CIGI QUALITA MOSIM 2023
Power Transformer Fault Prediction based on Support Vector Machine

HADIS HESABI[1], MUSTAPHA NOURELFATH[2], ADNÈNE HAJJI[3], THIERRY JUNG[4]

[1] Ph.D. in Mechanical Engineering
Department of Mechanical Engineering, Laval University, Canada
Interuniversity Research Center on Enterprise Networks, Logistics and Transportation (CIRRELT), Canada
Hadis.Hesabi.1@ulaval.ca

[2] Full Professor
Department of Mechanical Engineering, Laval University, Canada
Interuniversity Research Center on Enterprise Networks, Logistics and Transportation (CIRRELT), Canada
Mustapha.Nourelfath@gmc.ulaval.ca

[3] Full Professor
Department of Operations and Decision Systems, Laval University, Canada
Interuniversity Research Center on Enterprise Networks, Logistics and Transportation (CIRRELT), Canada
Adnene.Hajji@fsa.ulaval.ca

[4] Senior SME, APM and Digital Solutions
GE Renewable Energy, Canada
Thierry.Jung@ge.com

Abstract

Power transformers (PTs) play a crucial role in power generation, transmission, and distribution systems. Monitoring the health of these transformers is essential to ensure an uninterrupted power supply. Dissolved Gas Analysis (DGA) is a widely used technique to examine the condition of PTs. However, predicting dissolved gas content in PTs is challenging due to non-linearity, high dimensionality, and limited training datasets.

This paper presents a novel approach to predict PT faults using the Support Vector Machine (SVM) algorithm based on DGA data. The proposed method employs SVM to achieve accurate and timely fault diagnosis, which is essential for preventing faults, as manual diagnosis is time-consuming and expensive. The real data does not include all the desired labels, so Gaussian simulation generates new data that provides all labels. The new data is generated using the Inverse Cumulative Distribution Function (ICDF) to convert the Gaussian samples to samples from the specified distributions.

The proposed approach achieves a probabilistic output for the fault diagnosis of oil-immersed transformers, overcoming the limitations of traditional DGA methods that often provide inaccurate diagnosis results and cannot summarize the fault development rule inductively. The case study results demonstrate the effectiveness of the proposed approach in predicting PT faults. Furthermore, this paper contributes a new method that utilizes SVM based on DGA data, which can help maintenance managers detect faults accurately and promptly, contributing to the PT fault diagnosis field.

***Keywords* – power transformer, dissolved gas analysis, fault diagnosis, support vector machine, machine learning.**

1 INTRODUCTION

Power transformers are a strategic and significant investment that is designed to operate efficiently for several decades. The growing demand for electricity often exceeds predictions, which puts a strain on the equipment and causes malfunctions and failures. In recent years, the importance of power quality has been a widely discussed topic due to the increased use of renewable energy sources, distributed power generation, power electronics, and electric vehicles. Effective electricity transmission is crucial for ensuring the stability of the future power grid, and new approaches have been proposed to address these challenges. However, existing electricity networks were designed long ago, which leads to power transformers operating at or beyond their limits under capacitive loads. This highlights the importance of monitoring and maintaining the health of power transformers to avoid malfunctions and failures, which can lead to power outages and financial losses. Power transformers, mainly oil-immersed, offer high electric power management with a small size compared to dry-type transformers. They transport and distribute electrical energy point to point by voltage step-up or step-down. Still, their insulation often deteriorates due to the high electric field in the high- and ultra-high-voltage electrical transportation system and environmental factors such as temperature, humidity, surface dust, and UV radiation. Therefore, to extend the lifetime of power transformers, it is critical to monitor their insulation conditions. The insulation of high and medium-power transformers is a critical factor that can impact their condition due to physicochemical reactions. This insulation comprises various components such as solid insulation parts, liquid insulation, and insulating oil. The most common parameters used for monitoring the condition of power transformer insulation are Dissolved Gas analysis (DGA), tangent delta, partial discharges, and insulation resistance.

DGA is the most commonly used technique for monitoring power transformer insulation and can be performed online or offline. It analyzes insulating oil for the presence of specific key gases, which can indicate different types of insulation-related problems. Various DGA techniques, such as the IEC ratio method, Doerenburg ratio method, Rogers' ratio method, and Duval's triangle method, use these gases in the form of ratios to analyze the data.

However, traditional DGA methods have limitations in detecting subtle faults, and the diagnosis could vary among experts due to their dependence on subjective knowledge and experience. To overcome these limitations, artificial intelligence (AI) techniques, such as machine learning (ML) and pattern recognition, have been increasingly adopted in the power industry. AI techniques can enhance the accuracy and generalizability of power transformer fault diagnosis, reducing the risk of failure and improving the reliability and efficiency of the electrical power system. SVM is one of the classifiers used to solve power transformer fault classification problems and is proposed in this paper for condition monitoring and fault diagnosis of power transformers based on DGA data.

This paper aims to highlight the importance of the effects of insulation in power transformers and how DGA and AI techniques can be utilized for accurate and comprehensive fault diagnosis. Section 2 provides a literature review to emphasize the contribution of this paper, and Section 3 presents the study's general framework, focusing on the dataset. Section 4 briefly introduces the SVM classifier used in the study, and Section 5 discusses the experimental results and model evaluations. Finally, Section 6 concludes the paper, emphasizing the effectiveness of the proposed SVM-based approach for power transformer fault diagnosis and its potential to improve the reliability and efficiency of the electrical power system.

2 Literature Review

Prognostic and Health Management (PHM) is an integrated technology that detects anomalies, diagnoses occurring failures, and predicts the future health state of a system to estimate its Remaining Useful Life (RUL). It was first proposed by the US Air Force (Timothy, F., et al., 2009; Losik, Ph. D, L., 2012). In recent years, data-driven prognostic studies have focused on determining the relationship between the monitoring data of systems and their health condition.

Li et al. (Li, J., et al., 2016) presented an intelligent method for power transformers' fault diagnosis based on selected gas ratios and Support Vector Machine (SVM). This method outperformed conventional ratio threshold-based methods. Wei et al. (Wei, X., et al., 2006) proposed a classifier using a back-propagation neural network that showed strong learning capability and generalization in terms of higher accuracy of fault diagnosis than classic SVM methods. Guo et al. (Guo, Y. J., et al., 2007) implemented an improved radial basis function neural network for detecting power transformer faults. Fang et al. (Fang, J., et al., 2011) applied principal component analysis as a pre-processing procedure to enhance the data quality and achieved higher recognition accuracy than the former methods based on back-propagation neural networks only.

SVM is a good candidate among different Machine Learning (ML) algorithms to be implemented in the fault diagnosis of power transformers. It overcomes the difficulties of small samples, the curse of dimensionality, local minimum, and overfitting. Although the vast amount of data can help to identify better schemes to detect hidden patterns among equipment faults and data, collecting transformer samples is still complicated, and the increases in the dataset's size pose a significant challenge in data management and computation time. It should be noted that the accuracy results of ML algorithms also depend on the dataset's quality, and most articles' original datasets considered are very few, meaning that data balancing techniques have been used, which exhibits a considerable bias on the algorithm's accuracy (Rao, U. M., et al., 2021).

In recent years, various fault diagnosis techniques have been proposed, including the conventional key gas method, the ratio method, and the graphical representation method. However, the identification of the faulted location by the traditional method is not always an easy task due to the variability of gas data and operational natures. Recently, artificial intelligence techniques have been extensively used to develop more accurate diagnostic tools based on DGA data.

Shintemirov et al. (Shintemirov, A., Tang, W., & Wu, Q. H., December 2008) proposed the Genetic Programming (GP) method for transformer fault detection. The fuzzy logic method is applied to three transformers to diagnose the fault by analyzing the dissolved oil based on fuzzy logic in Ref. (Flores, W. C., Mombello, E. E., Jardini, J. A., Rattá, G., & Corvo, A. M. , July 2011) The back-propagation neural nets described in Ref. (Sun, Y. J., Zhang, S., Miao, C. X., & Li, J. M., March 2007) can identify complicated relationships among dissolved gas contents in transformer oil and corresponding fault types.

Intelligent techniques, such as artificial neural networks and evolutionary algorithms, have gained increasing attention due to their ability to simulate the behavior of living beings and make optimal decisions for solving complex problems. These methods are particularly effective in addressing problems that lack structural information and require a more general approach. One such problem is transformer fault diagnosis, which is challenging due to the high dimensionality and non-linearity of the data. By leveraging these intelligent techniques, it is possible to develop accurate and efficient fault diagnosis models that can effectively analyze the complex data from power transformers. AI techniques can deal with complex and nonlinear problems and implement empirical risk minimization to minimize the error in the training data. At present, SVM has been applied successfully to solve fault classification problems. Overall, intelligent techniques have been widely used in transformer fault diagnosis with convincing results.

3 Methodology

The three major types of power transformer faults which can be reliably identified during a visual inspection are partial discharges, thermal overheating, and arcing (Duval, M., & DePabla, A., 2001). These faults can be classified as electrical or thermal, with partial discharges and arcing resulting from high electrical stress and thermal faults arising from abnormal temperature rises that cause deterioration of the insulation system. The abnormal temperature rises can be due to various factors such as overheating of conductors, short circuits, Foucault's currents causing overheating of windings, loose connections, or insufficient cooling. The fault types and codes addressed in this paper are presented in Table 1.

Table 1. DGA fault types and dataset distribution

Fault Types	Acronyms	Label
Low energy discharge	LD	0
High energy discharge	HD	1
Partial discharge	PD	2
Low thermal faults $T < 300$ °C	LT	3
Medium thermal faults $T = 300$ °C	MT	4
High thermal faults 300 °C $< T < 700$ °C	HT	5

In a power transformer, gas formation occurs due to two primary causes: electrical and thermal stresses. Each fault type affects the oil or paper differently and produces varying amounts of dissolved gas, with the quantities being more or less significant depending on the severity of the fault. The nature and proportions of the gases generated provide valuable information on the type and intensity of the stress, as well as the affected materials. For instance, when an electric arc discharge takes place, significant amounts of hydrogen and acetylene are produced, with minor quantities of methane and ethylene. Acetylene typically accounts for 20% to 70% of the total hydrocarbons, while hydrogen ranges from 30% to 90%. Carbon dioxide and carbon monoxide may also be formed if cellulose is present at the fault site, and in some cases, the oil may carbonize. On the other hand, thermal faults result in the degradation of oil and paper. Overheating of the oil produces ethylene and methane, along with small amounts of hydrogen and ethane. If the fault is serious or involves electrical contacts, traces of acetylene may also form. When thermal faults attack cellulose, large quantities of carbon dioxide and carbon monoxide are produced. Understanding the types and amounts of gases produced in each type of fault is crucial for accurate fault diagnosis and timely maintenance of power transformers.

The conventional gas ratio methods use key gas ratios for fault diagnosis. In this paper, we used Doernenburg's ratio method to extract the fault labels from the actual data that General Electric provides.

The Doernenburg ratio method was one of the earliest techniques developed to analyze dissolved gases in transformers. Initially introduced in 1974, it aimed to evaluate the three primary types of transformer faults. Table 2 (IEEE Guide for the Interpretation of Gases Generated in Mineral OilImmersed Transformers, 2019) summarizes the Doernenburg ratios associated with each fault type and corresponding diagnostic interpretations. This method is typically employed when the concentration of one of the four gases (H2, CH4, C2H4, and C2H2) exceeds twice the limit values (as specified in Table 3), and one of the other gases also surpasses these limit values (Duval, M., & DePabla, A., 2001).

Table 2. Doernenburg ratio method

Fault Types	R1	R2	R3	R4
Thermal Decomposition	> 1.0	< 0.75	< 0.3	> 0.4
Corona	< 0.1	-	< 0.3	> 0.4
Arcing	0.1 – 1.0	> 0.75	> 0.3	< 0.4

$*R_1 = {CH_4}/{H_2}, R_2 = {C_2H_2}/{C_2H_4}$

$*R_3 = {C_2H_2}/{CH_4}, R_4 = {C_2H_6}/{C_2H_2}$

Table 3. Acceptable limits for for the Doernenburg Ratio Method

Key gases	Limit (ppm)
Hydrogen	100
Methane	120
Carbon monoxide	350
Acetylene	1
Ethylene	50
Ethane	65

After obtaining the fault labels, it was discovered that the provided dataset lacked several essential labels. To generate new data, the first step was to extract the distribution of features present in the real-time dataset. Next, the Inverse Cumulative Distribution Function (ICDF) was used to convert Gaussian samples into samples from the specified distributions. By doing this, the generated data retained the same statistical properties as the real dataset, and the missing labels were added to the generated data. The generated data was then combined with the real dataset to form a larger and more comprehensive dataset, which was used for training the SVM algorithm. The proposed approach achieved a probabilistic output for the fault diagnosis of oil-immersed transformers, overcoming the limitations of traditional DGA methods that often provide inaccurate diagnosis results and cannot summarize the fault development rule inductively. The effectiveness of the proposed approach was demonstrated in the case study results, which showed that the generated data significantly improved the performance of the SVM algorithm in predicting PT faults.

3.1 Developing and implementation of SVM

The development of any ML classification algorithm involves two crucial steps, namely building the classifier and applying it for classification purposes. In the first step, the ML classifier learns patterns, dependencies, and features from historical monitoring data, known as the training dataset. This dataset is used to derive classification rules, such as defining Class 0 as the gas amount being less than or equal to a predetermined time window, and other Classes for greater amounts of gases. In the second step, the model parameters are adjusted, and the classifier is utilized to predict the fault types of the validation set. The validation set helps to estimate the generalization error during or after training, and the hyperparameters are updated accordingly. Finally, the constructed classifier is evaluated on the test set to assess the accuracy of the classification rules on new datasets. If the performance metrics indicate that the ML classifier is reliable, it can be used to predict future fault types with high accuracy.

The ultimate goal of developing SVM is to assign discrete labels to dataset based on their features or characteristics. For power transformers' fault prediction, SVM is trained to assign a random variable Y that represents the possible fault occurrence based on the observation of a set of input variables X. In this paper, the SVM is implemented to detect the fault labels of power transformers. The aim is to learn a mapping from inputs X to outputs Y, where Y is an integer value ranging from 1 to C, with C being the number of classes. In this paper, the number of classes (C) is equal to six, making the problem a multiclass classification task. The developed SVM is evaluated on a test set to determine its performance for future fault prediction.

SVM is based on the statistical learning theory and aims to determine the location of decision boundaries that produce the optimal separation of classes (Vapnik, V. N., 1995). SVM is proposed initially for binary class classification. There are two main methods to generalize SVM for multi-class classification: One-versus-rest and One versus-one approach.

3.1.1 One-versus-rest

This approach is also called winner-take-all classification. Suppose the dataset is to be classified into *C* classes. Therefore, *C* binary SVM classifiers may be created where each classifier is trained to distinguish one class from the remaining *C-1* classes. For example, the class one binary classifier is designed to discriminate between class one data vectors and the data vectors of the remaining classes. Other SVM classifiers are constructed in the same manner. Data vectors are classified during the testing or application phase by finding the margin from the linear separating hyperplane. The final output is the class that corresponds to the SVM with the largest margin.

3.1.2 One-versus-one

This method creates SVM classifiers for all possible pairs of classes (Knerret al., 1990; Hastie and Tibshirani, 1998). Therefore, for C classes, there will be binary classifiers. The output from each classifier in the form of a class label is obtained. The class label that occurs the most is assigned to that point in the data vector. In case of a tie, a tie-breaking strategy may be adopted. A common tie-breaking approach is to select one of the class labels that are tied randomly. The number of classifiers created by this method is generally much larger than the previous method - *C(C-1)/2* classifiers in total-. In practice, one-versus-rest classification is usually preferred since the results are mostly similar, but the runtime is significantly less.

3.2 Experimental results

To implement the model, this study utilized a personal computer with Intel Core i7-10510U (1.80GHz) CPU, 16.0 GB memory, and Microsoft Windows 10 operating system. The model was implemented using Python 3.7 and the Keras (Chollet, F., 2015) library with the open-source software library TensorFlow (Abadi, M., et al., 2015) as a backend.

The most common performance metrics to evaluate the multiclass classification are precision, recall, F-Score, true positive rate, false-positive rate, and confusion matrix from the expected and predicted classes' matrix. These metrics are explained as follows.

Precision: The percentage of predicted anomaly records that are actual anomalies.

Recall: The percentage of the total number of correctly classified anomalous records to the total number of positive records. High recall value specifies the class is correctly recognized with a small number of false negatives.

Accuracy: The percentage of correctly classified records over total records. If the false positives and false negatives have similar costs, accuracy will work best.

F-Score: This is the harmonic mean (in percentile) of precision and recall and always has a value near the smaller value of precision or recall. Thus, it provides a more realistic measure of a test's accuracy by using precision and recall. If the costs of false positives and false negatives are very different, F-Score works the best.

Confusion Matrix: As illustrated by Table 4, this matrix explains the performance of classification on a set of test data where the true values are known.

Table 4: Confusion matrix for binary classification

		Predicted Label	
		Class 1	Class 0
Actual Label	Class 1	True Positive	False Negative
	Class 0	False Positive	True Negative

Table 5 shows the best results of the evaluation metrics for SVM classifier. It has achieved a considerable overall accuracy (>95%).

Table 5: Results of the evaluation metrics

Label	Precision	Recall	F1-Score
0	1.00	1.00	1.00
1	0.93	0.99	0.96
2	0.99	1.00	1.00
3	0.86	0.94	0.90
4	1.00	0.92	0.96
5	0.93	0.82	0.87
Accuracy	0.95		
macro avg	0.95	0.95	0.95
weighted avg	0.95	0.95	0.95

Fig.1 represents the results of evaluating the test set as a confusion matrix. In our case, *C* = 6 representing the various classes available in the dataset and it shows the definite positives, definite negatives, projected positives, and projected negatives. The running diagonal represents the values of the correctly predicted instances and is also known as true positive. The values in the off diagonal represents the incorrectly classified instances and manifest as falsely positive or falsely negative. The falsely positive are referred to as the Type I errors while the falsely negative instances are Type II errors. It is observed that the specificity values for all class labels is above 0.81%, which imply fewer false negatives. Interestingly, labels 0,1 and 2 are performing well though they are from minority classes.

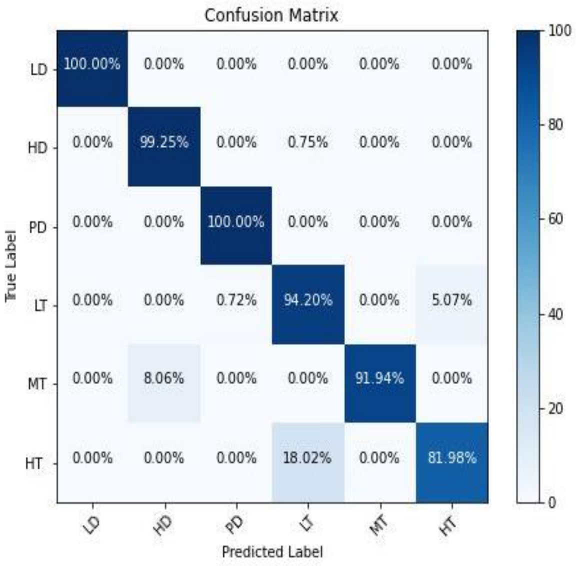

Fig. 1. Confusion matrix for SVM

4 Conclusion

In conclusion, this paper proposes a novel approach for power transformer fault prediction using SVM multi-classification, which effectively recognizes the six leading operating labels of transformers. The proposed approach was validated using General Electric data and transformer online monitoring data from various locations, and it provided promising results that recommend its use in the industry. However, as a future challenge, it is important to further investigate the impact of data pre-processing and data balancing methods on DGA data. In prospective studies, the authors plan to explore the implementation of sequence-to-sequence deep learning classifiers for time series datasets using IoT-enabled intelligent sensors for real-time observation of transformers. These efforts will further improve the accuracy and efficiency of fault diagnosis for power transformers.

5 ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Mitacs and the Mitacs Accelerate program, as well as the Institut de valorisation des données (IVADO) for their generous support in making this research possible. Their funding and resources were invaluable in completing this project and I am grateful for the opportunity to have participated in their programs.

6 References

Abadi, M., et al. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*. Retrieved from https://www.tensorflow.org/

Chollet, F. (2015). *Keras*. Retrieved from https://keras.io/

Duval, M., & DePabla, A. (2001). Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases. *IEEE Electrical Insulation Magazine*, *17*(2), 31-41.

Fang, J., et al. (2011). Application of BP neural network based on principal component analysis. *Transformer*, *48*(1), 47-51.

Flores, W. C., Mombello, E. E., Jardini, J. A., Rattá, G., & Corvo, A. M. . (July 2011). Expert system for the assessment of power transformer insulation condition based on type-2 fuzzy logic systems. *Expert Systems with Applications*, *38*(7), 8119-8127.

Guo, Y. J., et al. (2007). The fault diagnosis of power transformer based on improved RBF neural network. *2007 International Conference on Machine Learning and Cybernetics* (pp. 1111-1114). Hong Kong, China: IEEE.

(2019). *IEEE Guide for the Interpretation of Gases Generated in Mineral OilImmersed Transformers.* IEEE.

Li, J., et al. (2016). Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine. *IEEE Transactions on Dielectrics and Electrical Insulation*, *23*(2), 1198-1206.

Losik, Ph. D, L. (2012). Using the Prognostic Health Management Program on the Air Force Next Generation Reusable Launch Vehicle. *In AIAA SPACE 2012 Conference & Exposition*, (pp. 2012-5197). Pasadena, California.

Rao, U. M., et al. (2021). Identification and application of machine learning algorithms for transformer dissolved gas analysis. *IEEE Transactions on Dielectrics and Electrical Insulation*, *28*(5), 1828-1835.

Shintemirov, A., Tang, W., & Wu, Q. H. (December 2008). Power transformer fault classification based on dissolved gas analysis by implementing bootstrap and genetic programming. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, *39*(1), 69-79.

Sun, H. C., Huang, Y. C., & et al. (2012). A review of dissolved gas analysis in power transformers. *Energy Procedia*, *14*, 1220-1225.

Sun, Y. J., Zhang, S., Miao, C. X., & Li, J. M. (March 2007). Improved BP neural network for transformer fault diagnosis. *Journal of China University of Mining and Technology*, *17*(1), 138-142.

Timothy, F., et al. (2009). F-35 joint strike fighter structural prognosis and health management an overview. *Proceedings of the 25th Symposium of the International Committee on Aeronautical Fatigue*, (pp. 1215-1219). Rotterdam, The Netherlands.

Vapnik, V. N. (1995). *The nature of statistical learning theory.* New York, NY: Springer. doi:https://doi.org/10.1007/978-1-4757-2440-0

Wei, X., et al. (2006). Power transformer fault integrated diagnosis based on improved PSO-BP neural networks and DS evidential reasoning. *Automation of Electric Power Systems*, *30*(7), 46-49.

Answers to the Comments
Manuscript Number: 7057
Title: Power Transformer Fault Prediction based on Support Vector Machine

We would like to thank Co-présidents du comité scientifique de CIGI QUALITA MOSIM 2023 and the Reviewers for their feedback and valuable recommendations to improve the paper.

Reviewer #1

Dans le r´esum´e, il est dit : “Real-time data from General Electric is used to implement the Support Vector Machine, but due to inadequate data, Gaussian simulation is employed to generate new data”. Ceci est confirm´e Page 3, o`u il est dit “A Gaussian simulation technique was utilized to generate additional data around the real-time dataset”. Il me semble n´ecessaire d’apporter des pr´ecisions sur les donn´ees issues de l’´etude de cas. Au moins leurs carac ´eristiques et pourquoi elles sont inadapt´ees et insuffisantes. De plus, il faudrait des pr´ecisions sur comment sont g´en´er´ees les nouvelles donn´ees. Dire quelle est la proportion de donn´ees r´eelles par rapport aux donn´ees g´en´er´ees. La fa¸con de g´en´erer des donn´ees peut introduire des biais (favorables ou d´efavorables) impactant les r´esultats de la m´ethode. Ceci m´erite d’ˆetre discut´e plus en profondeur.
Remarques mineures
• Page 3, Table 3 : for for dans la l´egende

Reviewer #2

Interesting paper on fault detection and prediction for power transformers. More detailed information should be given about the data used in the SVM implementation (type of data, volume, distribution between learning and testing). Especially in the abstract it is written that "due to inadequate data, Gaussian simulation is employed to generate new data". What does it mean? What is the justification of this hypothesis? Is the study developed without real data?
To help improve the paper:
- introduce the acronym DGA earlier (it is employed from the beginning but defined on the second column of page 2)
- the first sentence of the last paragraph of section 2 is unclear. Please rewrite it.

Answer

Thank you very much for your positive feedback. The suggestions and comments have been closely followed, and revisions have been made accordingly. In the revised paper, we have marked in blue all the changes.