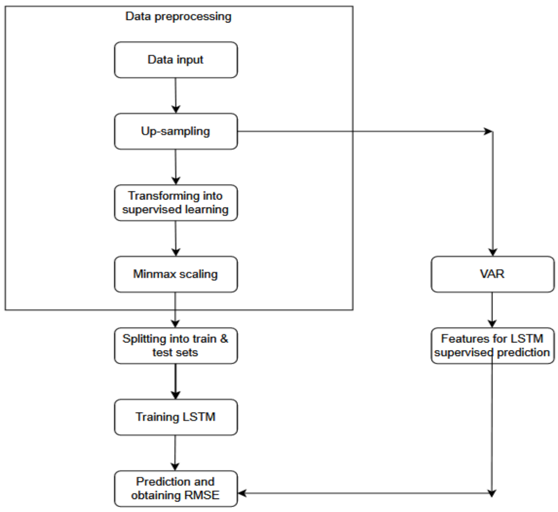

Figure 3 LSTM-VAR model

V. RESULTS AND DISCUSSIONS

The results of prediction by the different methods used in the project are analysed by comparing with the actual values of DGA for each of the transformers. Root mean square error (RMSE) is the commonly sed error metric in regression problems including time series forecasting. Here, the RMSE values of the models are compared to evaluate the performance of the models.

Root mean square error is calculated by:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y_i} - y_i)^2}{n}} \qquad (7)$$

where $\hat{y_i}$ represents the predicted value, y_i is the actual value, and n is the number of values.

A. ARIMA Model

The ARIMA model was found to have lower magnitude of RMSE when it was used to predict the test data. The RMSEcorresponding to different cases of predictions are shown in table 2. While these error values are quite low, this cannot be taken as a conclusive evidence to justify the performance of the model. There are two reasons for this. One, the gas concentrations in the dataset appears to be having a straight-line behavior towards the recent time instances, as observed in figure 4. Two, the ARIMA model is entirely dependent on the previous values of the gas concentration, and the prediction curve tends to saturate into a straight line, if no significant patterns are observed in the previous samples. Both these factors indicate that this model is unreliable for prediction of future trends.

H_2	CH_4	C_2H_2	C_2H_4	C_2H_6	CO	CO_2
0.069	0.13	0	0.054	0.046	21.035	24.347

Table 2 RMSE (in ppm) of ARIMA Model

This makes it necessary to implement a multivariate approach for prediction, and LSTM is used for this purpose.

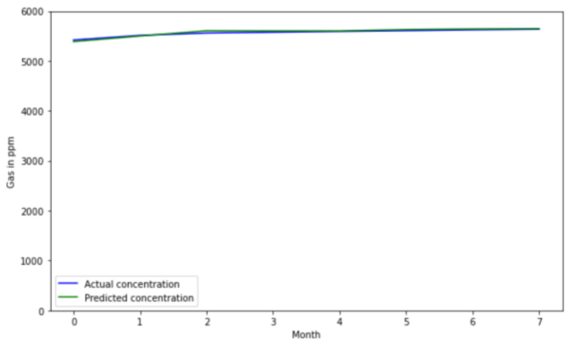

Figure 4 Predicted and actual values of CO2 using ARIMA model

B. LSTM Model

LSTM model gives single-step forecasts with a low value of error. The RMSE values obtained are given in table 3. Only one value is obtained at a time, and it requires all the features of prediction up to the previous value to obtain a prediction. Thus, the use of this method is very limited.

H_2	CH_4	C_2H_4	C_2H_2	C_2H_6	CO	CO_2
1.288	0.847	0.213	0.001	0.046	111.91	50.64

Table 3 RMSE (in ppm) of LSTM Model

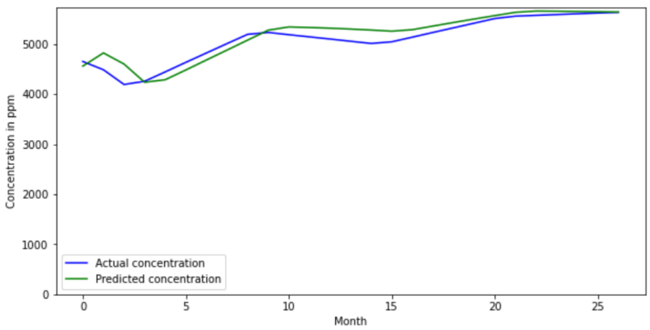

Figure 5 Predicted and actual values of CO2 using LSTM model

C. LSTM multiple-step prediction models

To predict multiple steps of gas concentrations, the encoder-decoder model of LSTM and LSTM-VAR models were implemented. The LSTM-VAR combined model uses VAR to predict features, which are incorporated to LSTM in order to predict gas concentrations. The predictions obtained by both these multiple-step prediction models are compared in table 4, by means of the average RMSE of forecast in all three transformer datasets. A comparison of forecasts by both of these models of concentration of CO_2 is shown in figure 4.

GAS	LSTM-VAR	ENCODER-DECODER
H_2	1.62	2.10
CH4	1.24	3.44
C_2H_4	0.04	0.04
C_2H_2	1.38	3.705
C_2H_6	0.36	1.96
CO	173.17	133.69
CO_2	256.78	266.10

Table 4 Average RMSE (in ppm) of Multiple-step prediction models

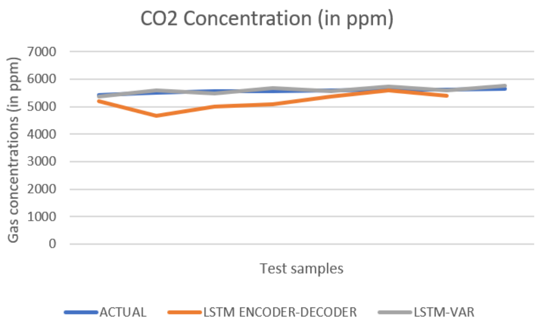

Figure 6 Comparison of CO2 concentration prediction

The RMSE for the encoder-decoder model is seen to be higher than that of the LSTM-VAR model for a majority of the gas concentrations. This is because the encoder-decoder model relies on periodic patterns present in a dataset to give predictions. It gives lesser importance to the magnitude of values of the features used. Since the DGA dataset is devoid of such periodically repeating patterns, it does not give a good prediction.

The real-life DGA values corresponding to the sample date 30/11/2020, 6 months from the previously available dataset, were obtained from PGCIL. The proposed multiple step prediction models were used to predict this sample and the results are shown in table 5.

GAS	ACTUAL VALUE (ppm)	ARIMA (ppm)	LSTM ENCODER-DECODER(ppm)	LSTM-VAR (ppm)
CH_4	30	34.05	27.23	25.95
C_2H_4	5	6.05	4.92	5.43
C_2H_6	8	8	6.35	7.98
CO_2	5710	5769	5375	5698

Table 5 Comparison with actual data for 30-11-2020

The ARIMA model predicts the given sample most accurately, followed by the LSTM-VAR model, as seen from the tables above. But since ARIMA is a univariate forecast algorithm, this low error value cannot be inferred to observe it as a reliable model, but it is due to the linear pattern of the dataset itself. LSTM-VAR is already observed to be better at predicting non-linear variations.

VI. CONCLUSION

DGA is a commonly used tool for determining any faults in transformers from the values of dissolved gas concentrations in transformer oil. The paper proposes a time series forecasting approach for forecasting gas concentration values in DGA datasets of 3 in-service transformers, which can then be used to determine any incipient faults occurring in a transformer. Four types of machine learning models were used in a phased manner, and the performance of these was compared using root mean squared error (RMSE) as the accuracy metric. The Auto-Regressive Integrated Moving Average (ARIMA) model gives a univariate prediction with low RMSE, which is unreliable and often saturates to linear regression. A reliable forecasting model should incorporate multiple variables which determine the operating state of the transformer. A Long Short Term Memory (LSTM) model was proposed for multivariate forecasting, but it can only be used for a one-step prediction. The encoder-decoder model and LSTM combined with Vector Auto-Regression (VAR) were the two alternatives used in LSTM based multiple-step prediction. The former showed a dependency on repeating patterns, and thus LSTM-VAR proves to be a better approach. The accuracy of forecasting can improve when the available training data increases. The LSTM-VAR model is proposed for forecasting DGA concentrations for 6 months in the future, thereby helping to estimate any fault occurring in the transformer in advance, without depending on the data from the DGA sensors for the corresponding months.

VII. FUTURE SCOPE

DGA data for faulted transformers have been acquired from PGCIL. The proposed models are being implemented for these data as well to evaluate the performance. The utility of advanced forms of RNN such as bidirectional LSTM, and other approaches for multiple step prediction using LSTM can be explored. A model for reducing the error in prediction can be developed by giving the previous error in each step as feedback to the next prediction. Also, the determination of the fault type from the forecasted values of gas concentrations can be performed by building a machine learning model for classification, with the DGA parameters as inputs.

ACKNOWLEDGMENT

We acknowledge the guidance and support rendered by Power Grid Corporation of India Ltd and express sincere thanks for providing the required data for this work.

REFERENCES

[1] *Report of Standing Committee of Experts on Failure of 220 Kv & Above Voltage Class Substation Equipment (September 2015 - December 2016),* Central Electricity Authority, Ministry of Power, Government of India

[2] *Report of Standing Committee of Experts on Failure of 220 Kv & Above Voltage Class Substation Equipment (April 2018 - March 2019),* Central Electricity Authority, Ministry of Power, Government of India

[3] J. Golarz *Understanding Dissolved Gas Analysis (DGA) Techniques and Interpretations*, 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D), 2016, pp. 1-5

[4] P. Mirowski and Y. LeCun, *Statistical Machine Learning and Dissolved Gas Analysis: A Review*, IEEE Transactions on Power Delivery, vol. 27, no. 4, pp. 1791-1799, Oct. 2012

[5] H. MehdipourPicha, R. Bo, H. Chen, M. M. Rana, J. Huang and F. Hu, *Transformer Fault Diagnosis Using Deep Neural Network*, 2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), 2019, pp. 4241-4245

[6] S. Hmood, A. Abu-Siada, M. A. S. Masoum and S. M. Islam, *Standardization of DGA interpretation techniques using fuzzy logic approach*, 2012 IEEE International Conference on Condition Monitoring and Diagnosis, 2012, pp. 929-932

[7] G.Peter Zhang, Time series forecasting using a hybrid ARIMA and neural network model, Neurocomputing, Volume 50, 2003, Pages 159-175, ISSN 0925-2312

[8] Aparnna A., Sabeena Beevi, Serene Benson, Ahamad Dilshad, and Deepa S. Kumar, *A Modified CNN for Detection of Faults During Power Swing in Transmission Lines*, 2020 International Conference on Power, Instrumentation, Control and Computing (PICC), pp. 1-5, IEEE, 2020

[9] S. Siami-Namini, N. Tavakoli and A. Siami Namin, *A Comparison of ARIMA and LSTM in Forecasting Time Series*, 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 2018, pp. 1394-1401

[10] B. Qi, Y. Wang, P. Zhang, C. Li and H. Wang, *A Novel Deep Recurrent Belief Network Model for Trend Prediction of Transformer DGA Data*, IEEE Access, vol. 7, pp. 80069-80078, 2019

[11] S. Zhang, Y. Bai, G. Wu and Q. Yao, *The forecasting model for time series of transformer DGA data basedon WNN-GNN-SVM combined algorithm*, 2017 1st International Conference on Electrical Materials and Power Equipment (ICEMPE), 2017, pp. 292-295

[12] J. Dai, H. Song, G. Sheng and X. Jiang, *LSTM networks for the trend prediction of gases dissolved in power transformer insulation oil*, 2018 12th International Conference on the Properties and Applications of Dielectric Materials (ICPADM), 2018, pp. 666-669