CIGRE-160 **2019 CIGRE Canada Conference**
Montréal, Québec, September 16-19, 2019

Transformer cooling system monitoring using neural networks

A. ZINFLOU, P. PICHER, L. VOULIGNY and A. MARCHAND
Hydro-Québec
Canada

EXECUTIVE SUMMARY

Dynamic thermal models can be used to monitor transformer cooling system condition by comparing estimated oil temperature with the top-oil measurement commonly available for all units. When cooling system performance deteriorates, the measured top-oil temperature exceeds the one estimated by the model. Several physical models for transformer dynamic thermal modelling have been proposed in the literature, with various levels of accuracy depending on their complexity.

A neural network based model has been applied to represent the dynamic thermal behavior of transformers in service on Hydro-Québec's transmission system. To cover Canadian winter and summer conditions, at least one year of measurements from transformers in service have been used to train the model. When using physical models, it is important to consider the various stages of cooling, e.g., the number of fans in operation in air-forced cooling mode. Neural networks trained from historical measurements can inherently consider this aspect.

This paper presents the performance and limitations of this approach. The models were trained using measurements from 54 units, with natural and forced oil flow in the windings and the cooling system and with nominal ratings ranging from 120 kV to 315 kV and 22.5 MVA to 450 MVA. To demonstrate the capability of the selected neural networks to detect a faulty condition, the cooling stage of a transformer was manually operated to create a temperature change that would be captured by the monitoring system. The next stage of development will include on-line training of the networks to be initiated automatically as soon as enough data is available.

KEYWORDS

Transformer Monitoring, Machine Learning, Artificial Neural Networks

zinflou.arnaud@ireq.ca

1. INTRODUCTION

Hydro-Québec's power transformer fleet is composed of more than 2 300 transformers (generator step-up, transmission and transformers feeding the distribution network) with a total installed capacity upwards of 200 000 MVA. An integrated transformer monitoring strategy has been implemented, with sensors installed on transformers to monitor selected parameters that can provide useful condition assessment information. Transforming measurement data into meaningful condition assessment information requires the use of data analytics that are based on the knowledge of the physical failure modes of the main components (active part, cooling system, bushings and tap changer). Every element of the strategy, from the sensors to the corrective maintenance actions (shown in Figure 1), are essential to capture the benefits of the strategy. At the time of preparing this paper, 293 transformers (with at least one connected sensor) are available via the centralized monitoring system presented in Figure 1. This number will grow in the coming years when more substations will be connected to the data acquisition system. This paper focuses on analysis of the data to assess cooling system performance.

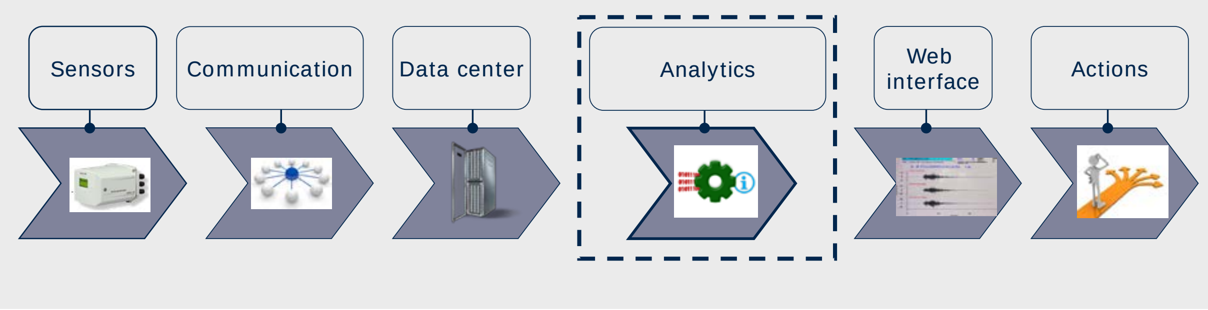


Figure 1: Integrated strategy for transformer monitoring

Temperature monitoring devices are used to control the cooling system and to trigger high temperature alarms. The use of modern digital devices allows the transmission of temperature data to a centralized location where measurements from other sensors are also available. Ambient temperature, oil temperature, winding temperature and the load are correlated. Recent researches using 3D numerical simulations and imbedded fiber optics in transformer windings indicate that the fluid flow occurring in the winding is an extremely complex phenomenon to describe by simple equations.

This paper presents a machine learning (ML) approach that can be used to monitor cooling system performance. This approach is compared with standard loading guide equations used to calculate the top-oil temperature as a function of ambient temperature and load. A deviation from normal operation has been created by manually operating the cooling stages to demonstrate the capacity of the model to detect abnormal behavior.

2. LOADING GUIDE DYNAMIC THERMAL MODELS AND THEIR LIMITATIONS

A detailed review of state-of-the-art dynamic thermal models can be found in [1]. The applicability of IEC and IEEE loading guides [2, 3] for power transformers is limited to ambient temperature above 0°C because the thermal model does not account for variations in the oil viscosity and winding resistance [4]. In other words, in the loading guide models, the ultimate temperature rise of oil and winding is independent of the ambient temperature. Recent developments in transformer dynamic thermal modelling concentrate on including these two parameters and on better describing the thermal overshoot phenomenon associated with the local convective heat transfer, which depends, among other factors, on the distribution of oil flow in the cooling channels. The direct modelling of thermal overshoot can help reproduce more adequately the transient temperature obtained after a step-load increase that can happen for instance when a parallel transformer is removed from service [5]. Authors in [6, 7] have developed metrics and compared various models with data measured on two transformers, one with forced oil circulation and the other with natural oil convection. They conclude that the most promising models should be further investigated using larger data sets. Authors in [8, 9] created an improved model that was tested and validated on one transformer with forced oil circulation. The model was tested in summer and winter conditions with good accuracy even when keeping the same ultimate oil temperature rise.

A three-phase transformer (in La Suète substation located in Québec City) rated 66 MVA, 225/26.4 kV with natural oil convection and air flow through radiators (ONAF) was continuously monitored to improve our knowledge of thermal behavior. Measurements of ambient temperature, load and oil temperature (top and bottom), as well as inside the windings using fiber optic probes, were taken every minute for more than three years. Table 1 summarizes the data when the transformer was operating at rated load, with all cooling fans in operation, at ambient temperatures varying from -29 °C in January to +23 °C in July.

When the bottom oil temperature falls below 0°C, the temperature difference between top and bottom oil increases significantly, thus indicating a lower oil flow rate in the windings that can be explained by an exponential increase of the oil viscosity [10]. The formulation in [10] includes the oil viscosity, but it is based on the top-oil temperature that is much higher than 0°C even at -29°C ambient, thus creating a systematic error of more than 20°C. This model could possibly be improved by using the bottom oil temperature as a reference for the oil viscosity calculation. However, the bottom oil temperature is usually not measured in normal operation, and the heat run test data does not always integrate this information for older transformers.

In summary, research in this field is ongoing. A new CIGRE Working Group (WG A2.60) was created in 2019 to collect existing experience and make recommendations for further improvements.

Table 1: Comparison of measurement and calculation for rated load operating condition of a 66 MVA transformer with natural oil convection

	Measurements from transformer in service at rated load				Loading guide models
Month	θ_{amb} **(°C)**	θ_{to} **(°C)**	θ_{bo} **(°C)***	$\theta_{to} - \theta_{amb}$ **(°C)**	$\Delta\theta_{to,rated}$ **(°C)**
January	-29.0	45.2	-15.5	**74.2**	**50.2**
March	-3.7	52.1	13.3	**55.8**	
April	9.7	65.0	29.8	**55.3**	
September	15.6	70.3	35.5	**54.7**	
July	23.0	74.1	41.4	**51.1**	

θ_{amb} is the ambient temperature
θ_{to} is the top-oil temperature
θ_{bo} is the bottom oil temperature
$\Delta\theta_{to,rated}$ is the top-oil temperature rise over ambient at rated load

3. THERMAL MODELLING THROUGH MACHINE LEARNING APPROACHES

3.1 Literature overview

Dynamic thermal models limitations and the increased interest in machine learning (ML) approaches lead many authors to propose fit-for-purpose thermal prediction models, such as artificial neural networks (ANN) [11-14], thermal models with error correction via ANN [15], thermal models with parameter optimization via genetic algorithms (GA) and particle swarm optimization (PSO) [16], adaptive network-based fuzzy inference systems (ANFIS) [17], evolving Gaussian fuzzy systems [18], neuro-fuzzy systems [19], genetic programming [20] and ensembles of quantile regression models [21].

In our study, different models have been designed and tested, including random forests (RF) [22], feedforward multilayer perceptron (MLP) [23], support vector regression (SVR) [24], ANFIS, and non-linear autoregressive exogenous neural networks (NARX) [25]. These approaches are well known in the literature and have their own strengths and weaknesses. Our experimental results show that the NARX method generates the most accurate predictions to solve this problem. The NARX model is explained in further detail in Section 3.2. The following describes briefly the other tested approaches.

- Schematically, RF are an ensemble of recursive trees. Each tree is generated from a bootstrapped sample and a random subset of descriptors is used at the branching of each node in the tree. The approach creates many trees by repeatedly resampling training data and averaging differences through voting [22].

- An MLP is a three-layer neural network (input, hidden, output) composed of fully connected neurons. Each neuron performs a weighted sum of its inputs and passes the results through an activation function. Simple neural networks like MLP can be easily adapted to process time series through an input tapped-delay line, giving rise to the well-known time delay neural network (TDNN) [26].

- The SVR algorithm basically maps input data into an m-dimensional feature space using a kernel function. The kernel translates a nonlinearly separable problem into a feature space, which is linearly separable by a hyperplane. The SVR defines a $\mathcal{E}$ loss function that ignores the errors situated within a certain distance of the true value.

- ANFIS are a class of adaptive networks that incorporate both neural networks and fuzzy logic principles. This approach is essentially a rule-based fuzzy logic model whose rules are developed during the training process of the model. The training process is data-based, in other words, ANFIS constructs a fuzzy inference system whose membership function parameters are derived from the training examples.

3.2 Description of the selected approach

A NARX is a type of recurrent neural network in which the output of a given time step depends on both exogenous inputs and outputs of the preceding time steps. The dynamic behavior of a NARX can be expressed as:

$$y(t) = f(y(t-1), \dots, y(t-n_y), x(t), x(t-1), \dots, x(t-n_x)) \quad (1)$$

where $x(t)$ and $y(t)$ are the input and the output at time step t, n_x and n_y are user-defined delays for the input and the output, and $f(.)$ is a non-linear function. $f(.)$ is generally modelled as a feedforward neural network. NARX are formalized in [25] and used for prediction in a wide range of domains on systems with non-linear dynamic behavior, such as heat exchangers, wastewater treatment plants, catalytic reforming systems in a petroleum refinery and time series prediction. However, to the best of our knowledge, this is the first time NARX have been used to predict transformer top-oil temperature validated with many units in real operation for several years.

NARX inputs are the ambient temperature and the electrical load. A cross-correlation analysis shows that there is a non-linear relationship between these variables and the top-oil temperature. This indicates that these variables may be good predictors because they add new information to the model. The output is the top-oil temperature, which is compared to the top-oil temperature measurement used for controlling the cooling system.

A drop in performance of the cooling system may be due, for instance, to failure of a fan or pollution on the radiator fins. In this case, the measured temperature will be higher than the temperature predicted by the model. Then, a warning is sent to the maintenance personnel, and corrective actions are carried out before a high temperature alarm is sent to the network operators.

The cooling stage information is not necessary for NARX models: comparison experiments have demonstrated that adding this information as input to a NARX model has no significant impact on its performance. This may be due to the correlation between the cooling stages and NARX inputs: cooling stages depend on the winding temperature, which is directly correlated with the ambient temperature and the load.

Time steps are defined as five-minute periods. The feedback delay n_{y} corresponds to the highest auto-correlation value for the top-oil temperature signal. The input delay n_{x} is calculated as the maximum of input signal time delays, where each input signal time delay corresponds to the highest cross-correlation value with the top-oil temperature signal. $f(.)$ is represented as a feedforward neural network; it has been calibrated to one hidden layer of 10 neurons.

NARX has been implemented in MATLAB with the Neural Network Toolbox [27] (see Figure 2). The chosen performance function is the mean-squared error. A model has been trained for each transformer on one year of historical data with the Levenberg-Marquardt algorithm [28] during 1 000 iterations.

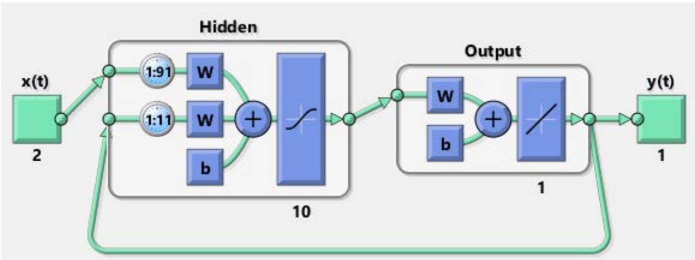


Figure 2: Schema of NARX for a transformer with $n_{\mathrm{x}} = 91$ and $n_{\mathrm{y}} = 11$ (figure taken from MATLAB)

4. FIELD EXPERIMENTS

4.1 Training data selection, cleaning and preparation

As discussed in Section 2, thermal performance of transformers with natural oil convection is strongly influenced by the oil viscosity and its dependence with temperature. A suitable training period should then include all four seasons. During training, the cooling system is assumed to be in good condition. If the training is carried out while the cooling system is deteriorated and maintenance is performed after training, the measured temperature would probably become lower than the predicted temperature. In this case, a new training should be initiated.

Before training the NARX, data cleaning and preparation must be performed. The measured data samples are not sent simultaneously. They are stored using an exception technique, i.e., they are stored only if the value has changed more than a predetermined threshold (also called dead band). A simple linear interpolation between registered samples is made to feed the model with data having a fixed 5-minute time step. Communication or sensor issues may interrupt the data flow to the centralized server so interpolation is stopped when the time between two consecutive measurement samples exceeds a predetermined threshold. The same interpretation rules apply when the model is used in monitoring mode. Inconsistent data that may be transmitted during testing of the sensors or bad data (due to noise or some other unknown) are removed from the training data set.

Considering these elements, it was possible to apply training on 54 transformers, and a total of 140 transformer-years of good monitoring data were available when this paper was prepared. Transformer cooling uses natural and forced oil flow in windings with natural or forced-air circulation through radiators. The ratings of the transformers vary from 22.5 to 450 MVA and from 120 to 315 kV.

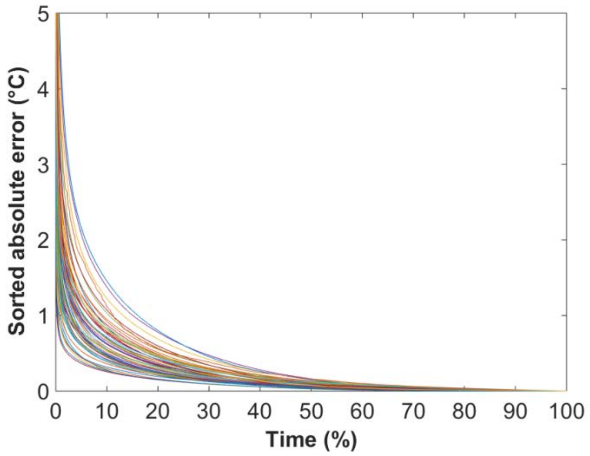
4.2 Performance evaluation

A good way to evaluate the performance of NARX prediction is to plot the sorted absolute error as a function of normalized time. Hence, it is possible to compare data from transformers with different durations of historical data, as shown in Figure 3. Figure 4 illustrates the 90th percentile error for each transformer as a function of the historical data duration. It can be seen that for 49 transformers out of 54, the error is below 1°C for 90% of the data samples.

NARX performance is similarly good at all temperatures and load conditions. Figure 5 shows the sorted errors for four transformers in La Suète substation, with data clustered depending on ambient temperature and load. This demonstrates that NARX performance is not influenced by these parameters. Moreover, the transformers were operated in various cooling stage conditions, and they showed no influence on estimation accuracy.

Figure 6 and 7 show the comparison between the developed NARX model and the model from Swift [8] with constant rated temperature rise (corresponding to the maximum cooling stage rating), which is very similar to the loading guide top-oil models [2, 3]. The graphs show the number of running fans, the load (with different profiles in summer and winter) and the ambient temperature. The NARX model estimation is essentially superposed with the measured top-oil temperature. The loading guide model does not consider the number of fans running and assumes that all fans are running at all times. In order to assess its performance, some areas are identified when all the fans are running. It can be noted that in summer, the loading guide model performs better, however not as well as the NARX model. In winter, the loading guide model deviation is increased due to the effect of increased oil viscosity. This clearly shows that using a loading guide dynamic model that does not correctly account for the low ambient temperature behavior can produce a false alarm regarding cooling system performance.

The NARX model has shown significantly reduced performance after a transformer step-load change or when the input data (ambient temperature or load) is contaminated with noise, as shown in Figure 8 and 9. In these cases, the NARX estimation becomes unstable and diverges from the normal value for a duration that can last a few hours in the case of a step-load change or longer if the deviation is caused by noisy input values. Since the deviation is significantly different from the normal behavior, it cannot be misinterpreted as a cooling system failure. It is then possible to post-process the data to eliminate these deviations from the analysis.



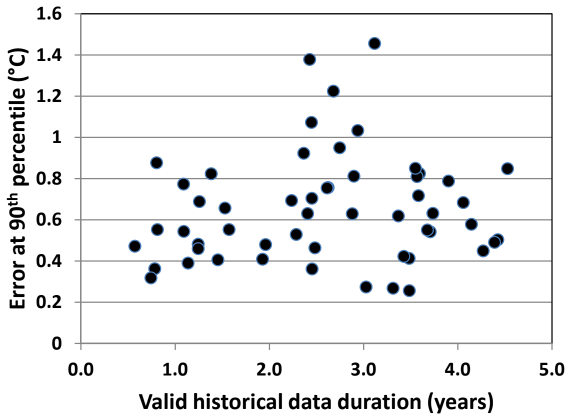


Figure 3: Sorted absolute error for 54 transformers representing 140 transformer-years of data

Figure 4: Error at 90% percentile for each transformer as a function of duration of historical data (90% of the data has lower error)

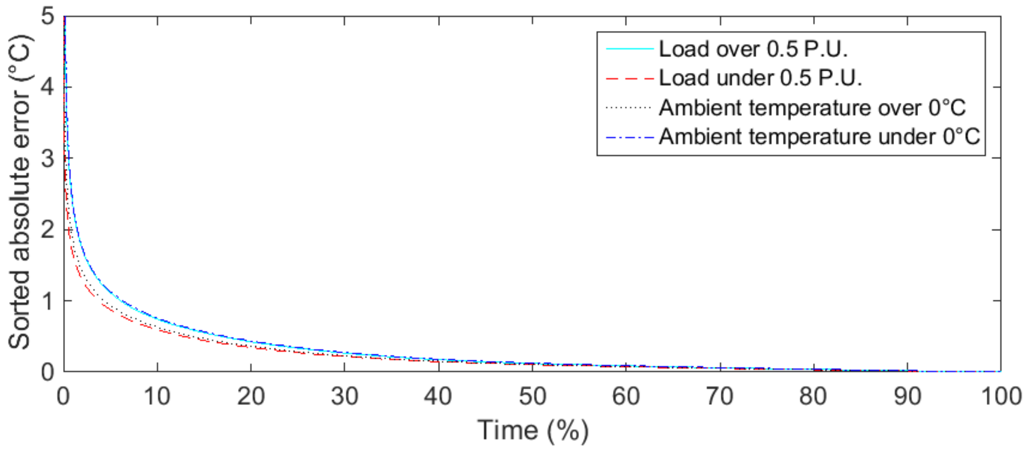


Figure 5: Sorted absolute error for four transformers in La Suète substation clustered as a function of ambient temperature and load

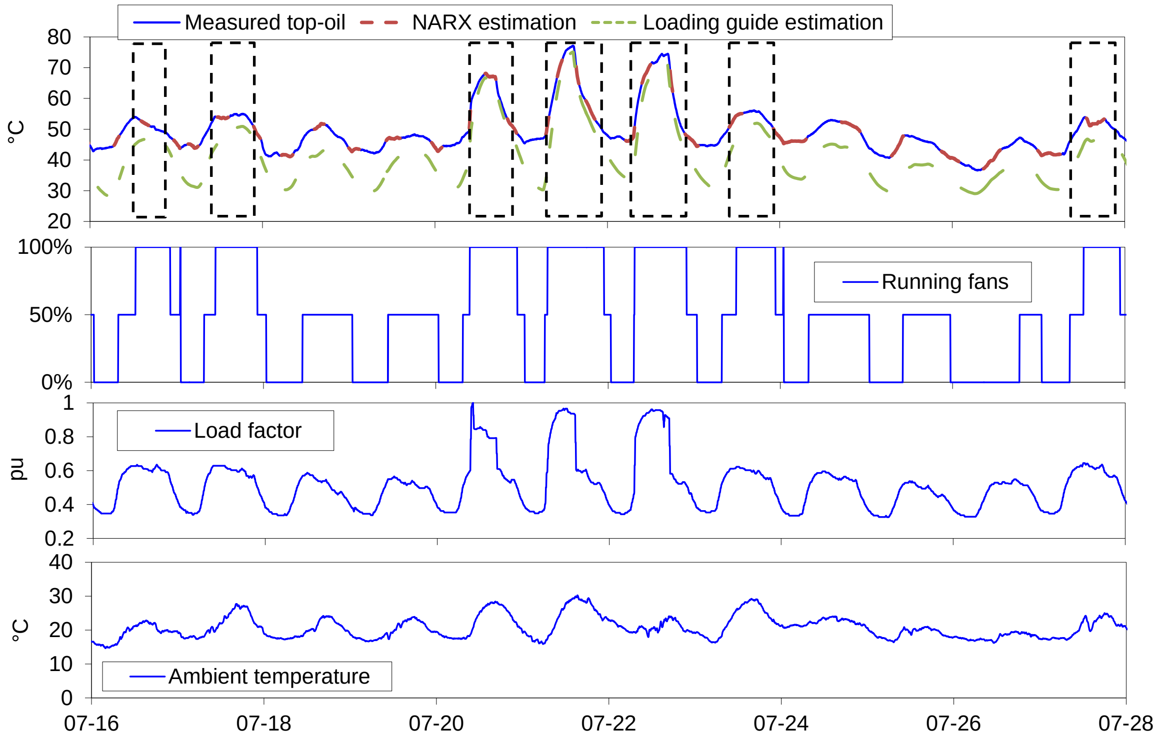


Figure 6: Comparison of loading guide and NARX estimations for a 66 MVA transformer operated in summer

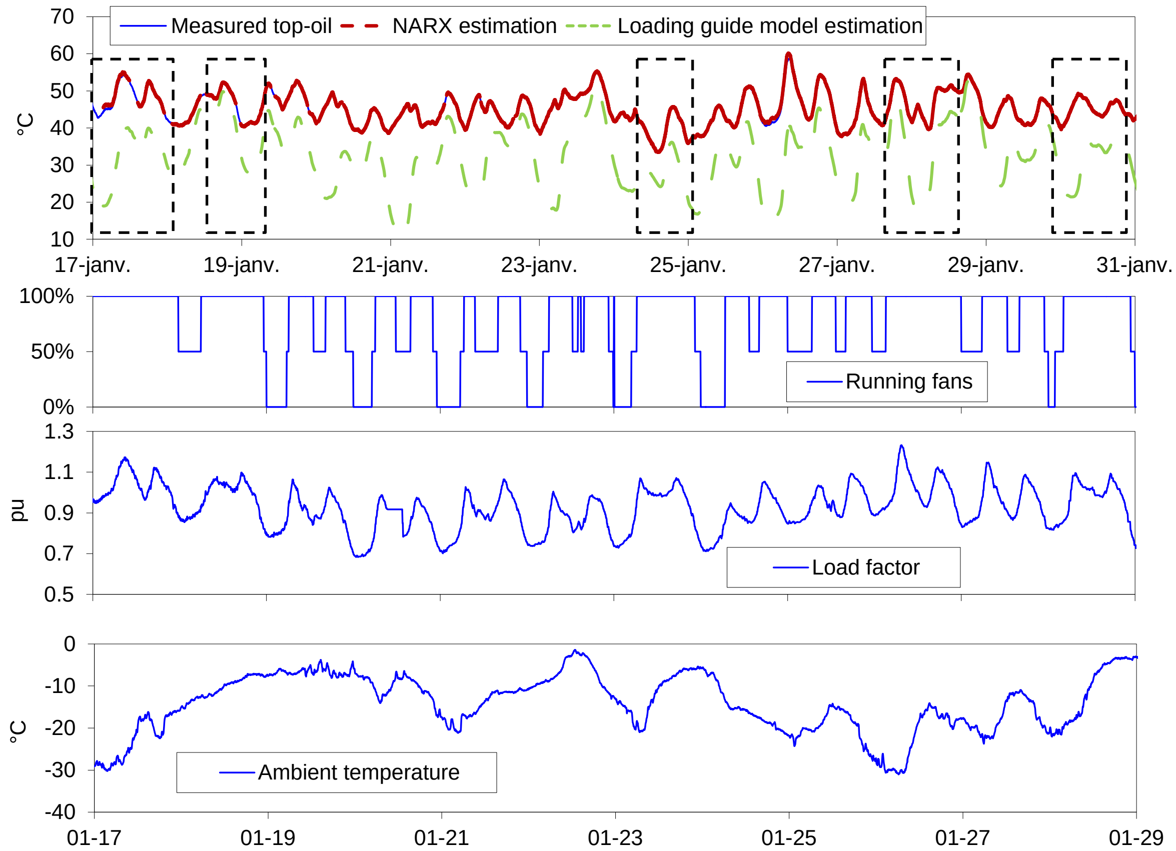


Figure 7: Comparison of loading guide and NARX estimations for a 66 MVA transformer operated in winter

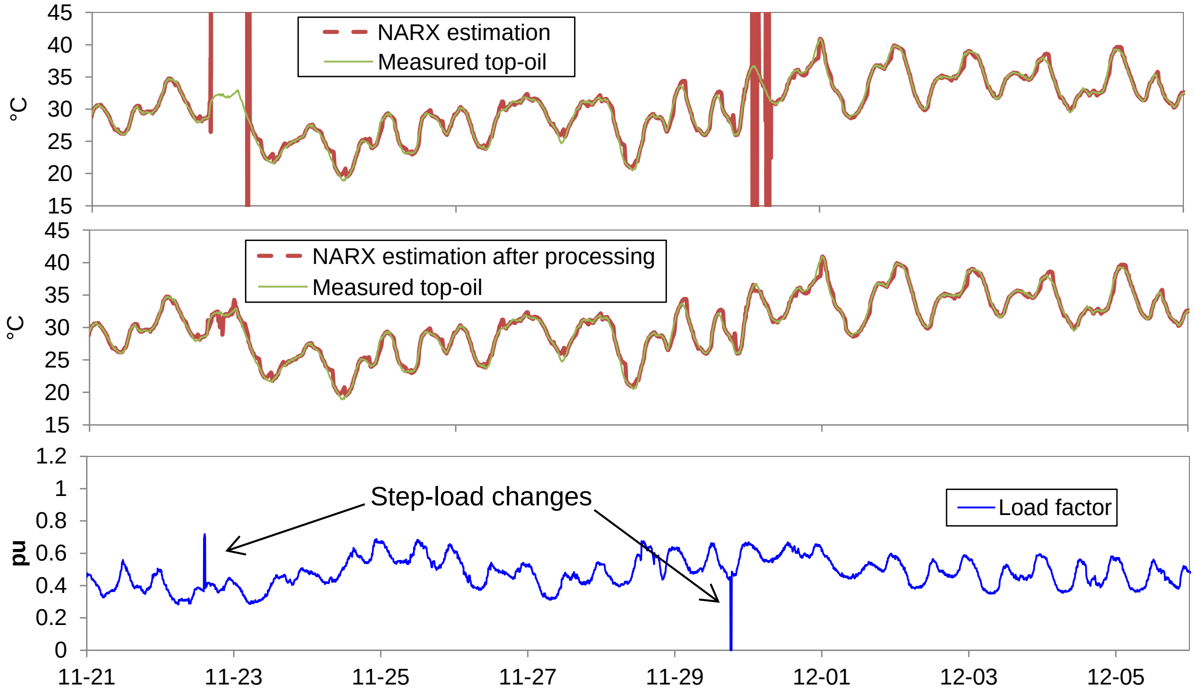


Figure 8: Step-load changes leading to NARX estimation instability

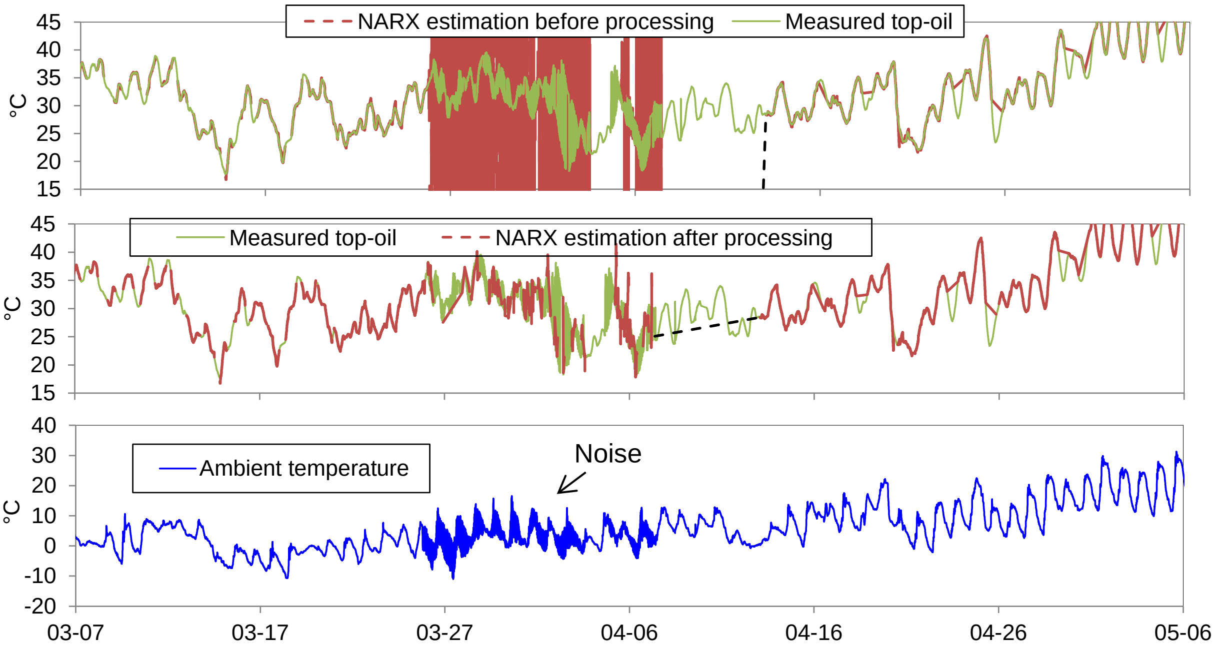


Figure 9: Noisy ambient temperature measurements leading to NARX estimation instability

5. FAILURE SIMULATION: MANUAL CHANGE OF COOLING CONDITION

A simple way to detect a failure of the cooling system could be based on the daily average error on estimated top-oil temperature. If the top-oil temperature is estimated accurately, this error will be low during normal operation of the transformer. In order to induce a variation of normal cooling system behavior, the fans of a transformer were manually activated for some weeks. The top-oil temperature was then lower than in normal operation, as shown in Figure 10. As expected, it is observed that the NARX model is able to capture this abnormal deviation. Indeed, the estimated top-oil temperature is consistently higher than the measured temperature while the fans are manually operated. Improved deviation algorithms can be applied using moving average or more advanced statistical analysis of the deviations, as proposed in [29].

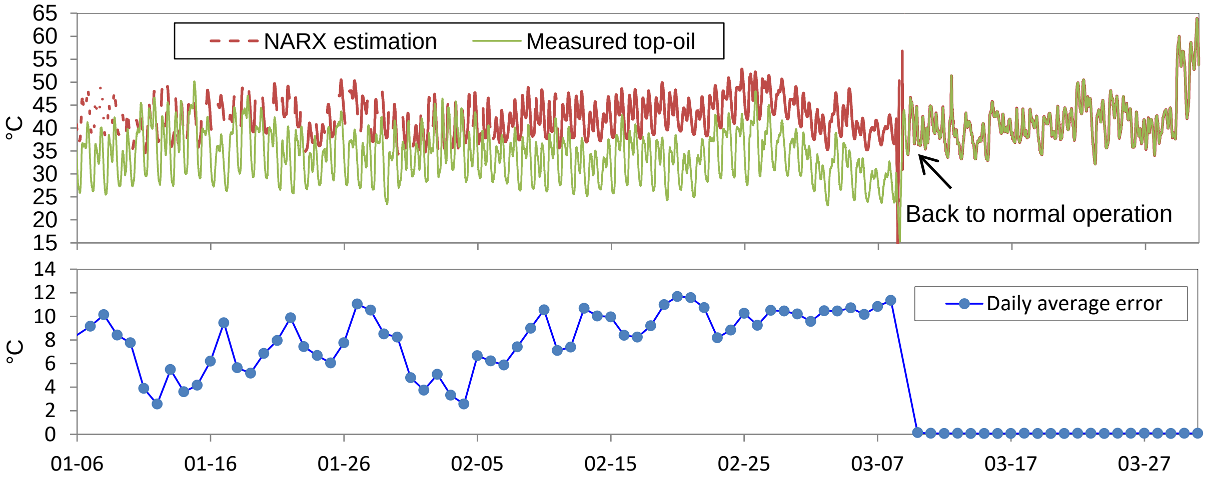


Figure 10: Measurement and NARX estimation during manual operation of the fans

6. CONCLUSION

A dynamic thermal model can be used to monitor transformer cooling performance. The model must be able to predict the temperature in a wide range of ambient temperatures and loads. This is still a challenge for models based on the physics of thermal phenomena.

To tackle this complex problem, the machine learning approach has been applied more frequently in recent years. The NARX model presented in this study demonstrates an excellent performance on many transformers with natural and forced oil circulation in the windings and the cooling system. The model showed similar performance in summer and winter conditions and with load varying from 0 to 1.4 pu. However, the NARX model showed inadequate performance for sudden load changes because the training is based mostly on slow load fluctuations. Similar deviations were observed in the case of noisy input signals. These abnormal estimations were identified and removed with post-processing of the output. As the purpose of the model is to detect cooling performance changes, these isolated events should not diminish the potential of the proposed approach.

Future work will concentrate on on-line training that would be automatically initiated when sufficient input data is available.

7. ACKNOWLEDGMENT

The authors would like to acknowledge the contribution of J. Picher, trainee at IREQ, for assistance with the data processing and for comments that improved the overall manuscript.

8. BIBLIOGRAPHY

[1] CIGRE Working Group A2.38, "Tranformer thermal modelling," *Brochure 659*, 2016.

[2] "Loading guide for oil-immersed power transformers," *IEC 60076-7 (2005-12)*, 2005.

[3] "IEEE Guide for Loading Mineral-Oil-Immersed Transformers and Step-Voltage Regulators - Redline," *IEEE Std C57.91-2011*, pp. 1-172, 2012.

[4] J. Aubin and Y. Langhame, "Effect of oil viscosity on transformer loading capability at low ambient temperatures," *IEEE Transactions on Power Delivery*, vol. 7, pp. 516-524, 1992.

[5] H. Nordman, N. Rafsback, and D. Susa, "Temperature responses to step changes in the load current of power transformers," *IEEE Transactions on Power Delivery*, vol. 18, pp. 1110-1117, 2003.

[6] L. Jauregui-Rivera and D. J. Tylavsky, "Acceptability of four transformer top-oil thermal models-part I: Defining metrics," *IEEE Transactions on Power Delivery*, vol. 23, pp. 860-865, 2008.
[7] L. Jauregui-Rivera and D. J. Tylavsky, "Acceptability of four transformer top-oil thermal models - Part II: Comparing metrics," *IEEE Transactions on Power Delivery*, vol. 23, pp. 866-872, 2008.
[8] G. Swift, T. S. Molinski, and W. Lehn, "A fundamental approach to transformer thermal modeling - Part I: Theory and equivalent circuit," *IEEE Transactions on Power Delivery*, vol. 16, pp. 171-175, 2001.
[9] G. Swift, T. S. Molinski, R. Bray, and R. Menzies, "A fundamental approach to transformer thermal modeling - Part II: Field verification," *IEEE Transactions on Power Delivery*, vol. 16, pp. 176-180, 2001.
[10] D. Susa, M. Lehtonen, and H. Nordman, "Dynamic thermal modelling of power transformers," *IEEE Transactions on Power Delivery*, vol. 20, pp. 197-204, 2005.
[11] M. K. Pradhan and T. S. Ramu, "Online monitoring of temperature in power transformers using optimal linear combination of ANNs," in *Conference Record of the 2004 IEEE International Symposium on Electrical Insulation*, 2004, pp. 70-73.
[12] D. J. Tylavsky, "Transformer top-oil temperature modeling and simulation," *IEEE Transactions on Industry Applications*, vol. 36, pp. 1219-1225, 2000.
[13] J. L. Velasquez-Contreras, M. A. Sanz-Bobi, and S. Galceran Arellano, "General asset management model in the context of an electric utility: Application to power transformers," *Electric Power Systems Research*, vol. 81, pp. 2015-2037, 2011/11/01/ 2011.
[14] 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, 2000.
[15] D. Villacci, G. Bontempi, A. Vaccaro, and M. Birattari, "The role of learning methods in the dynamic assessment of power components loading capability," *IEEE Transactions on Industrial Electronics*, vol. 52, pp. 280-290, 2005.
[16] M. Ali Taghikhani, "Power Transformer Top Oil Temperature Estimation with GA and PSO Methods," *Energy and Power Engineering*, vol. 04, 2012.
[17] M. Hell, P. Costa, and F. Gomide, "Participatory Learning in Power Transformers Thermal Modeling," *IEEE Transactions on Power Delivery*, vol. 23, pp. 2058-2067, 2008.
[18] L. M. Souza, A. P. Lemos, W. M. Caminhas, and W. C. Boaventura, "Thermal modeling of power transformers using evolving fuzzy systems," *Engineering Applications of Artificial Intelligence*, vol. 25, pp. 980-988, 2012.
[19] G. Rigatos and P. Siano, "Power transformers' condition monitoring using neural modeling and the local statistical approach to fault diagnosis," *International Journal of Electrical Power and Energy Systems*, vol. 80, pp. 150-159, 2016.
[20] A. Seier, P. D. H. Hines, and J. Frolik, "Data-Driven Thermal Modeling of Residential Service Transformers," *IEEE Transactions on Smart Grid*, vol. 6, pp. 1019-1025, 2015.
[21] A. Bracale, G. Carpinelli, M. Pagano, and P. D. Falco, "A Probabilistic Approach for Forecasting the Allowable Current of Oil-Immersed Transformers," *IEEE Transactions on Power Delivery*, vol. 33, pp. 1825-1834, 2018.
[22] L. Breiman, "Random Forests " *Mach. Learn*, vol. 45, pp. 5-32, 2001.
[23] M. R. G. Meireles, P. E. M. Almeida, and M. G. Simoes, "A comprehensive review for industrial applicability of artificial neural networks," *IEEE Transactions on Industrial Electronics*, vol. 50, pp. 585-601, 2003.
[24] A. J. Smola, B. J. S. Schölkopf, and Computing, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, pp. 199-222, August 01 2004.
[25] H. T. Siegelmann, B. G. Horne, and C. L. Giles, "Computational capabilities of recurrent NARX neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 27, pp. 208-215, 1997.
[26] J. Principe, N. Euliano, and W. Curt Lefebvre, *Neural and Adaptive Systems : Fundamentals through Simulations*: Wiley New York, 2000.
[27] M. H. Beale, M. T. Hagan, and H. B. J. U. s. G. Demuth, MathWorks, "Neural network toolbox," vol. 2, pp. 77-81, 2010.
[28] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, pp. 431-441, 1963.
[29] M. Djamali and S. Tenbohlen, "A validated online algorithm for detection of fan failures in oil-immersed power transformers," *International Journal of Thermal Sciences*, vol. 116, pp. 224-233, 2017.