International Journal of Electrical and Computer Engineering (IJECE)
Vol. 16, No. 1, February 2026, pp. 1~9
ISSN: 2088-8708, DOI: 10.11591/ijece.v16i1.pp1-9

Machine learning-based prediction of moisture and oxygen in a large power transformer with online monitoring validation

Osama T. Ghazal, Mohammed S. Assaf
Department of Maintenance, ATTARAT Operation and Maintenance Company (OMCO), Amman, Jordan

Article Info

Article history:

Received Apr 5, 2025
Revised Sep 19, 2025
Accepted Nov 23, 2025

Keywords:

Dissolved gases
Moisture migration
Moisture prediction
Oxygen prediction
Power transformers

ABSTRACT

This study presents a predictive modeling approach for monitoring moisture and dissolved oxygen dynamics in a newly commissioned high-capacity power transformer. Using over 48,000 real-time observations collected across three years via an advanced online monitoring device installed on a 326 MVA generator step-up transformer (GSUT), machine learning models were developed to estimate moisture and oxygen concentrations based on correlated operational parameters. Multiple regression-based algorithms were trained and evaluated using performance metrics including root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2). Linear regression achieved superior performance with an RMSE values as low as 0.05888 ppm for oxygen and 0.0153 ppm for moisture. The models were further validated using data from a sister transformer, demonstrating generalizability and reliability across similar transformer units. This work contributes a scalable and accurate solution for real-time transformer health assessment, with practical implications for predictive maintenance strategies in power utilities.

Corresponding Author:

Mohammed S. Assaf
Department of Maintenance, ATTARAT Maintenance and Operation Company (OMCO)
Level 8, 213 Zahran Street, Amman, 11822, Jordan
Email: Mohammad.assaf@aom.com.jo

1. INTRODUCTION

The reliable operation of power transformers is a cornerstone of stable and secure electrical power systems. These large and capital-intensive assets play a pivotal role in voltage transformation and load distribution across power networks. Given their criticality, any failure can lead to prolonged outages, costly repairs, and safety risks. One of the most significant threats to transformer longevity and reliability is the degradation of its insulation system, especially due to moisture ingress and oxidation. Moisture in transformer oil adversely affects the electric field distribution, reduces dielectric strength, and accelerates the degradation of cellulose-based solid insulation. Furthermore, dissolved oxygen acts as a catalyst in oxidation processes, leading to the formation of acids, sludges, and other aging byproducts that compromise both oil and solid insulation performance.

Numerous studies have established the importance of monitoring moisture content as a core parameter of transformer health [1], [2]. Variations in oil temperature, often caused by thermal loading, influence the migration of moisture between oil and paper insulation layers, further complicating condition assessment [3]. The dynamic interaction between temperature, moisture, and dissolved oxygen has a direct impact on the transformer's dielectric behavior, thermal performance, and aging rate. Understanding and accurately predicting this interaction is essential for implementing predictive maintenance strategies and avoiding premature failures.

Recent advancements in artificial intelligence have opened new avenues for transformer diagnostics. Machine learning (ML) models, in particular, have shown promising results in estimating moisture content and insulation condition using laboratory and field data [4]–[6]. These models can capture complex, nonlinear relationships among operational parameters, such as gas concentrations and temperature, which are difficult to quantify using traditional analytical techniques. Some approaches employ dielectric frequency response (DFR) or dissolved gas analysis (DGA) as inputs, while others use neural networks or hybrid fuzzy logic to enhance prediction accuracy. However, many of these studies are either limited in scale, rely on synthetic datasets, or lack validation under actual field conditions.

In this paper, we address this gap by presenting a comprehensive case study involving a newly commissioned generator step-up transformer (GSUT) rated at 326 MVA at ATTARAT Power Company (APCO), Jordan. An online monitoring device was installed to continuously track dissolved gases and moisture concentrations over a three-year period. Using this large and high-resolution dataset (over 48,000 observations), we developed and evaluated multiple machine learning models to predict the levels of moisture and dissolved oxygen under varying operational conditions. The models were further validated using data from a sister transformer to assess generalizability and robustness. This approach enhances condition-based maintenance practices and provides valuable insights into transformer insulation behavior under dynamic thermal and environmental conditions [7], [8].

2. MOISTURE, DISSOLVED OXYGEN, HYDROGEN IN TRANSFORMER INSULATION SYSTEM

2.1. Sources

High-quality manufacturing of power transformers necessitates strict control of moisture content, especially during the critical drying phase of cellulose insulation. This process aims to reduce the water content to below 1.0% in accordance with industry standards [9], [10]. Despite these measures, the water content in the insulation system often increases post-manufacture due to several operational factors. Exposure to ambient air during transportation, storage, and maintenance can lead to moisture ingress. Additionally, moisture migration occurs between the oil and solid insulation under thermal cycling, as the solubility of water in oil changes with temperature. Dissolved oxygen and hydrogen enter the transformer oil through multiple pathways, including air ingress from maintenance activities, breathing actions caused by temperature-induced oil expansion and contraction, leaks, and degradation of insulating materials. These gases can also form during the oxidation of oil and cellulose, especially in the presence of moisture, accelerating aging and deteriorating dielectric performance [11], [12].

2.2. Effect on the transformer insulation system

Moisture and oxygen are among the most influential factors driving the aging and degradation of both liquid and solid transformer insulation. High moisture levels reduce the breakdown voltage, increase dielectric losses, and significantly decrease the mechanical integrity of cellulose insulation. Prolonged exposure leads to hydrolytic and oxidative degradation, resulting in the generation of acids, furans, and sludge, which compromise transformer reliability [11], [13].

Moisture promotes the formation of polar contaminants and corrosive byproducts, which negatively affect heat dissipation and insulation strength. Moreover, the presence of dissolved oxygen accelerates oil oxidation and increases the acidity level, further degrading the dielectric properties of the system. This degradation leads to increased power losses, reduced insulation life, and a higher likelihood of incipient faults [7]. Due to these critical impacts, monitoring moisture and dissolved oxygen content is essential in high-voltage transformers to ensure safe operation and to enable timely interventions [11], [13].

2.3. Migration and equilibrium characteristics

The movement of moisture within oil-paper insulation systems is governed by temperature gradients and vapor pressure differentials. These forces drive moisture migration between solid and liquid phases, especially during daily loading cycles. The rate of migration is characterized by the diffusion coefficient (D), which depends on temperature and insulation condition.

As the transformer oil heats up, its ability to dissolve water increases, temporarily reducing the relative water saturation in oil. This results in moisture being released from the paper insulation into the oil. During cooling, the reverse migration can occur. These cyclical exchanges highlight the dynamic nature of moisture distribution and the importance of continuous monitoring.

Oommen's moisture equilibrium curves, developed in 1983, are widely used to describe the relationship between moisture content in paper and oil under thermal equilibrium. These curves form the basis for estimating the moisture content in the solid insulation based on oil measurements and temperature

[14]. They have since been refined and validated through both experimental studies and field measurements [15], [16].

2.4. Water solubility (S) and relative saturation (RS)

Water content in transformer oil is typically measured in parts per million (ppm) using Karl Fischer titration, following IEC 60814 or ASTM D1533 standards. However, this method alone does not account for the saturation state of the oil at a given temperature. Therefore, relative saturation (RS) and solubility (S) are used to better assess moisture risk. Solubility (S) is defined as the maximum amount of water that can dissolve in oil at a specific temperature and varies exponentially with temperature [17]. Relative saturation (RS) represents the percentage of this solubility that is currently occupied by dissolved moisture.

RS is a preferred parameter for evaluating transformer moisture condition because it normalizes the ppm value across different oil volumes and transformer sizes. It also shows a strong correlation with dielectric breakdown strength and is a reliable metric for defining moisture thresholds in condition-based monitoring systems [17], [12]. Furthermore, RS is highly sensitive to temperature fluctuations, making it essential to interpret moisture data in conjunction with thermal profiles. By combining RS with dissolved gas analysis (DGA), a more comprehensive picture of insulation health can be achieved.

3. MACHINE LEARNUNG PREDICTION MODELS

3.1. Introduction

Machine learning (ML) has emerged as a powerful data-driven approach for predictive maintenance and diagnostics in electrical systems, including power transformers. Unlike conventional rule-based modeling, ML algorithms can learn complex relationships from empirical data through adaptive processes, offering enhanced accuracy in forecasting operational conditions. ML frameworks are broadly categorized into supervised, unsupervised, and reinforcement learning. In this study, supervised learning was employed to build regression models that predict two critical indicators of transformer insulation degradation: moisture (M) and dissolved oxygen (O_2). The models were developed using real-time data collected over a 3-year period, encompassing 48,000+ high-frequency observations. Inputs included temperature, hydrogen concentration, relative saturation, and solubility—features selected based on physical relevance and statistical correlation.

3.2. Performance measurements

To evaluate the performance of the predictive models, several well-established regression metrics were used:

a. Root mean squared error (RMSE): Measures the square root of the average squared prediction errors; penalizes large deviations.
b. Mean absolute error (MAE): Computes the average of absolute differences between predicted and actual values; interpretable and less sensitive to outliers.
c. Mean squared error (MSE): Square of RMSE; provides scale-dependent error magnitude.
d. Coefficient of determination (R^2): Indicates the proportion of variance in the dependent variable explained by the model.

R^2 values close to 1 reflect strong model performance. All metrics were computed using cross-validation to ensure robustness and prevent overfitting [18]–[20].

3.3. Model validation

To ensure reliability and generalization, two validation techniques were implemented:

a. Hold-out validation: the dataset was initially split into training (80%) and testing (20%) sets.
b. K-Fold cross-validation: the training data was further evaluated using 5-fold cross-validation, where the data was divided into five subsets. The model was iteratively trained on four subsets and validated on the fifth. This process mitigates model variance and ensures consistent evaluation [21].

All modeling and validation processes were conducted using MATLAB's regression learner app, which supports streamlined model comparison, hyperparameter tuning, and visualization [22].

3.4. Machine learning algorithms

Regression analysis is used to investigate the relationship function between variables, which is express several regression algorithms were evaluated to identify the optimal model for each response variable (moisture and oxygen). These include:

a. Linear models

Linear regression (LR) and its variants (interaction, robust, and stepwise linear) are interpretable and computationally efficient. They model the response as a linear combination of predictors [20]:

$$\hat{y} = f(x1, x2, \ldots, xp) + \mathcal{E} \quad (1)$$

whereas; $\hat{y}$ is response variable, $x1, x2, \ldots, xp$ are predictors, and $\mathcal{E}$ is random error.

b. Regression trees (RTs)

Fine, medium, and coarse decision trees split the dataset based on predictor thresholds. These models are easy to interpret but may overfit small datasets.

c. Ensembles of trees (EoTs)

Bagged trees and boosted trees combine multiple decision trees to improve predictive accuracy. Bagging reduces variance, while boosting reduces bias—though both increase computation time.

d. Support vector machines (SVMs)

Linear, polynomial (quadratic/cubic), and Gaussian kernel SVMs map inputs into higher-dimensional spaces to find optimal regression boundaries. While powerful, SVMs are more computationally demanding and sensitive to parameter tuning.

e. Gaussian process regression (GPR)

GPR models use kernel-based Bayesian learning to make probabilistic predictions. Despite their high accuracy, especially with limited noisy data, they incur high computational cost, particularly with large datasets [22].

4. CASE STUDY AND RESULTS

4.1. Transformer under study

This case study focuses on the generator step-up transformer unit 1 (GSUT1) at Attarat power company (APCO), Jordan, a newly commissioned unit energized in July 2020. The transformer is rated at 340 MVA, with a voltage rating of (420±8×1.25%)/20 kV, and employs a YNd11 vector group with ONAN/ONAF cooling. It is filled with approximately 70,230 kg of naphthenic-based insulating oil and manufactured by TBEA. Detailed technical specifications are presented in Table 1, providing a clear operational and structural profile of the transformer.

The GSUT1 is equipped with an advanced online monitoring system (HAOZHI ELECTRIC model W-PD2M) as shown in Figure 1. This device continuously tracks key fault gases and moisture content in the oil, including: Hydrogen (H_2), Carbon Monoxide (CO), Methane (CH_4), Acetylene (C_2H_2), Ethylene (C_2H_4), Ethane (C_2H_6), Carbon Dioxide (CO_2), and Oxygen (O_2). It also monitors oil temperature and relative humidity. Its compact size, low cost, and high data resolution make it well-suited for predictive maintenance applications.

Table 1. Unit1 GSUT technical data

Rated Power	340 MVA
Voltage Rating	(420±8×1.25%)/20 kV
Vector Group	YNd11
Rated Frequency	50 Hz
Cooling Mode	ONAN/ONAF (62/100%)
Manufacture	TBEA
Insulation oil mass	70230 kg
Oil Base	Naphthenic

(a) (b)

Figure 1. ATTARAT power plant (a) unit generator step-up transformer GSUT and (b) online key fault gas and moisture monitoring device

4.2. DGA and moisture monitoring device

The monitoring device captures readings every 30 minutes, generating over 48,000 observations spanning from July 2020 to August 2023. Raw parameters include: dissolved oxygen (O_2), dissolved hydrogen (H_2), moisture (M), and oil temperature (OT).

To enhance the feature set, two derived variables were calculated:

a. Solubility (S) [ppm]: The maximum water solubility in oil at a given temperature, calculated using empirical temperature-solubility relations [17].
b. Relative saturation (RS) [%]: Indicates moisture level as a fraction of saturation, providing temperature-normalized insight.

These features offer a deeper representation of transformer internal conditions and were integrated as predictors in the ML model.

4.3. Observations statistical analysis

Basic statistical descriptors for the monitored variables are summarized in Table 2, including minimum, median, maximum, and percentage of outlier values. Key observations:

a. Moisture (M) exhibited a narrow operational range (4.88 to 12.6 ppm), with minimal outliers (0.65%).
b. Solubility (S) and oil temperature (OT) showed wide variability, reflecting fluctuating environmental/load conditions.
c. Relative saturation (RS) displayed high stability and consistency, suggesting effective moisture equilibrium.
d. Oxygen (O_2) ranged significantly (514.6 to 1023.8 ppm) with higher outlier presence (24%).

This statistical foundation supports the development of robust predictive models and highlights key dynamic parameters.

Table 2. Unit1 GSUT data statistical summary

	Min	Median	Max	Outliers [%]
H2 [ppm]	0.1	1.4	11.32	5.7
O2 [ppm]	514.64	635.12	1023.8	24
M [ppm]	4.88	7.66	12.6	0.65
OT [°C]	24.48	30.27	48.74	25
S [ppm]	66.362	83.654	165.62	24.6
RS [%]	6.3999	8.465	10.669	0

4.4. Observations correlation and moisture equilibrium

To identify interdependencies among parameters, a Pearson correlation matrix was generated Table 3. Notable findings:

a. Oxygen (O_2) exhibited strong positive correlations with moisture (r=0.772), temperature (r=0.999), and solubility (r=0.996).
b. Moisture (M) correlated moderately with oil temperature (r=0.746) and solubility (r=0.744), suggesting a thermal coupling effect.
c. Hydrogen (H_2) had modest correlations with other variables, reflecting its possible generation from separate fault mechanisms.
d. Relative saturation (RS) showed weak negative correlations with all other parameters, particularly with temperature and solubility, consistent with moisture equilibrium behavior.

To further investigate moisture balance in the oil–paper insulation system, values were projected on Oommen's equilibrium curve for low-moisture regions Figure 2. This illustrates how moisture content follows predictable thermal dynamics based on equilibrium theory [23], [24]. The curve reinforces the interpretation that increased temperature drives moisture into the oil phase, which is then captured by real-time monitoring. These findings also justify the inclusion of solubility and RS as features in the predictive models.

Table 3. Unit1 GSUT data correlation matrix

	H2	O2	M	OT	S	RS
H2	1	0.528	0.392	0.529	0.536	-0.162
O2	0.528	1	0.772	0.999	0.996	-0.257
M	0.392	0.772	1	0.746	0.744	0.4093
OT	0.529	0.999	0.746	1	0.997	-0.294
S	0.536	0.996	0.744	0.997	1	-0.296
RS	-0.162	-0.257	0.409	-0.294	-0.29	1

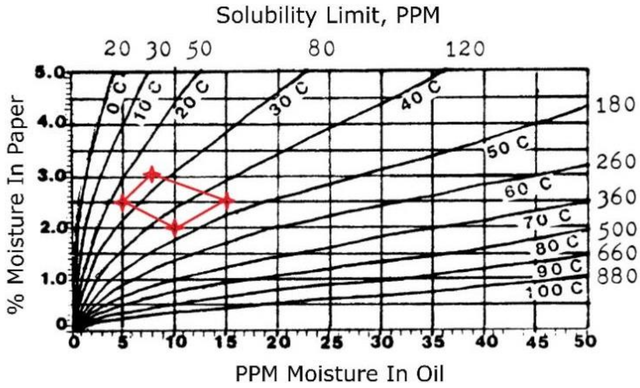


Figure 2. Moisture concentration projection on Oommen's curve for low moisture region of moisture equilibrium for a paper-oil-system

5. RESULTS AND DISCUSSION

5.1. Overall model performance

Table 4 presents key performance metrics—R^2, RMSE, and MAE—for all evaluated machine learning models, including linear regression (LR), regression trees (RT), Ensembles, SVMs, and GPR. Linear regression achieved exceptional accuracy with $R^2 \approx 0.995$, RMSE ≈ 0.0589 ppm for O_2, and RMSE ≈ 0.0153 ppm for moisture (M). More complex models like GPR produced similar accuracy but incurred significantly higher computational cost and training time.

The simplicity and computational efficiency of LR, combined with its accuracy, make it the most practical choice for real-time deployment. The hold-out test results, validated using K-fold cross-validation, indicate robust generalizability.

Table 4. Oxygen/moisture prediction machine learning algorithm training results

	Oxygen				Moisture			
Model	RMSE	R^2	MSE	MAE	RMSE	R^2	MSE	MAE
				LR				
Linear	0.05888	1	0.003467	0.050686	0.015304	1	0.00023422	0.013142
Interactions linear	0.058877	1	0.003467	0.050683	2.4183e-5	1	5.843e-10	2.069e-5
Robust linear	0.058881	1	0.003467	0.050687	0.015305	1	0.00023426	0.013138
Stepwise linear	0.058878	1	0.003467	0.050682	2.4428e-5	1	5.9672e-10	2.097e-5
				RTs				
Fine tree	1.2486	1	1.559	0.93961	0.039634	1	0.0015708	0.028681
Medium tree	1.2595	1	1.5862	0.94167	0.042293	1	0.0017887	0.029627
Coarse tree	1.3284	1	1.7647	0.96316	0.062231	1	0.0038728	0.039983
				EoTs				
Boosted Trees	28.802	0.88	829.57	28.492	0.3486	0.93	0.12152	0.32664
Bagged Trees	0.65222	1	0.42538	0.40312	0.026713	1	0.00071357	0.017165
				SVM				
Linear SVM	3.9199	1	15.366	3.5416	0.056043	1	0.0031409	0.044646
Quadratic SVM	2.9097	1	8.4661	2.1629	0.096878	0.99	0.0093853	0.08876
Cubic SVM	4.4189	1	19.527	4.1349	0.10277	0.99	0.010561	0.089467
Fine Gaussian SVM	6.6679	0.99	44.461	4.4679	0.11236	0.99	0.012626	0.092983
Medium Gaussian SVM	3.3794	1	11.42	2.6374	0.066318	1	0.0043981	0.05519
Coarse Gaussian SVM	4.902	1	24.03	4.5941	0.08080461	1	0.0064739	0.0064739
				GPR				
Rational Quadratic	0.058988	1	0.003480	0.05741	0.0012582	1	1.5829e-6	0.00093856
Squared Exponential	0.058982	1	0.003479	0.050745	0.0010864	1	1.1804e-6	0.00082484
Matern 5/2	0.41441	1	0.17174	0.063827	0.0064526	1	4.1636e-5	0.004764
Exponential GPR	0.058988	1	0.003480	0.05741	0.0052514	1	2.7577e-5	0.000559964

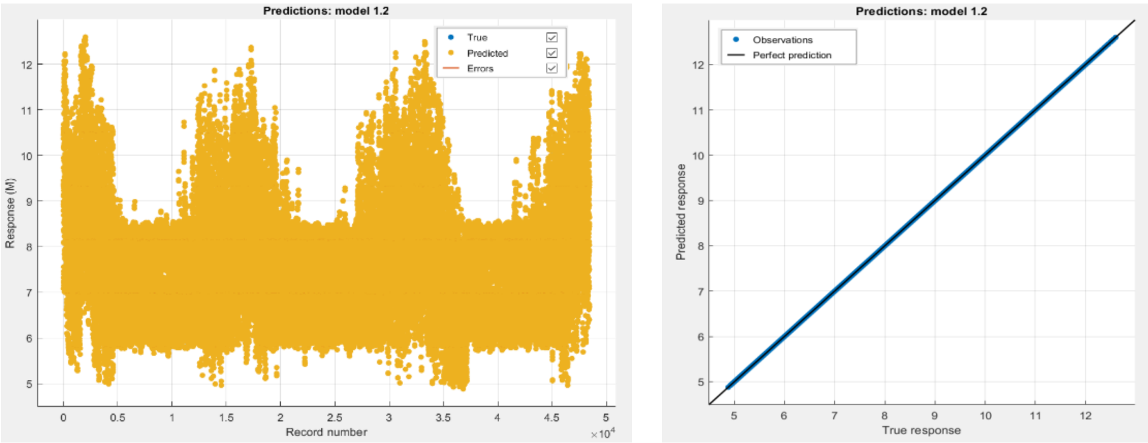
5.2. Comparative analysis with recent studies

Our results demonstrate meaningful improvements over recent literature, study [25] conducted a comprehensive review of moisture detection techniques and noted limitations in applying ML to operational transformer data. By contrast, our work leverages real-time, long-term monitoring data and advanced feature engineering (RS, solubility), surpassing previous approaches in realism and accuracy. This contribution positions our study as a pioneering application of ML with long-term, field-based transformer data.

5.3. Time-series patterns and physical insights

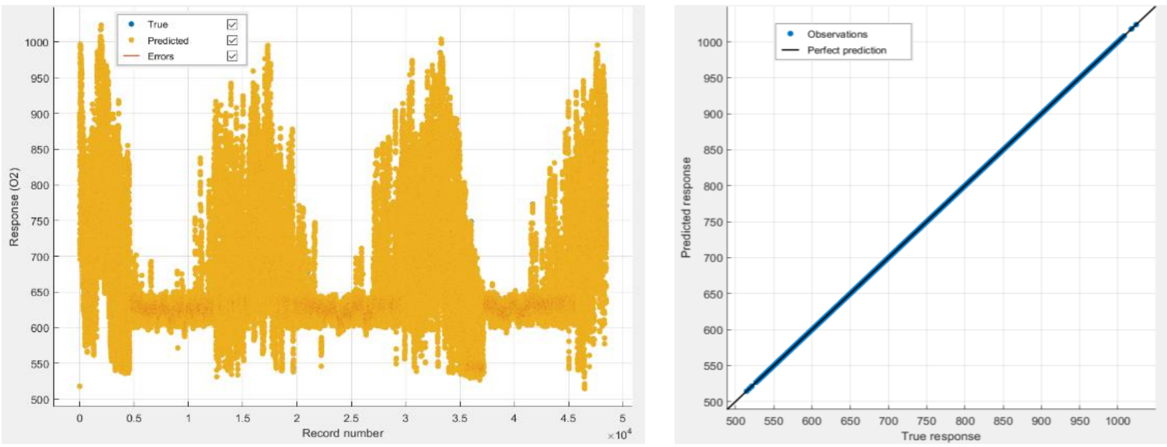
Figures 3 and 4 present the temporal profiles of moisture (M) and oxygen (O_2): moisture generally increases in cooler months, consistent with temperature-driven solubility changes and the observed relative saturation (RS) behavior, while oxygen shows periodic fluctuations likely attributable to system breathing and ambient ingress/contamination.

The strong negative correlation between RS and OT in Table 3 supports the thermodynamic principle: higher oil temperatures increase solubility, thereby reducing relative saturation. These findings align with early theoretical predictions about moisture equilibrium in transformer oil–paper systems. Our ML models effectively capture these dynamics, enhancing interpretability.



(a) (c)

Figure 3. Performance of the linear regression model for moisture prediction (a) moisture linear regression response plot and (b) moisture linear regression predicted vs actual plot linear regression response plot



(a) (c)

Figure 4. Performance of the linear regression model for dissolved oxygen O_2 prediction (a) dissolved oxygen (O^2) linear regression response plot, (b) dissolved oxygen, (o^2) linear regression predicted vs actual plot linear regression response plot

6. CONCLUSION

In conclusion, the proposed ML framework—led by a transparent, high-fidelity linear regression baseline—delivers reliable, real-time prediction of transformer moisture and oxygen that translates directly into operations: scheduling dehumidification or oil filtration when thresholds are approached, initiating sealing interventions when oxygen trends rise, and optimizing maintenance intervals to reduce avoidable costs. While findings are based on a single GSUT and would benefit from validation across diverse units and cooling schemes, the approach is readily extensible through additional features (e.g., furan, partial discharge indicators) and more expressive temporal models (e.g., transformer-style architectures) for longer-horizon forecasting and anomaly detection. Overall, this work advances beyond laboratory-centric studies by exploiting lifetime field data, demonstrating practical viability for condition-based maintenance, and laying a clear path toward scalable, intelligent transformer asset management.

ACKNOWLEDGMENTS

The authors would like to thank ATTARAT Power Company (APCO) and ATTARAT Operation and Maintenance Company (OMCO) management for their encouragement and support in accomplishing this study and their approval of submitting this paper for publication.

REFERENCES

[1] S. R. Arsad *et al.*, "Water content in transformer insulation system: a review on the detection and quantification methods," *Energies*, vol. 16, no. 4, p. 1920, Feb. 2023, doi: 10.3390/en16041920.
[2] R. Liao, Y. Lin, P. Guo, H. Liu, and H. Xia, "Thermal aging effects on the moisture equilibrium curves of mineral and mixed oil-paper insulation systems," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, no. 2, pp. 842–850, Apr. 2015, doi: 10.1109/TDEI.2015.7076783.
[3] W. Sikorski, K. Walczak, and P. Przybylek, "Moisture migration in an oil-paper insulation system in relation to online partial discharge monitoring of power transformers," *Energies*, vol. 9, no. 12, p. 1082, Dec. 2016, doi: 10.3390/en9121082.
[4] A. Vatsa, A. S. Hati, V. Bolshev, A. Vinogradov, V. Panchenko, and P. Chakrabarti, "Deep learning-based transformer moisture diagnostics using long short-term memory networks," *Energies*, vol. 16, no. 5, p. 2382, Mar. 2023, doi: 10.3390/en16052382.
[5] C. Weigen and L. Qiang, "On-line monitoring model based on neural network for moisture content in transformer oil (in Chinese)," *High Voltage Engineering*, vol. 33, no. 5, pp. 73–78, 2007.
[6] V. V. Vasylevsky, O. O. Kaplienko, and S. I. Shylo,, "Application of neural networks to predict the moisture content in oil-filled power transformers insulation," *Bulletin of NTU "KhPI". Series: Problems of Electrical Machines and Apparatus Perfection. The Theory and Practice*, vol. 0, no. 2, pp. 8–12, Dec. 2019, doi: 10.20998/2079-3944.2019.2.02.
[7] M. Wang, A. J. Vandermaar, and K. D. Srivastava, "Review of condition assessment of power transformers in service," *IEEE Electrical Insulation Magazine*, vol. 18, no. 6, pp. 12–25, Nov. 2002, doi: 10.1109/MEI.2002.1161455.
[8] N. Bakar, A. Abu-Siada, and S. Islam, "A review of dissolved gas analysis measurement and interpretation techniques," *IEEE Electrical Insulation Magazine*, vol. 30, no. 3, pp. 39–49, May 2014, doi: 10.1109/MEI.2014.6804740.
[9] S. Siswanto, A. Indarto, R. Rahmatullah, and C. Hudaya, "The effects of drying time during manufacturing process on partial discharge of 83.3 mva 275/160 kv power transformer," *MATEC Web of Conferences*, vol. 218, p. 4008, 2018, doi: 10.1051/matecconf/201821804008.
[10] J. Aubin *et al.*, "Moisture equilibrium and moisture migration within transformer insulation systems," *Electra*, no. 238, 2008.
[11] S. Abdi, L. Safiddine, A. Boubakeur, and A. Haddad, "The effect of water content on the electrical properties of transformer oil," in *Lecture Notes in Electrical Engineering*, vol. 599 LNEE, Springer International Publishing, 2020, pp. 518–527, doi: 10.1007/978-3-030-31680-8_52.
[12] I. S. C57.106™-2002, "IEEE guide for acceptance and maintenance of insulating oil in equipment," *Communications Society*, vol. 2002, no. April, pp. 1–34, 2003.
[13] S. S. Dessouky, A. E. Kalas, R. A. A. El-Aal, and A. M. M. Hassan, "Study and examination of transformer oil while exposed to air during operation," in *2017 19th International Middle-East Power Systems Conference, MEPCON 2017 - Proceedings*, Dec. 2017, vol. 2018-February, pp. 445–450, doi: 10.1109/MEPCON.2017.8301218.
[14] Y. Du, "Moisture equilibrium in transformer paper-oil systems," *IEEE Electrical Insulation Magazine*, vol. 15, no. 1, pp. 11–20, Jan. 1999, doi: 10.1109/57.744585.
[15] CIGRE, "Moisture equilibrium and moisture migration within transformer insulation systems," CIGRE, no. 349, 2008.
[16] S. Tenbohlen and M. Koch, "Aging performance and moisture solubility of vegetable oils for power transformers," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 825–830, Apr. 2010, doi: 10.1109/TPWRD.2009.2034747.
[17] L. R. Lewand, "Understanding water in transformer systems: The relationship between relative saturation and parts per million (ppm)," *NETA World Journal*, vol. 1, no. 4, 2002.
[18] J. Brownlee, *Machine learning mastery with Python*. Melbourne, Australia: Machine Learning Mastery, 2016.
[19] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014, doi: 10.5194/gmd-7-1247-2014.
[20] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning: with applications in R*, vol. 103. New York: Springer, 2013, doi: 10.1007/978-1-4614-7138-7.
[21] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*, 1995, pp. 1137–1143.
[22] MathWorks, "Regression learner app," *MathWorks*, 2023. https://www.mathworks.com/help/stats/regression-learner-app.html
[23] O. Roizman, V. Davydov, and B. Ward, "Water-in-paper activity: A new approach for moisture management in transformers," *Energies*, vol. 16, no. 3, p. 1052, 2023, doi: 10.3390/en16031052.
[24] B. García, J. C. Burgos, Á. M. Alonso, and J. Sanz, "A moisture-in-oil model for power transformer monitoring - Part I: Theoretical foundation," *IEEE Transactions on Power Delivery*, vol. 20, no. 2 II, pp. 1417–1422, Apr. 2005, doi: 10.1109/TPWRD.2004.832366.
[25] Y. Du, A. V. Mamishev, B. C. Lesieutre, M. Zahn, and S. H. Kang, "Moisture solubility for differently conditioned transformer oils," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 8, no. 5, pp. 805–811, 2001, doi: 10.1109/94.959706.

BIOGRAPHIES OF AUTHORS

Osama T. Ghazal is a senior OT cybersecurity and DCS engineer with over 14 years of experience in the power generation sector. He holds an M.Sc. in mechatronics engineering and a B.Sc. in electrical power engineering from Al-Balqa'a Applied University, Jordan. Currently serving at Attarat Operation and Maintenance Company, he leads cybersecurity and control system efforts for a 2×270 MW oil shale thermal power plant. His expertise includes OT cybersecurity, DCS, PLCs, SCADA, and machine learning for predictive maintenance and process optimization. Osama is a GICSP-certified professional, has completed advanced ICS cybersecurity training from DHS-CISA and SANS, and has authored several publications in computer science and machine learning. He can be contacted at email: osama.ghazal@aom.com.jo.

Mohammed S. Assaf is a senior electrical maintenance engineer at Attarat Operation and Maintenance Company (OMCO), with over 13 years of experience in power plants. He holds a B.Sc. in electrical power engineering from Al-Balqa'a Applied University in Jordan. His expertise covers key power plant equipment, including generators, transformers, protection systems, MV switchgears, and emergency power supplies. Mohammed is a certified maintenance and reliability professional (CMRP) and a certified emergency power systems specialist (CEPSS). He is the author of two books on transformer testing, including power transformers electrical testing handbook: mastering seven electrical tests. He can be contacted at email: mohammad.assaf@aom.com.jo.