

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Marcos Vinícius Gonçalves da Rocha

**Power Transformers Thermal Modeling using an Enhanced Set-Membership
Multivariable Gaussian Evolving Fuzzy System**

Juiz de Fora

2020

Marcos Vinícius Gonçalves da Rocha

**Power Transformers Thermal Modeling using an Enhanced Set-Membership
Multivariable Gaussian Evolving Fuzzy System**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Sistemas Eletrônicos

Orientador: Prof. Dr. Eduardo Pestana de Aguiar

Juiz de Fora

2020

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Rocha, Marcos Vinícius Gonçalves da.

Power Transformers Thermal Modeling using an Enhanced Set-Membership Multivariable Gaussian Evolving Fuzzy System / Marcos Vinícius Gonçalves da Rocha. – 2020.

47 f. : il.

Orientador: Eduardo Pestana de Aguiar

Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Faculdade de Engenharia. Programa de Pós-Graduação em Engenharia Elétrica, 2020.

1. Enhanced Set-Membership. 2. Evolving Fuzzy Systems. 3. Power Transformers. 4. Thermal Modeling. I. Aguiar, Eduardo Pestana, orient. II. Título.

Marcos Vinícius Gonçalves da Rocha

**Power Transformers Thermal Modeling using an Enhanced Set-Membership
Multivariable Gaussian Evolving Fuzzy System**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Sistemas Eletrônicos

Aprovada em 29 de setembro de 2020

BANCA EXAMINADORA

Eduardo Pestana de Aguiar

Prof. Dr. Eduardo Pestana de Aguiar - Orientador
Universidade Federal de Juiz de Fora

Frederico Gadelha Guimarães

Prof. Dr. Frederico Gadelha Guimarães
Universidade Federal de Minas Gerais

André Luís Marques Marcato

Prof. Dr. André Luís Marques Marcato
Universidade Federal de Juiz de Fora

To my parents, Antonio and Geni.

ABSTRACT

Knowledge of temperature distribution in power transformers is essential for the management of electrical distribution systems. Monitoring the hot-spot temperature of a power transformer can extend its lifetime. In this work, we present two new models based on Set-Membership filtering: the Set-Membership evolving Multivariable Gaussian and the Enhanced Set-Membership evolving Multivariable Gaussian. Both approaches are acting by adjusting the learning rate in the evolving fuzzy modeling system. To evaluate its performance were applied synthetic data sets, as benchmarks, and data for thermal modeling of real power transformers, under two load conditions: with and without an overload condition. The obtained results are compared with the performance of the original evolving Multivariable Gaussian and with other classical models suggested in the literature. Both proposed models obtained lower errors and presenting a competitive number of rules, suggesting that the models are flexible and efficient approaches in these scenarios.

Keywords: Enhanced Set-Membership. Evolving Fuzzy Systems. Power Transformers. Thermal Modeling.

RESUMO

O conhecimento da distribuição de temperatura em transformadores de potência é essencial para o gerenciamento de sistemas de distribuição elétrica. O monitoramento da temperatura do ponto quente de um transformador de energia pode estender sua vida útil. Neste trabalho, apresentamos dois novos modelos baseados na filtragem *Set-Membership*: o *Set-Membership* evolutivo Gaussiano Multivariado e o *Enhanced Set-Membership* evolutivo Gaussiano Multivariado. Ambas as abordagens agem ajustando a taxa de aprendizagem no sistema de modelagem fuzzy evolutivo. Para avaliar seu desempenho foram aplicados conjuntos de dados sintéticos, como *benchmarks*, e dados para modelagem térmica de transformadores de potência reais, sob duas condições de carga: com e sem sobrecarga. Os resultados obtidos são comparados com o desempenho do modelo evolutivo Gaussiano Multivariado original e com outros modelos clássicos sugeridos na literatura. Ambos os modelos propostos obtiveram erros menores e apresentam número competitivo de regras, sugerindo que os modelos são abordagens flexíveis e eficientes nestes cenários.

Palavras-chave: *Enhanced Set-Membership*. Sistemas Fuzzy Evolutivos. Transformadores de Potência. Modelagem Térmica.

LIST OF FIGURES

Figure 1 – Mechanism of learning of eMG model	21
Figure 2 – Comparison of Mackey-Glass time series forecasting	28
Figure 3 – Comparison of Mackey-Glass time series forecasting, with $\tau = 10$	30
Figure 4 – Comparison of Mackey-Glass time series forecasting, with $\tau = 20$	30
Figure 5 – Comparison of Mackey-Glass time series forecasting, with $\tau = 70$	31
Figure 6 – Comparison of nonlinear system identification	33
Figure 7 – Comparison of high-dimensional system identification	35
Figure 8 – Sensor’s location in the experimental transformer	37
Figure 9 – Learning data set	38
Figure 10 – Data set 1: without overload condition	38
Figure 11 – Data set 2: with overload condition	39
Figure 12 – Hot-spot estimation, operation without overload condition	41
Figure 13 – Hot-spot estimation, operation with overload condition	41

LIST OF TABLES

Table 1 – Performance of Mackey-Glass time series forecasting approaches .	28
Table 2 – Results of the MGN test, Mackey-Glass time series forecasting .	29
Table 3 – Performance of Mackey-Glass time series forecasting approaches .	29
Table 4 – Performance of nonlinear system identification methods	32
Table 5 – Results of the MGN test, nonlinear system identification	33
Table 6 – Performance of high-dimensional system identification methods .	34
Table 7 – Results of the MGN test, high-dimensional system identification	35
Table 8 – Characteristics of the experimental power transformer	36
Table 9 – Performance of thermal modeling of real power transformers, operation without overload condition	40
Table 10 – Performance of thermal modeling of real power transformers, operation with overload condition	40
Table 11 – Results of the MGN test, operation without overload condition .	42
Table 12 – Results of the MGN test, operation with overload condition . . .	42

SUMMARY

1	INTRODUCTION	9
1.1	MOTIVATION	10
1.2	JUSTIFICATION	11
1.3	AIMS AND OBJECTIVES	11
1.4	STRUCTURE OF THE WORK	12
2	PROBLEM FORMULATION	13
3	THE ENHANCED SET-MEMBERSHIP EVOLVING MUL-	
	TIVARIABLE GAUSSIAN	15
3.1	EVOLVING FUZZY SYSTEMS	15
3.2	GAUSSIAN PARTICIPATORY EVOLVING CLUSTERING	16
3.3	THE EVOLVING MULTIVARIABLE GAUSSIAN MODEL	20
3.4	THE SET-MEMBERSHIP EVOLVING MULTIVARIABLE GAUSSIAN MODEL	23
3.5	THE ENHANCED SET-MEMBERSHIP EVOLVING MULTIVARIA- BLE GAUSSIAN MODEL	25
4	EXPERIMENTAL RESULTS	26
4.1	SYNTHETIC DATA SETS	27
4.1.1	Mackey-Glass Time Series Forecasting	27
4.1.1.1	Comparative Mackey-Glass Time Series Forecasting for Different Values of τ	29
4.1.2	Nonlinear System Identification	31
4.1.3	High-Dimensional System Identification	33
4.2	THERMAL MODELING OF REAL POWER TRANSFORMERS	35
5	CONCLUSION	43
5.1	FUTURE WORKS	43
	REFERENCES	44
	APPENDIX A – Publications	47

1 INTRODUCTION

Power transformers are essential equipment in electrical distribution systems. These equipment are used in the generation, through transmission and distribution to delivery to the end customer, and require cautious operation and management. Due to their high price and critical function in the power system, they need more diligence in their operation in order to ensure safety and efficiency.

Regarding the use of power transformers, maintenance and reliability are essential issues. An important factor concerns the working temperature (i.e., the hot-spot temperature), which is directly related to the aging of the insulation and the service life of the power transformers. Power transformer life is determined by its insulation aging rate, moreover, monitoring the hot-spot temperature is a crucial factor in transformer life expectancy (JAFARBOLAND, M., 2019) and to develop overload protection systems (GOMIDE, F., 2007). Unexpected failures can seriously affect the power system, causing economic and social losses (GOMIDE, F., 2008).

Adequate knowledge of the temperature distribution inside a transformer, as well as all thermal aspects, are important to management to maximize the service life and the power rating and minimize the operation cost (LEBRETON, R., 2018). The most common model used in practice for the prediction of the hot-spot temperature of power transformers is based on the IEEE Standard C57.91-2011, which uses transient heating equations and specific thermal characteristics and parameters of power transformers (IEEE, 2012). However, this modeling can become an inaccurate method due to its simplifications. Therefore, it is necessary to employ more advanced techniques to optimize the use of the power transformer capacity and increase its useful life, maintaining its functionality and safety (GOMIDE, F., 2007).

There is a considerable amount of research on monitoring power transformers in the literature. For instance, (VILLACCI, D., 1996) presents a method to monitor and estimate the heating of the transformers based on neural identification techniques. In (VACCARO, A., 2000), the authors intend to predict the hot-spot temperature of the winding of a power transformer in the presence of overload conditions using a radial basis function network (RBFN). Reference (IPPOLITO, L., 2004) deals with a fuzzy Takagi-Sugeno-Kang (TSK) model capable of reproducing the temperature distribution of power transformers to implement an overload protection system and in (BIRATTARI, M., 2005) discusses an architecture to integrate the modeling of physical knowledge with machine learning techniques, focusing on forecasting of the hot-spot temperature of a transformer. Moreover, (GOMIDE, F., 2007) presents a recurrent neurofuzzy hybrids network to model the thermal condition of power transformers and (GOMIDE, F., 2008) uses the concept of Participatory Learning (PL) to train a hybrid neurofuzzy network, to

the same purpose.

1.1 MOTIVATION

Thus, motivated by the need for adaptable systems, evolving intelligent systems (FILEV, D., 2004) have received great attention in recent years, being widely applied in problems of modeling, control, forecasting, classification, and processing of data in dynamic and non-stationary environments (BALLINI, R., 2018).

A wide variety of these models can be found in the literature, with the evolving Takagi-Sugeno (eTS) model as a precursor (FILEV, D., 2004). This model updates the system recursively, adding new rules, or updating existing rules. The antecedents are determined through a subtractive grouping (CHIU, S., 1994) based on the notion of the potential function and the parameters of the consequents are updated through Recursive Least Squares (YOUNG, P., 2012).

However, the aforementioned models, as noted by (GOMIDE, F., 2011), are not robust in the presence of imprecise or noisy data, i.e., this kind of data are also incorporated by the learning process which can result in a model that does not truly represent the real system.

To overcome this problem, in (LIMA, E., 2010) was introduced an evolving system based on the concept of Participatory Learning proposed by (YAGER, R., 1990). This approach assumes that learning and beliefs about an environment depend on what the system already knows about the environment (YAGER, R., 2004).

In this way, the model proposed in (LIMA, E., 2010), called evolving Participatory Learning (ePL), uses an unsupervised dynamic fuzzy clustering algorithm that captures many of the salient features of human learning where the update of the rule base is a function of existing rules and new information received. A similar process occurs with human learning, which is amplified if there is already prior knowledge on a given subject. This means that a new data sample that are very distant from what has already been learned by the system tend to be discarded or have their effect smoothed making the model more robust to outliers.

More recently a new evolving fuzzy model, named evolving Multivariable Gaussian (eMG) (GOMIDE, F., 2011), was proposed as an evolution of the ePL model. In the eMG each fuzzy cluster is represented by a multivariable Gaussian membership function instead of a single variable Gaussian membership function as in ePL. Also, clusters in the eMG approach are estimated using a normalized distance measure (similar to the Mahalanobis distance) and trigger ellipsoidal clusters, whose axes are not necessarily parallel to the axes of the input variables thus preserving the information about interactions between input variables.

Due to its evolving structure that makes the model can adapt quickly to changes in the power transformer's operating characteristics, such as insulation aging, environmental changes, among others, the eMG model was applied to the hot-spot temperature modeling problem achieving extremely promising results (BOAVENTURA, W., 2012).

In another branch of studies, the Set-Membership (SM) adaptive filtering theory (HAYKIN, S., 1996) was combined with a fuzzy system trained by Steepest Descent method in order to reduce the computational complexity of the model and to increase the convergence speed during the training phase (RIBEIRO, M. V., 2017). The results presented by this study suggest that the combination of the SM framework with an evolving system can lead to models of less computational complexity and with the ability to adapt to changes in the environment over time. This characteristic makes this strategy an interesting candidate for the proposed problem, especially for real-time applications.

1.2 JUSTIFICATION

In this sense, this work suggests a synergism between the SM framework and its enhanced version, named Enhance Set-Membership (ESM), and the eMG evolving fuzzy model to create a power transformer's thermal model with low computational complexity, which can handle inaccurate or noisy data, which preserves the iterations between the input variables and which can adapt to changes in the dynamics of the transformers imposed by the aging of equipment, changes in the load profiles or environmental changes.

1.3 AIMS AND OBJECTIVES

The objectives of this work are listed as follows:

- To propose two evolving fuzzy models, called Set-Membership evolving Multivariable Gaussian (SM-eMG) and Enhanced Set-Membership evolving Multivariable Gaussian (ESM-eMG), based on the Set-Membership filtering (RIBEIRO, M. V., 2017, LAMARE, R. C., 2011, JIANG, T., 2016) to adjust the learning rate in the evolving Multivariable Gaussian (eMG) algorithm.
- To adopt synthetic time series forecasting and nonlinear system identification problems, and measured data set of power transformers for thermal modeling, which has two load conditions considered: with and without an overload condition, and we evaluate the results in terms of error metrics and the number of final rules.
- To compare the performance with other approaches suggested in the literature: evolving Participatory Learning with Kernel Recursive Least Squares (ePL-KRLS) (BALLINI, R., 2018), Set-Membership evolving Participatory Learning with Kernel Recursive Least Squares (SM-ePL-KRLS) and Enhanced Set-Membership evolving

Participatory Learning with Kernel Recursive Least Squares (ESM-ePL-KRLS) (ALVES, K. S., 2020), IEEE Deterministic Model (IEEE, 2012), Adaptive Neurofuzzy Inference System (ANFIS) (JANG, J., 1993) and Multilayer Perceptron (MLP) (STORK, D. G., 2012).

1.4 STRUCTURE OF THE WORK

This work is organized as follows: Chapter 2 deals with the problem formulation. Chapter 3 details the proposed models. Chapter 4 presents and discusses the results of computer simulations. Finally, Chapter 5 states the main conclusions.

2 PROBLEM FORMULATION

The problem of thermal modeling of power transformers consists of determining the hot-spot temperature from other variables that can be obtained more easily, such as the load current, the temperature on the surface of the transformer's oil (top-oil temperature) and the environment temperature.

Currently, the classic way to deal with this problem is to apply the deterministic model proposed in the IEEE Standard C57.91-2011 (IEEE, 2012). In this model, the load curves and operating conditions (especially the operating temperature) of the transformer are used to compute a series of differential equations for which conservative safety rates are generally applied (VACCARO, A., 2000). To summarize this deterministic model, as presented in (GOMIDE, F., 2007), the following variables and parameters are considered:

Variables (functions of time):

Θ_A	=	environment temperature, °C.
Θ_{TO}	=	top oil temperature, °C.
Θ_H	=	hot-spot temperature, °C.
$\Delta\Theta_H$	=	hot-spot rise above top oil temperature, °C.
$\Delta\Theta_{H,U}$	=	ultimate hot-spot temperature rise over top oil (for a given load current), °C.
$\Delta\Theta_{TO,U}$	=	ultimate top oil temperature rise over environment (for a given load current), °C.
I_L	=	load current, <i>per unit</i> .

Parameters (constants):

$\Theta_{A,R}$	=	rated environment temperature, °C.
$\Theta_{H,R}$	=	rated hot-spot winding temperature, °C.
$\Delta\Theta_{TO,R}$	=	rated top oil temperature rise over environment, °C.
$\Delta\Theta_{H,R}$	=	rated hot-spot temperature rise over top oil, °C.
τ_{TO}	=	top oil rise time constant, <i>hours</i> .
τ_H	=	hot-spot rise time constant, <i>hours</i> .
R	=	ratio of load loss at rated-load to no-load loss at applicable tap position, <i>dimensionless</i> .
m, q	=	empirically derived values, depend on the cooling method, <i>dimensionless</i> .

The first step of the hot-spot calculation by the IEEE deterministic model is to calculate the ultimate top oil temperature rise over the environment temperature ($\Delta\Theta_{TO,U}$),

as follows:

$$\Delta\Theta_{TO,U} = \Delta\Theta_{TO,R} \left[\frac{I_L^2 R + 1}{R + 1} \right]^q \quad (2.1)$$

The next step is to calculate the increment in the top oil temperature (Θ_{TO}) that is found from Equation (2.1) and the environment temperature (Θ_A) through the following differential equation:

$$\tau_{TO} \frac{d\Theta_{TO}}{dt} = [\Delta\Theta_{TO,U} + \Theta_A] - \Theta_{TO} \quad (2.2)$$

The values for the τ_{TO} parameter or a method for estimate them are described in detail in (IEEE, 2012). In sequence, the ultimate hot-spot rise over top oil ($\Delta\Theta_{H,U}$) is computed as follows:

$$\Delta\Theta_{H,U} = \Delta\Theta_{H,R} I_L^{2m} \quad (2.3)$$

To obtain the value of $\Delta\Theta_{H,R}$ is necessary to perform thermal tests on the equipment. Alternatively, the IEEE Standard C57.91-2011 presents another method to estimate this value.

Now, the increment in hot-spot rise above top oil temperature ($\Delta\Theta_H$) can be calculated using the value of $\Delta\Theta_{H,U}$ using the following differential equation:

$$\tau_H \frac{d\Delta\Theta_H}{dt} = \Delta\Theta_{H,U} - \Delta\Theta_H \quad (2.4)$$

Finally, the hot-spot temperature is calculated as a function of Θ_{TO} and $\Delta\Theta_H$ as follows:

$$\Theta_H = \Theta_{TO} + \Delta\Theta_H. \quad (2.5)$$

In the previous calculations, several simplification assumptions are made to ensure the safe operation of the transformer. These assumptions establish that the calculated operational rate of the power transformer can be 20% to 30% lower than its nominal capacity (VILLACCI, D., 1996). Thus, the use of the deterministic model in the operational planning of the electrical system implies the underutilization of the installed equipment, which can leads to financial losses for the energy companies (GOMIDE, F., 2008).

In this sense, the main objective of the models proposed in this work is to present an algorithm capable of predicting the temperature of the hot-spot with high accuracy. An advantage of these models is that they can adjust their structure to changes in the dynamics of the problem, in order to increase the operating margin of the system mainly in the presence of overload conditions (VAIDYANATHAN, S., 2015).

3 THE ENHANCED SET-MEMBERSHIP EVOLVING MULTIVARIABLE GAUSSIAN

The thermal modeling approach suggested in this work is based on the concepts of evolving fuzzy systems (FILEV, D., 2004), multivariable Gaussian participatory clustering (GOMIDE, F., 2011), Set-Membership (RIBEIRO, M. V., 2017, LAMARE, R. C., 2011) and Enhanced Set-Membership filtering (ALVES, K. S., 2020). For a better understanding, in this Chapter, these concepts are briefly reviewed and then the proposed model is introduced.

3.1 EVOLVING FUZZY SYSTEMS

Evolving fuzzy systems (eFS) are highly adaptive fuzzy systems whose models are self-developed from a stream of data through recursive methods of machine learning (FILEV, D., 2006). In general, these systems propose strategies to continuously evolve the antecedents/consequents parameters of a rule-based system by the use of recursive clustering algorithms.

In particular, evolving Takagi-Sugeno (eTS) based systems (FILEV, D., 2004) demonstrated to be highly applicable to nonlinear dynamic systems modeling in dynamically evolving environments, such as the modeling of the temperature distribution of power transformers problem (BOAVENTURA, W., 2012). These models are based on a first-order Takagi-Sugeno functional rule-based model whose fuzzy rules are as follows:

$$\mathcal{R}_i : \mathbf{IF} \ x \text{ is } \Gamma_i \ \mathbf{THEN} \ y_i = \gamma_{io} + \sum_{j=1}^p \gamma_{ij}x_j \quad (3.1)$$

where \mathcal{R}_i is the i -th fuzzy rule, for $i = 1, \dots, c^k$, c^k is the number of fuzzy rules at step k , $x \in [0, 1]^p$ is the input data, y_i is the output of the i -th rule, Γ_i is the vector of antecedent fuzzy rules, γ_{io} and γ_{ij} are the parameters of the consequent. The antecedent fuzzy sets in Γ_i are Gaussian membership functions:

$$\mu = e^{-r\|x^k - v_i\|} \quad (3.2)$$

where r is a positive constant that defines the zone of influence of the i -th local model and v_i is the respective cluster center or focal point, $i = 1, \dots, c^k$.

To find the cluster centers v_i , the eTS model requires online clustering procedure. For the original eTS this clustering procedure is a form of subtractive clustering, a variation of the Filev and Yager mountain clustering approach (FILEV, D., 1994). In this approach, a Cauchy function is used to measure the potential of a point to become a cluster center. If the potential of new data is higher than the potential of the current cluster centers, then the new data become a new center and a new rule is created. If the potential of

new data is higher than the potential of the current centers, but it is close to an existing center, then the new data replace the existing center. This makes the eTS model capable of continuously develop its structure and functionality through an online self-organization.

More recently, a variation of eTS model was introduced with the intent of being more robust to erroneous or anomalous observations in the input data. This new model, called evolving Multivariable Gaussian (eMG), implements a different clustering algorithm, based on the concept of Participatory Learning (YAGER, R., 1990), instead of the subtractive clustering of the original eTS model. In addition, for eMG, the antecedent fuzzy sets in Γ_i are multivariable Gaussian membership functions instead of single-variable Gaussian membership functions as presented in Equation (3.2). In the next Sections, these concepts are briefly discussed.

3.2 GAUSSIAN PARTICIPATORY EVOLVING CLUSTERING

The Gaussian Participatory Evolving Clustering (GPEC), proposed in (GOMIDE, F., 2011), is an evolving fuzzy clustering procedure in which each cluster is characterized as a multivariable Gaussian membership function of the form:

$$H(x) = \exp \left[-\frac{1}{2}(x - v)\Sigma^{-1}(x - v)^T \right] \quad (3.3)$$

where x and v are $1 \times m$ vectors containing the input and the cluster center, respectively, and Σ is a $m \times m$ symmetric, positive-definite matrix containing the dispersion of the cluster and represents the spread of $H(x)$.

If the input variables interact (e.g, if the problem uses lagged values of the input and/or output as inputs), the use of the multivariable Gaussian membership functions defined in Equation (3.3), avoid the information loss induced by the computation of the fuzzy relation through an aggregation operator (e.g., a t-norm) and the single-variable Gaussian membership functions defined in Equation (3.2) (GOMIDE, F., 2011). As will be seen later, this is exactly the case with thermal modeling of power transformers problem.

As mentioned earlier, the GPEC also implements a new kind of fuzzy clustering in which the cluster structure (number, shape and centers of clusters) is recursively updated at each step by an algorithm based on the Participatory Learning paradigm (YAGER, R., 1990). In this learning process, the influence of a new observation in the cluster structure depends on the knowledge previously learned by the model. The strength of this influence is calculated using a compatibility measure $\rho_i^k \in [0, 1]$, which indicates the degree of compatibility of an observation with the current cluster structure and an arousal index $a_i^k \in [0, 1]$, which defines whether the current structure should be changed. In addition, T_ρ and T_a , are, respectively, used as thresholds for both parameters ρ_i^k and a_i^k . Based on these thresholds, at each step, the learning process may create a new cluster, modify the

parameters of an existing one, or merge two similar clusters.

As detailed in (GOMIDE, F., 2011), if, at each step, the compatibility measure of the current observation is less than the threshold for all clusters, i.e., $\rho_i^k < T_\rho \forall i = 1, \dots, c_k$, and the arousal index of the cluster with the greatest compatibility is greater than the threshold, i.e., $a_i^k > T_a$ for $i = \arg \max_i \rho_i^k$, then a new cluster is created. Otherwise, the cluster center with the highest compatibility is adjusted as follows:

$$v_i^{k+1} = v_i^k + G_i^k (x^k - v_i^k) \quad (3.4)$$

where x^k is the input at step k , v_i^k is the cluster center, for $i = 1, \dots, c^k$, c^k is the number of clusters at step k , and G_i^k is defined as follows:

$$G_i^k = \alpha (\rho_i^k)^{1-a_i^k} \quad (3.5)$$

where $\alpha \in [0, 1]$ is the learning rate. This is the parameter that will be updated with the SM or ESM algorithm, in order to improve the model learning.

According to (YAGER, R., 1990) the compatibility ρ_i^k is a function that measures the compatibility between the current belief of the model, represented by each cluster center, and the current observation.

The compatibility measure ρ_i^k suggested in (GOMIDE, F., 2011) uses the squared value of the normalized distance between the new observation and cluster centers (*M-distance*), given by:

$$M(x^k, v_i^k) = (x^k - v_i^k)(\Sigma_i^k)^{-1}(x^k - v_i^k)^T \quad (3.6)$$

To compute the *M-distance*, the dispersion matrix of each cluster Σ_i^k must be estimated at each step. The recursive estimation of the dispersion matrix proceeds as follows:

$$\Sigma_i^{k+1} = (1 - G_i^k)(\Sigma_i^k - G_i^k(x^k - v_i^k)(x^k - v_i^k)^T) \quad (3.7)$$

The compatibility measure at each step k is given by:

$$\rho_i^k = F(x^k, v_i^k) = \exp \left[-\frac{1}{2} M(x^k, v_i^k) \right] \quad (3.8)$$

The function $F(x^k, v_i^k) \in [0, 1]$ is such that it should approach zero as observations become contradictory with the current belief, i.e., the cluster centers, and approach one as the observations become in complete agreement with the current belief.

To find a threshold value for the compatibility measure, we assume that the values $M(x^k, v_i^k)$ can be modeled by a chi-square distribution. Thus, given a significance level λ , the threshold can be computed as follows:

$$T_\rho = \exp \left[-\frac{1}{2} \chi_{m,\lambda}^2 \right] \quad (3.9)$$

where $\chi_{m,\lambda}^2$ is the λ upper unilateral confidence interval of a chi-square distribution with m degrees of freedom, where m is the number of inputs.

The arousal index is defined by the probability of observing less than nv violations of the compatibility threshold on a sequence of w observations.

To compute the arousal index for each observation, a related occurrence value o_i^k is found using the following expression:

$$o_i^k = \begin{cases} 0, & \text{for } M(x^k, v_i^k) < \chi_{n,\lambda}^2 \\ 1, & \text{otherwise} \end{cases} \quad (3.10)$$

Given a sequence assembled by the last w observations, the number of threshold violations nv_i^k is as follows:

$$nv_i^k = \begin{cases} \sum_{j=0}^{w-1} o_i^{k-j}, & k > w \\ 0, & \text{otherwise} \end{cases} \quad (3.11)$$

The discrete probability distribution of observing nv threshold violations on a window of size w is $p(NV_i^k = nv)$, with NV_i^k assuming the values $nv = 0, 1, \dots, w$. The binomial distribution gives the probability of observing nv threshold violations in a sequence of w observations.

$$p(NV_i^k = nv) = \begin{cases} \binom{w}{nv} \lambda^{nv} (1-\lambda)^{w-nv}, & nv = 0, \dots, w \\ 0, & \text{otherwise} \end{cases} \quad (3.12)$$

The arousal index is defined as the value of the cumulative probability of NV_i^k , i.e., $a_i^k = p(NV_i^k < nv)$. The multivariable Gaussian evolving clustering procedure is shown in Algorithm 1.

Algorithm 1 MGEC

```

1: Input:  $k, x^k, v, \Sigma, \lambda, w, \Sigma_{init}, c$ 
2: Output:  $v, \Sigma, cluster\_created, cluster\_merged, c, idx^k$ 
3: if  $k == 1$  then
4:   Initialize first cluster:
5:    $v_1 = x^1; \Sigma_1 = \Sigma_{init}$ 
6:    $c = 1$ 
7: end if
8: Compute  $\rho_i$  and  $a_i$  for all clusters
9: for  $i = 1, c$  do
10:   $M(x^k, v_i) = (x^k - v_i)(\Sigma_i)^{-1}(x^k - v_i)^T$ 
11:   $\rho_i = \exp\left[-\frac{1}{2}M(x^k, v_i)\right]$ 
12:  if  $\rho_i < T_\rho$  then
13:     $o_i^k = 1$ 
14:  else
15:     $o_i^k = 0$ 
16:  end if
17:  if  $k > w$  then
18:     $nv_i = \sum_{l=0}^{w-1} o_i^{k-l}$ 
19:     $a_i = p(NV_i^k < nv_i)$ 
20:  else
21:     $a_i = 0$ 
22:  end if
23: end for
24:  $idx = \arg \max_i \rho_i$ 
25: if  $\rho_i < T_\rho \forall i$  and  $a_{idx}^k > T_a$  then
26:   Create new cluster:
27:    $c = c + 1; v_c = x^k$ 
28:    $\Sigma_c = \Sigma_{init}; idx = c$ 
29:    $cluster\_created = c$ 
30: else
31:   Update an existing cluster:
32:    $G_{idx} = \alpha (\rho_{idx})^{1-a_{idx}}$ 
33:    $v_{idx} = v_{idx^k} + G_{idx} (x^k - v_{idx})$ 
34:    $\Sigma_{idx} = (1 - G_{idx})(\Sigma_{idx} - G_{idx}(x^k - v_{idx})(x^k - v_{idx})^T)$ 
35:    $\backslash\backslash$  SM or ESM is called here
36: end if
37: Check for redundant clusters
38: for  $i = 1, c$  do
39:   if  $\rho_{idx}(v_{idx}, v_i) > T_\rho$  or  $\rho_j(v_j, v_{idx}) > T_\rho$  then
40:     Merge two redundant clusters:
41:      $v_{idx} = \text{mean}(v_j, v_{idx}); \Sigma_{idx} = \Sigma_{init}$ 
42:      $c = c + 1; cluster\_merged = [idx \ j]$ 
43:   end if
44: end for

```

It is worth mentioning that, according to (GOMIDE, F., 2011), the clustering algorithm uses only three parameters:

- the learning rate α ;
- the window size w , which is used in the arousal mechanism;
- the confidence level λ , which is used to compute T_ρ and T_a .

The confidence level, λ , is chosen based on the value of w , suggested in (GOMIDE, F., 2011), according to the Equation (3.13). The window size, w , is chosen considering the number of consecutive observations in the arousal index calculation.

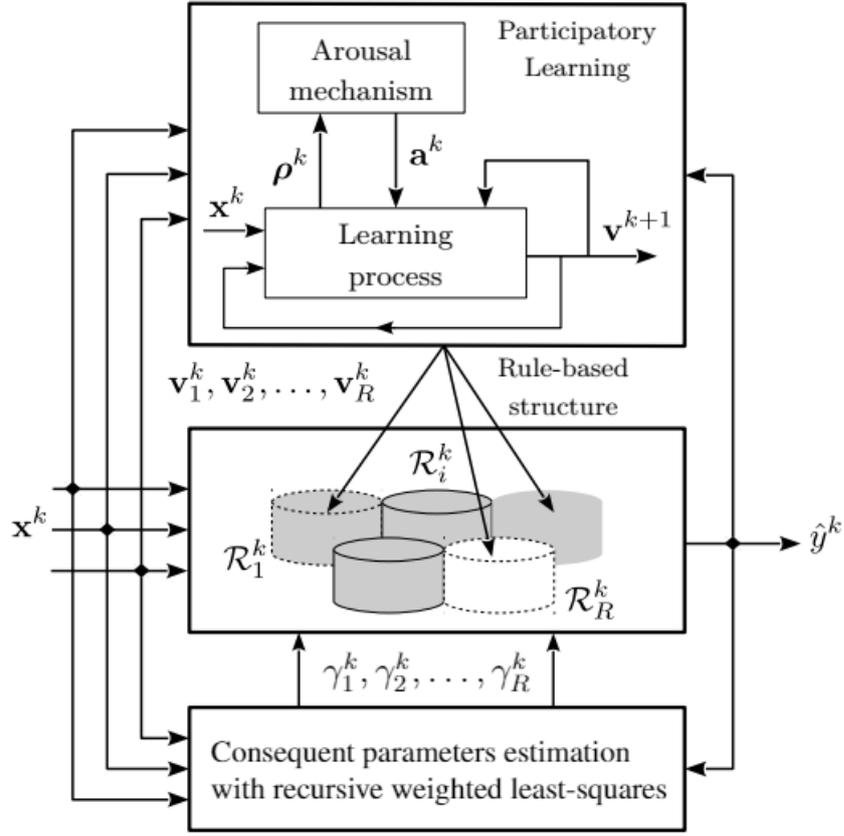
$$\lambda = \begin{cases} 0.01, & \text{if } w \geq 100 \\ 0.05, & \text{if } 20 \leq w < 100 \\ 0.1, & \text{if } 10 \leq w < 20 \end{cases} \quad (3.13)$$

For more details about GPEC algorithm see (GOMIDE, F., 2011).

3.3 THE EVOLVING MULTIVARIABLE GAUSSIAN MODEL

As mentioned before, the evolving Multivariable Gaussian (eMG) model is an evolving functional Takagi-Sugeno (eTS) model in which the rule base is continuously updated through the GPEC algorithm presented in Section 3.2 and the membership functions in the antecedent of the rules are the Multivariable Gaussian functions presented in Equation (3.3) that are adopted to preserve information between the interactions of the input variables. In addition, the eMG model uses the recursive weighted least squares algorithm to estimate the consequent parameters of the rules. As with (GOMIDE, F., 2011), the number of eMG rules is the same as the number of clusters determined by the clustering algorithm at each step. The learning structure of eMG is represented in Figure 1.

Figure 1: Mechanism of learning of eMG model



Source: Adapted from BALLINI, R. (2018), GOMIDE, F. (2011).

The model consists of a set of functional fuzzy rules, as follows:

$$\mathcal{R}_i : \mathbf{IF} \ x^k \text{ is } H_i \ \mathbf{THEN} \ y_i^k = \gamma_{io}^k + \sum_{j=1}^m \gamma_{ij}^k x_j^k \quad (3.14)$$

where x^k is a $1 \times m$ input vector, H_i is a fuzzy set with multivariable Gaussian membership function, \mathcal{R}_i is the i -th fuzzy rule, for $i = 1, \dots, c^k$, c^k is the number of rules, and γ_{io}^k and γ_{ij}^k are the parameters of the consequent at step k (GOMIDE, F., 2011).

The model output is the weighted average of the outputs of each rule (GOMIDE, F., 2011), i.e.,

$$\hat{y}^k = \sum_{i=1}^{c^k} \Psi_i(x^k) y_i^k \quad (3.15)$$

with normalized membership functions

$$\Psi_i(x^k) = \frac{\exp\left[\left(x^k - v_i^k\right) \Sigma_i^{-1} \left(x^k - v_i^k\right)^T\right]}{\sum_{i=1}^{c^k} \exp\left[\left(x^k - v_i^k\right) \left(\Sigma_i^{-1}\right)^{-1} \left(x^k - v_i^k\right)^T\right]} \quad (3.16)$$

where v_i^k and Σ_i^k are the center and dispersion matrix of the i -th cluster membership function at step k . Hence, the consequent parameters and matrix Q_i of the update formulas for rule i at each iteration k are as follows:

$$\gamma_i^{k+1} = \gamma_i^k + Q_i^{k+1} x^k \Psi_i(x^k) \left[y_i^k - \left((x^k)^T \gamma_i^k \right) \right] \quad (3.17)$$

$$Q_i^{k+1} = Q_i^k - \frac{\Psi_i(x^k) Q_i^k x^k \left(x^k\right)^T Q_i^k}{1 + \left(x^k\right)^T Q_i^k x^k} \quad (3.18)$$

In case a cluster is updated, the antecedent parameters of the corresponding rule are updated. In case a new cluster is created, the consequent parameters are computed as the weighted average of the parameters of the existing clusters

$$\gamma_{new}^k = \frac{\sum_{i=1}^{c^k} \gamma_i^k \rho_i^k}{\sum_{i=1}^{c^k} \rho_i^k} \quad (3.19)$$

In case two clusters i and j are merged, the consequent parameters are computed as follows:

$$\gamma_{new}^k = \frac{\gamma_i^k \rho_i^k + \gamma_j^k \rho_j^k}{\rho_i^k + \rho_j^k} \quad (3.20)$$

In both cases, the matrix Q is a set as $Q_{new}^k = \omega I_{m+1}$, where I_{m+1} is an $m+1$ identity matrix, and ω is a large real value, (GOMIDE, F., 2011) suggests $\omega \in [10^2, 10^4]$. The evolving Multivariable Gaussian fuzzy system modeling procedure is shown in Algorithm 2.

Algorithm 2 eMG

```

1: Input:  $x, y_d, \lambda, \omega, \Sigma_{init}$ 
2: Output:  $y_s$ 
3:  $\gamma_1 = [y^1 0 \dots 0]$ 
4:  $Q_1 = \omega I_{m+1}$ 
5: for  $k = 1, \text{length}(x)$  do
6:   Compute the output
7:   for  $i = 1, c$  do
8:      $\rho_i = \exp \left[ -\frac{1}{2} (x^k - v_i)^T \Sigma_i^{-1} (x^k - v_i) \right]$ 
9:      $y_i = x^k \gamma_i$ 
10:  end for
11:   $y_s = \frac{\sum_{i=1}^c \rho_i y_i}{\sum_{i=1}^c \rho_i}$ 
12:   $[v, \Sigma, \text{cluster\_created}, \text{cluster\_merged}, c, \text{id}x^k] =$ 
     $MGEC(k, x^k, v, \Sigma, \lambda, \omega, \Sigma_{init}, c)$ 
13:  if  $\text{cluster\_created}$  then
14:    Create a new rule
15:     $\gamma_c = \frac{\sum_{i=1}^{c-1} \gamma_i \rho_i}{\sum_{i=1}^{c-1} \rho_i}$ 
16:     $Q_c = \omega I_{m+1}$ 
17:  end if
18:  for  $i = 1, c$  do
19:    Update consequent parameters
20:     $\gamma_i = \gamma_i + Q_i x^k \Psi_i(x^k) \left[ y_i^k - \left( (x^k)^T \gamma_i \right) \right]$ 
21:     $Q_i = Q_i - \frac{\Psi_i(x^k) Q_i x^k (x^k)^T Q_i}{1 + (x^k)^T Q_i x^k}$ 
22:  end for
23:  if  $\text{cluster\_merged}$  then
24:    Merge two rules
25:     $[i, j] = \text{cluster\_merged}$ 
26:     $\gamma_i = \frac{\gamma_i \rho_i + \gamma_j \rho_j}{\rho_i + \rho_j}$ 
27:     $Q_i = \omega I_{m+1}$ 
28:  end if
29: end for

```

For a detailed discussion about the eMG procedure see (GOMIDE, F., 2011).

3.4 THE SET-MEMBERSHIP EVOLVING MULTIVARIABLE GAUSSIAN MODEL

The Set-Membership (SM) concept can be seen as an adaptive filtering technique that acts by adjusting a given parameter as a function of model errors (LAMARE, R. C., 2011). The update of the parameter is performed by comparing the error with a pre-set default value. In this work, the parameter to be adjusted was chosen as the eMG model's learning rate α , as mentioned earlier. In this way, if the error value is greater than a limit ($\bar{\gamma}$), then the variation in the learning rate (α) increases to improve the model's learning. Otherwise, α becomes zero, reducing computational complexity of the model (ALVES,

K. S., 2020). Some works in the literature (RIBEIRO, M. V., 2017, LAMARE, R. C., 2011) demonstrate that SM is a framework that limits the increase of the error, reduces computational complexity and improves the capacity of convergence in a learning process. So, for a learning procedure based on the SM concept, the learning rate α is updated at each step k as follows:

$$\alpha^k = \begin{cases} 1 - \frac{\bar{\gamma}}{|\tilde{e}^k|}, & \text{if } |\tilde{e}^k| > \bar{\gamma} \\ 0, & \text{otherwise} \end{cases} \quad (3.21)$$

where \tilde{e}_i^k is the error at i -th iteration. Thresholds, inferior limit (IL) and superior limit (SL) are set to keep α in a specified range, i.e., $\alpha \in [IL, SL]$, where $IL \geq 0$, $SL \leq 1$, and $IL \leq SL$.

To the SM-eMG model proposed in this work, we applied the concept of SM filtering, previously discussed, to the eMG model presented in Section 3.3. The main contribution of this model in relation to original eMG is the application of Equation (3.21) to produce a new strategy to update the learning rate parameter in the eMG algorithm. This results in a convergent learning method in which the model's error is asymptotically upper bounded in magnitude (LAMARE, R. C., 2011).

In this way, during the learning process, when an existing cluster is updated, after using Equations (3.4) and (3.5), the learning rate is also updated as a function of model error making an algorithm that requires less computational effort and with faster tracking without compromising the error performance (RIBEIRO, M. V., 2017, LAMARE, R. C., 2011).

The SM model is shown in Algorithm 3.

Algorithm 3 SM

- 1: $\tilde{e}^k = y^k - \hat{y}^k$
 - 2: **if** $\tilde{e}^k > \bar{\gamma}$ **then**
 - 3: $\alpha^k = 1 - \frac{\bar{\gamma}}{|\tilde{e}^k|}$
 - 4: **else**
 - 5: $\alpha^k = 0$
 - 6: **end if**
 - 7: **if** $\alpha^k < IL$ **then**
 - 8: $\alpha^k = IL$
 - 9: **end if**
 - 10: **if** $\alpha^k > SL$ **then**
 - 11: $\alpha^k = SL$
 - 12: **end if**
-

3.5 THE ENHANCED SET-MEMBERSHIP EVOLVING MULTIVARIABLE GAUSSIAN MODEL

The Enhanced Set-Membership (ESM) approach is an improvement of the SM method described in the previous Section. In this approach, the learning rate α is reduced when the error is lower than $\bar{\gamma}$ instead of becoming zero as in the SM. This adjustment mechanism of the learning rate is given at each step k , as follows:

$$\alpha^k = \begin{cases} \alpha^{k-1} + \frac{|\tilde{e}^k|}{10^{gr \times \bar{\gamma}}}, & \text{if } |\tilde{e}^k| > \bar{\gamma} \\ \alpha^{k-1} - \frac{|\tilde{e}^k|}{10^{dr \times \bar{\gamma}}}, & \text{otherwise} \end{cases} \quad (3.22)$$

where $gr, dr \in \mathbb{Z}$ are the rate of parameter increase and decrease, respectively.

In a similar way to the SM-eMG, the ESM-eMG model proposed suggests an adaptive way to update the learning rate α as a function of model error, through the application of Equation (3.22). Again, as proved in literature (LAMARE, R. C., 2011), the SM/ESM-based model provides a flexible trade-off between the computational complexity and the bound on the asymptotic model error in the learning process, which leads to a better performance of the model.

In particular, the conjunction of both, SM and ESM approaches, with the eMG model emerges as a promising alternative to the thermal modeling of power transformers. Since it combines the characteristic of the eMG to robustly adapt to dynamically evolving environments with the characteristic of the SM approaches to reduce the computational complexity of the resulting model. This statement is tested in the next Chapter.

The ESM model is shown in Algorithm 4.

Algorithm 4 ESM

```

1:  $\tilde{e}^k = y^k - \hat{y}^k$ 
2: if  $|\tilde{e}^k| > \bar{\gamma}$  then
3:    $\alpha^k = \alpha^{k-1} + \frac{|\tilde{e}^k|}{10^{gr \times \bar{\gamma}}}$ 
4: else
5:    $\alpha^k = \alpha^{k-1} - \frac{|\tilde{e}^k|}{10^{dr \times \bar{\gamma}}}$ 
6: end if
7: if  $\alpha^k < IL$  then
8:    $\alpha^k = IL$ 
9: end if
10: if  $\alpha^k > SL$  then
11:    $\alpha^k = SL$ 
12: end if

```

4 EXPERIMENTAL RESULTS

To evaluate the modeling performance of proposed models, the following parameters were calculated: Root Mean Squared Error (RMSE), Nondimensional Error Index (NDEI) and Mean Absolute Error (MAE), been expressed by Equations (4.1), (4.3) and (4.2), respectively.

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (y^k - \hat{y}^k)^2} \quad (4.1)$$

$$NDEI = \frac{RMSE}{std(y^k)} \quad (4.2)$$

$$MAE = \frac{1}{n} \sum_{k=1}^n |y^k - \hat{y}^k| \quad (4.3)$$

where n is the size of the data set used, y^k is the target output, \hat{y}^k is the obtained output, and $std()$ is the standard deviation function.

In order to statistically validate the performance of the proposed models we also perform the Morgan-Granger-Newbold test (MGN) introduced in (MARIANO, R. S., 2002).

The statistical test is performed as follow:

$$MGN = \frac{\hat{\rho}_{sd}}{\sqrt{\frac{1-\hat{\rho}_{sd}^2}{n-1}}} \quad (4.4)$$

where $\hat{\rho}_{sd}$ is the correlation coefficient between s and d , with $s = r_1 + r_2$, $d = r_1 - r_2$, r_1 is the residual of model 1 and r_2 is the residual of model 2. As can be seen in the Equation (4.4), this test is based on the correlation between the sum (r) and the difference (d) of the prediction errors. Therefore, the MGN values, and consequently the p -values, can vary significantly due to the differences in models outputs.

This statistical test is a Student's t -distribution with $n - 1$ degrees of freedom, considering a significant level (α_{MGN}) of 5%. If the p -value is lower than α_{MGN} , we reject the null hypothesis, which assumes the models have equal accuracy.

In eMG, SM-eMG and ESM-eMG some parameters are adopted in the simulations for all this models as specified as follows: $\omega = 10^3$, $\lambda = 0.05$ and the initial $\alpha = 0.01$. Furthermore, the SM-eMG and ESM-eMG have the following additional parameters: $IL = 0.00001$, $SL = 0.1$, $gr = 1$, $dr = 1$, these last two only for ESM-eMG. In short, all of these parameters cited were chosen from the references already cited throughout the work. The other parameters (w , Σ_{init} and $\bar{\gamma}$) were chosen heuristically.

For both models, SM-eMG and ESM-eMG, a $\bar{\gamma}$ specified was chosen as the best result between 700 simulations, starting at $\bar{\gamma} = 0.001$ and ending at $\bar{\gamma} = 0.7$. More information about the developed codes can be found at: <https://github.com/rocha-mvg/sm-esm-emg>.

4.1 SYNTHETIC DATA SETS

Considering the evolving models introduced, tests were performed to develop long-term forecasting of the Mackey-Glass time series, and using a classic nonlinear system identification problem, and a high-dimensional system identification problem.

The MAE values of some simulations will be omitted in this Section, cause the other related authors did not use them as a metric in their work.

4.1.1 Mackey-Glass Time Series Forecasting

Simulations were performed with the proposed algorithms to develop long-term forecasting of the Mackey-Glass time series, and its results are compared with other approaches.

This series is made through the time-delay differential equation:

$$\frac{dx(t)}{dt} = \frac{0.2x(t - \tau)}{1 + x^{10}(t - \tau)} \quad (4.5)$$

where $x(0) = 1.2$ and $\tau = 17$.

The goal is to predict the value x^{k+85} from the input vector $[x^{k-18} \ x^{k-12} \ x^{k-6} \ x^k]$ for any k .

The computational simulations were performed following the same roadmap presented in (GOMIDE, F., 2011). The data set was generated with $k = 8001$ samples. Thus, 3000 data samples were collected, considering $k \in [201, 3200]$, and they were used as inputs of the evolving learning procedure (training phase). Afterwards, 500 data samples, considering $k \in [5001, 5500]$, were collected to verify the performance during the test phase. The same training and test sets were adopted for all models portrayed in this Subsection.

The additional parameters adopted for eMG, SM-eMG and ESM-eMG are: $w = 40$ and $\Sigma_{init} = 2 \times 10^{-2} I_4$. Furthermore, the $\bar{\gamma} = 0.2160$ and $\bar{\gamma} = 0.1970$ were chosen for SM-eMG and ESM-eMG, respectively. Table 1 shows the results of simulations.

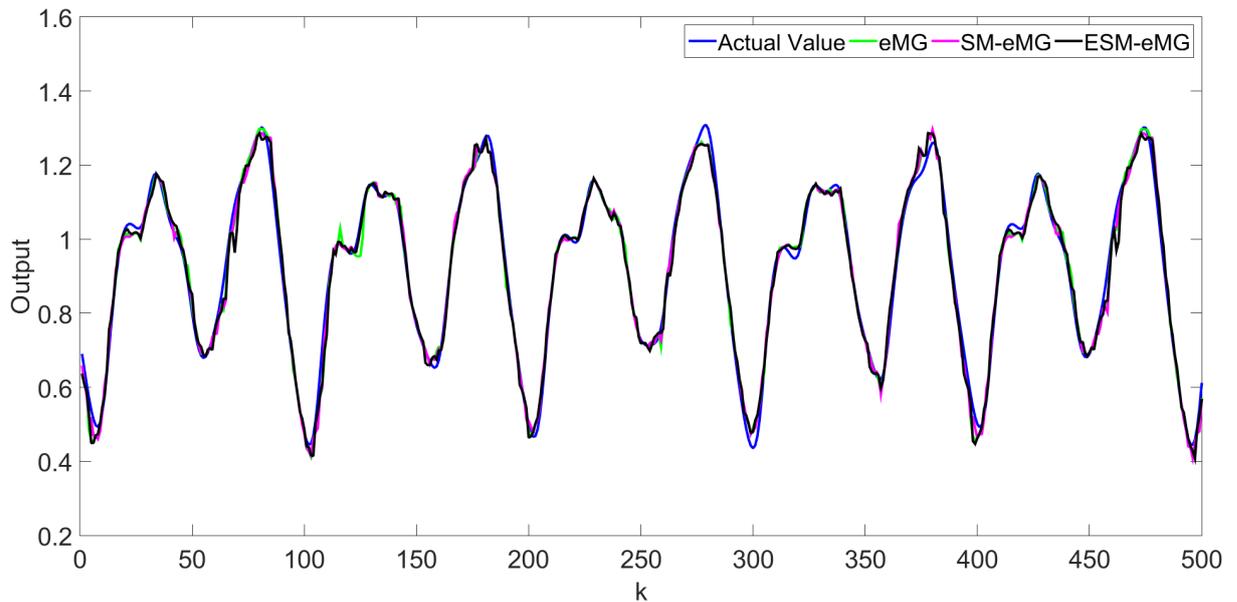
Table 1: Performance of Mackey-Glass time series forecasting approaches

Model	RMSE	NDEI	MAE	Rules
eTS (FILEV, D., 2004)	0.0858	0.3720	-	9
xTS (ZHOU, X., 2006)	0.0764	0.3310	-	10
DENFIS (QUN, S., 2002)	0.0637	0.2760	-	58
FLEXFIS <i>VarA</i> (LUGHOFER, E. D., 2008)	0.0475	0.2060	-	69
FLEXFIS <i>VarB</i> (LUGHOFER, E. D., 2008)	0.0362	0.1570	-	89
ESM-eMG	0.0270	0.1174	0.0194	43
eMG (GOMIDE, F., 2011)	0.0254	0.1103	0.0179	36
SM-eMG	0.0239	0.1040	0.0178	44

Source: Author (2020).

The result of the simulation is shown in Figure 2, comparing the actual value with the results of the ESM-eMG, SM-eMG and eMG.

Figure 2: Comparison of Mackey-Glass time series forecasting



Source: Author (2020).

Table 2 presents the results of the MGN tests considering the proposed models in relation to the original eMG model. Since the p -value is lower than 0.05, the null hypothesis is rejected, assuming that the models have different accuracy.

Table 2: Results of the MGN test, Mackey-Glass time series forecasting

Model 1 x Model 2	MGN	p -value
ESM-eMG x eMG	1.8061	0.0357
SM-eMG x eMG	2.1757	0.0150

Source: Author (2020).

4.1.1.1 Comparative Mackey-Glass Time Series Forecasting for Different Values of τ

In this Subsection, a comparative study of the performance of the proposed models is made considering the Mackey-Glass time series under different values of the τ parameter. Considering equation (4.5), the τ parameter defines how chaotic the Mackey-Glass time series will be. As τ increases, keeping the other parameters, the series becomes more chaotic.

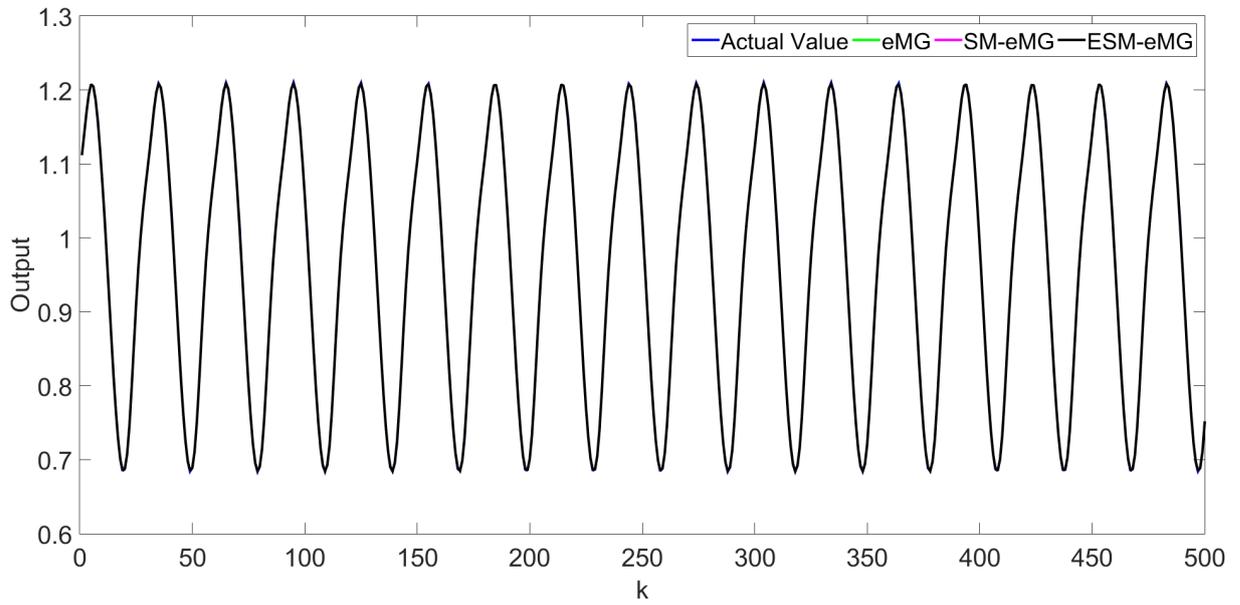
The values of $\tau = \{10, 20, 70\}$ were adopted to evaluate the performance of the proposed models. Furthermore, $x(0) = 1.2$ was maintained. The goal is to predict the value x^{k+85} from the input vector $[x^{k-18} \ x^{k-12} \ x^{k-6} \ x^k]$ for any k .

The computational simulations were performed following the same roadmap presented in the previous Subsection, varying only the database for each τ . The additional parameters adopted for eMG, SM-eMG and ESM-eMG are: $w = 40$ and $\Sigma_{init} = 2 \times 10^{-2} I_4$. Furthermore, the $\bar{\gamma} = 0.2160$ and $\bar{\gamma} = 0.1970$ were chosen for SM-eMG and ESM-eMG, respectively. Table 3 shows the results of simulations. Figures 3, 4 and 5 show the series for $\tau = \{10, 20, 70\}$, respectively.

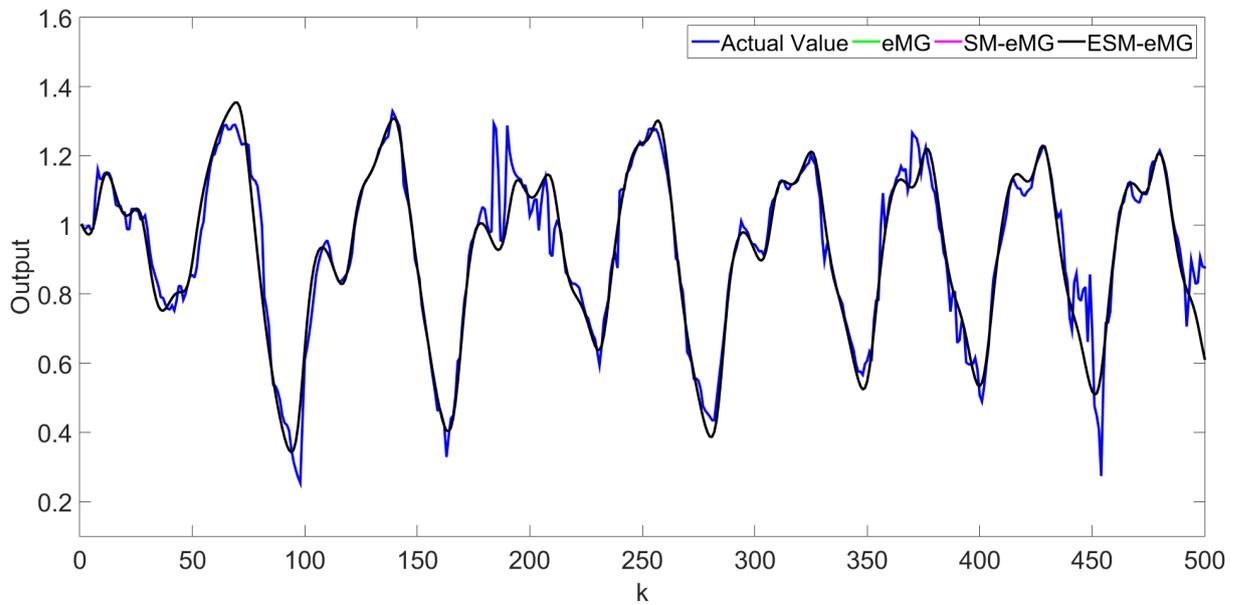
Table 3: Performance of Mackey-Glass time series forecasting approaches

Model	τ	RMSE	NDEI	MAE	Rules
eMG (GOMIDE, F., 2011)	10	0.000633	0.003562	0.000458	10
SM-eMG	10	0.000626	0.003520	0.000420	14
ESM-eMG	10	0.000625	0.003513	0.000419	14
eMG (GOMIDE, F., 2011)	20	0.064503	0.268862	0.038642	60
SM-eMG	20	0.070043	0.291956	0.042115	58
ESM-eMG	20	0.089415	0.372700	0.052973	51
eMG (GOMIDE, F., 2011)	70	0.314659	0.999031	0.247178	73
SM-eMG	70	0.314097	0.997247	0.254866	73
ESM-eMG	70	0.346286	1.099445	0.283625	59

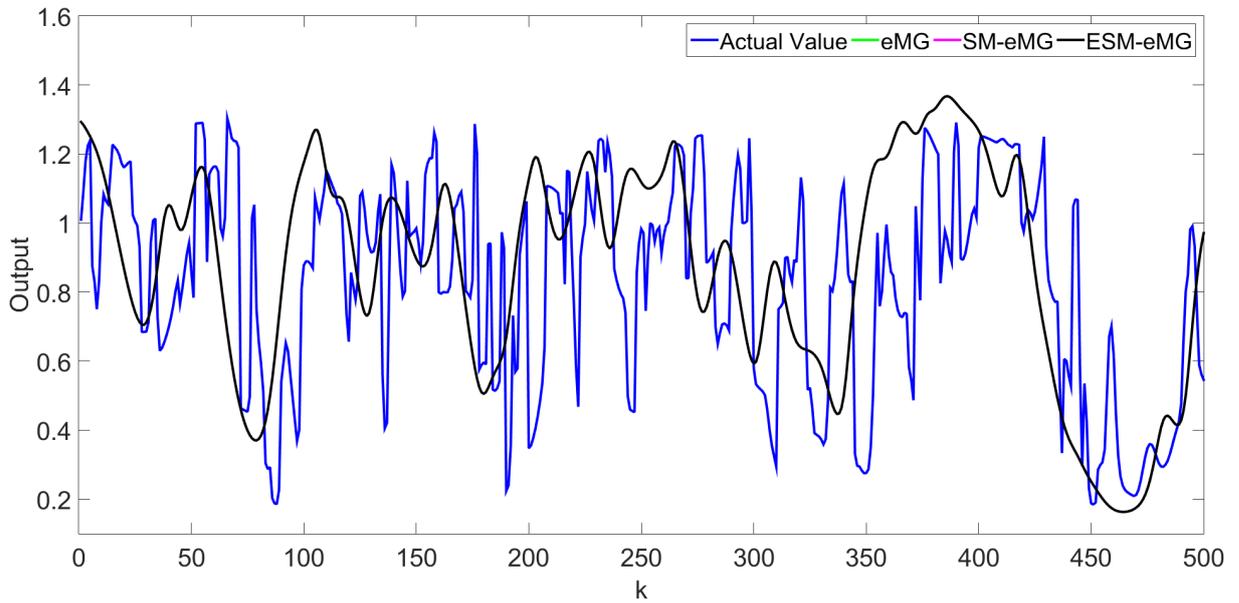
Source: Author (2020).

Figure 3: Comparison of Mackey-Glass time series forecasting, with $\tau = 10$ 

Source: Author (2020).

Figure 4: Comparison of Mackey-Glass time series forecasting, with $\tau = 20$ 

Source: Author (2020).

Figure 5: Comparison of Mackey-Glass time series forecasting, with $\tau = 70$ 

Source: Author (2020).

It is noticed that the proposed models performed better in scenarios with lower tau value. As expected, as the value of τ increases, the error also increases. In addition, it can be noted that even when the time series are more chaotic, the SM-eMG or ESM-eMG achieved better or at least competitive results compared to the original eMG.

4.1.2 Nonlinear System Identification

Simulations were performed with the proposed algorithms using a classic nonlinear system identification problem, and its results are compared with other approaches.

The nonlinear system to be identified is defined by:

$$y^k = \frac{y^{k-1}y^{k-2}(y^{k-1} - 0.5)}{1 + (y^{k-1})^2 + (y^{k-2})^2} + u^{k-1} \quad (4.6)$$

where $u^k = \sin(2\pi k/25)$, and $y^0 = y^1 = 0$.

The intention is to predict the current output of past inputs and outputs. The model of this data set has the following format:

$$\hat{y}^k = f(y^{k-1}, y^{k-2}, u^{k-1}) \quad (4.7)$$

where \hat{y}^k is the output.

The computational simulations were performed following the same roadmap presented in (GOMIDE, F., 2011). The data set was generated with $k = 5200$ samples. Thus, 5000 data samples were collected, considering $k \in [1, 5000]$, and they were used as inputs of the evolving learning procedure (training phase). Afterwards, 200 data samples, considering $k \in [5001, 5200]$, were collected to verify the performance during the test phase. The same training and test sets were adopted for all models portrayed in this Subsection.

The additional parameters adopted for eMG, SM-eMG and ESM-eMG are: $w = 40$ and $\Sigma_{init} = 10^{-2}I_3$. Furthermore, the $\bar{\gamma} = 0.0030$, was chosen for both SM-eMG and ESM-eMG. Table 4 shows the results of simulations.

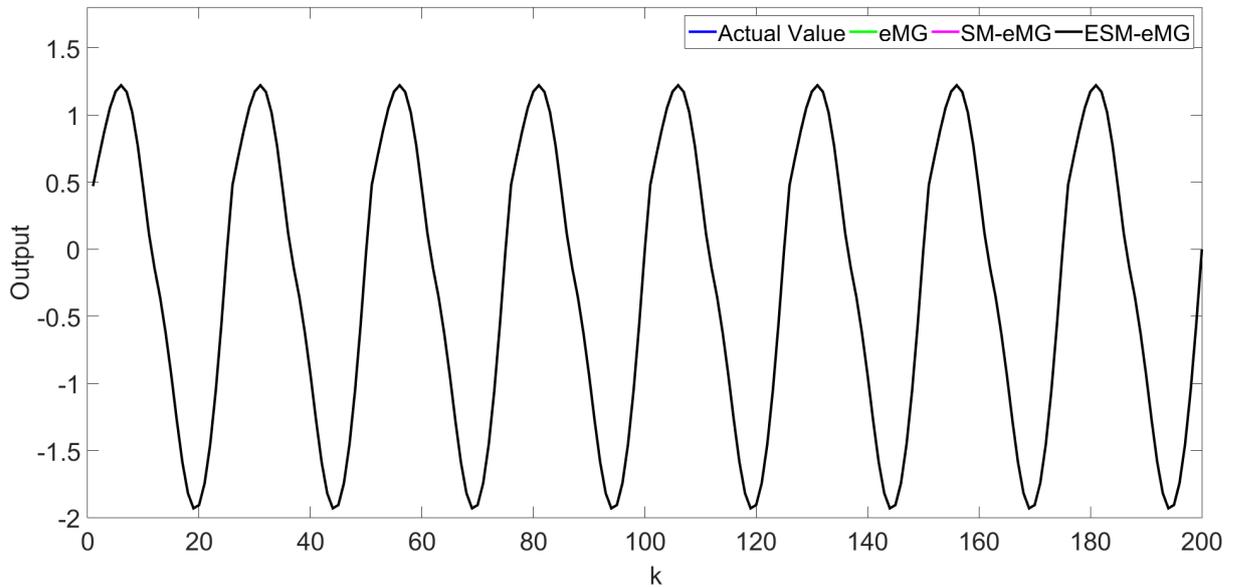
Table 4: Performance of nonlinear system identification methods

Model	RMSE	NDEI	MAE	Rules
SAFIS (SARATCHANDRAN, P., 2006)	0.0221	0.0202	-	17
SOFMLS (RUBIO, J., 2009)	0.0201	0.0183	-	5
FLEXFIS <i>VarA</i> (LUGHOFER, E. D., 2008)	0.0176	0.0161	-	5
FLEXFIS <i>VarB</i> (LUGHOFER, E. D., 2008)	0.0171	0.0156	-	8
xTS (ZHOU, X., 2006)	0.0063	0.0057	-	5
SM-eMG	0.00058	0.00053	0.00021	22
eMG (GOMIDE, F., 2011)	0.00022	0.00020	0.00013	21
ESM-eMG	0.00010	0.00009	0.00004	22

Source: Author (2020).

The result of the simulation is shown in Figure 6, comparing the actual value with the results of the ESM-eMG, SM-eMG and eMG.

Figure 6: Comparison of nonlinear system identification



Source: Author (2020).

Table 5 presents the results of the MGN tests considering the proposed models in relation to the original eMG model. Since the p -value is lower than 0.05, the null hypothesis is rejected, assuming that the models have different accuracy.

Table 5: Results of the MGN test, nonlinear system identification

Model 1 x Model 2	MGN	p -value
ESM-eMG x eMG	19.6879	0.0000
SM-eMG x eMG	23.0536	0.0000

Source: Author (2020).

4.1.3 High-Dimensional System Identification

Simulations were performed with the proposed algorithms using a high-dimensional system identification problem, and its results are compared with other approaches.

The high-dimensional system to be identified is defined by:

$$y^k = \frac{\sum_{i=1}^m y^{k-i}}{1 + \sum_{i=1}^m (y^{k-i})^2} + u^{k-1} \quad (4.8)$$

where $u^k = \sin(2\pi k/20)$, and $y^j = 0$, for $j = 1, \dots, m$ and $m = 10$.

The intent is to predict the current output from past input and outputs. The data set has the form as follow:

$$\hat{y}^k = f(y^{k-1}, y^{k-2}, \dots, y^{k-10}, u^{k-1}) \quad (4.9)$$

where \hat{y}^k is the model output.

The computational simulations were performed following the same roadmap presented in (GOMIDE, F., 2011). The data set was generated with $k = 3300$ samples. Thus, 3000 data samples were collected, considering $k \in [1, 3000]$, and they were used as inputs of the evolving learning procedure (training phase). Afterwards, 300 data samples, considering $k \in [3001, 3300]$, were collected to verify the performance during the test phase. The same training and test sets were adopted for all models portrayed in this Subsection.

The additional parameters adopted for eMG, SM-eMG and ESM-eMG are: $w = 25$, $\Sigma_{init} = 2 \times 10^{-2} I_{11}$. Furthermore, that the $\bar{\gamma} = 0.0270$ and $\bar{\gamma} = 0.1360$ were chosen for SM-eMG and ESM-eMG, respectively. Table 6 shows the results of simulations.

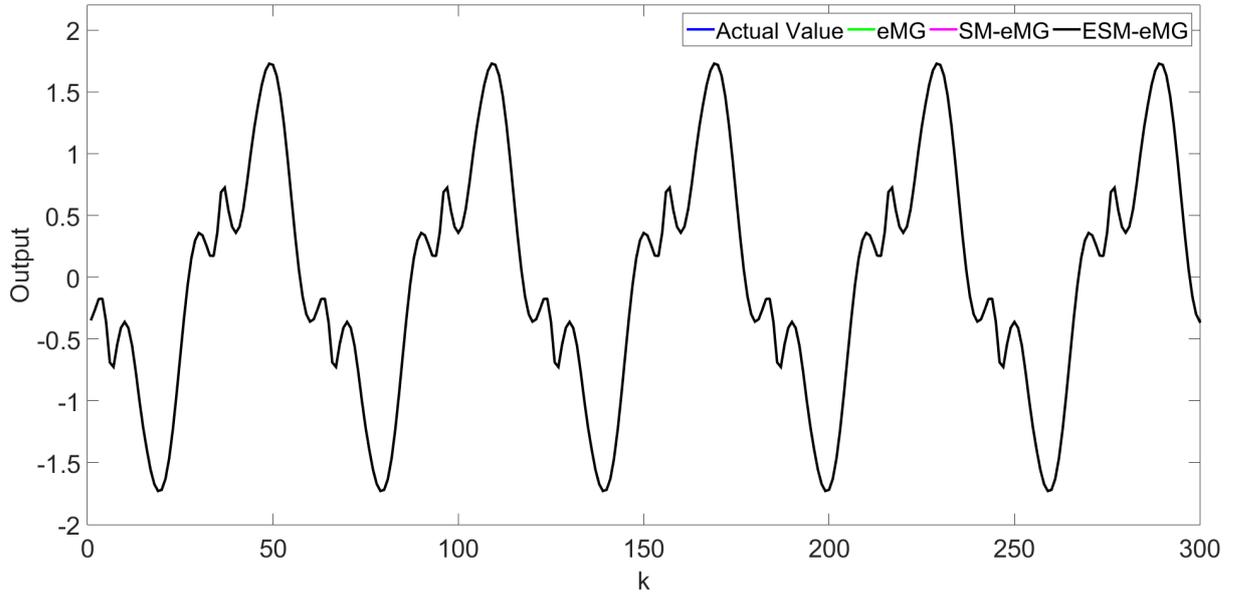
Table 6: Performance of high-dimensional system identification methods

Model	RMSE	NDEI	MAE	Rules
xTS (ZHOU, X., 2006)	0.0331	0.0351	-	9
FLEXFIS <i>VarA</i> (LUGHOFER, E. D., 2008)	0.0085	0.0090	-	15
eTS (FILEV, D., 2004)	0.0075	0.0080	-	14
eMG (GOMIDE, F., 2011)	0.000016	0.000017	0.000009	62
SM-eMG	0.000015	0.000016	0.000009	62
ESM-eMG	0.000013	0.000013	0.000006	62

Source: Author (2020).

The result of the simulation is shown in Figure 7, comparing the actual value with the results of the ESM-eMG, SM-eMG and eMG.

Figure 7: Comparison of high-dimensional system identification



Source: Author (2020).

Table 7 presents the results of the MGN tests considering the proposed models in relation to the original eMG model. Since the p -value is lower than 0.05, the null hypothesis is rejected, assuming that the models have different accuracy.

Table 7: Results of the MGN test, high-dimensional system identification

Model 1 x Model 2	MGN	p -value
ESM-eMG x eMG	11.0879	0.0000
SM-eMG x eMG	3.4033	0.0004

Source: Author (2020).

4.2 THERMAL MODELING OF REAL POWER TRANSFORMERS

To evaluate the effectiveness of the method proposed in this work with actual data, the models presented in previous Chapter are also applied in the estimation of the hot-spot temperature of a real transformer whose characteristics are shown in Table 8.

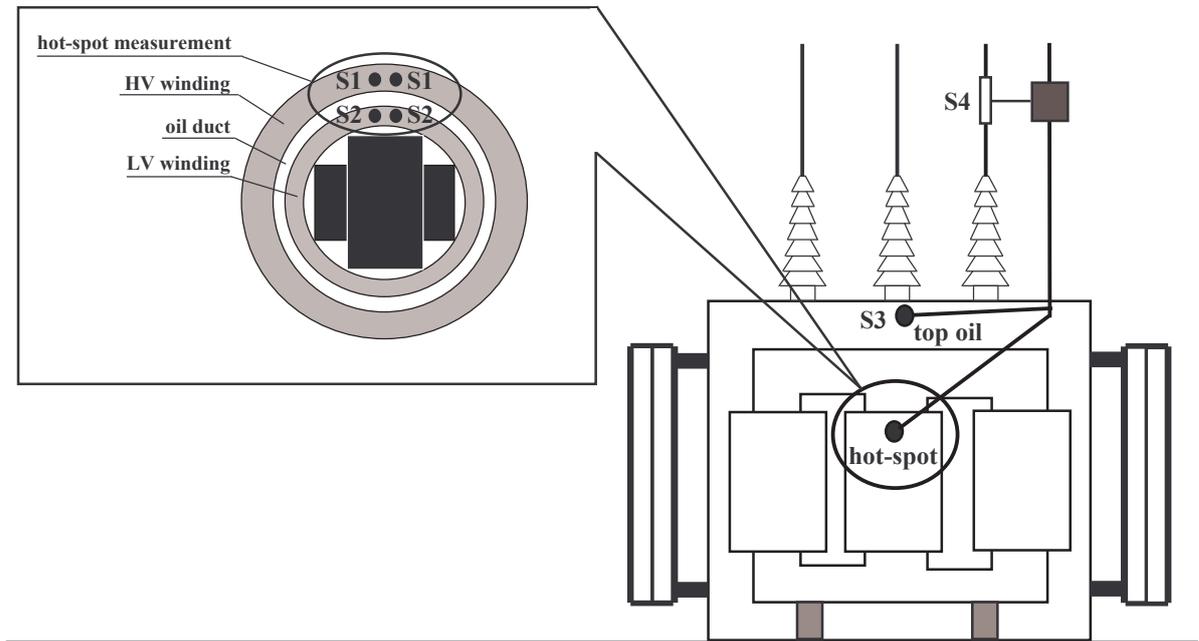
Table 8: Characteristics of the experimental power transformer

Copper losses	776 W
Factory year	MACE/1987
Iron losses	195 W
Nameplate rating	25 kVA
Tank dimensions	$64 \times 16 \times 80 \text{ cm}^3$
Top oil temperature rise at full load	73.1 °C
Type of cooling	ONAN
$V_{primary}/V_{secondary}$	10 kV / 380 kV
Weight of core and coil assembly	136 kg
Weight of oil	62 kg

Source: GOMIDE, F. (2008).

The data sets collected from this transformer are the same presented in (VACCARO, A., 2000) and were obtained through a measurement system composed of three fiber-optical-based temperature sensors and a hall-effect current sensor. The first two temperature sensors (S1 and S2) were inserted in the spacer between the disks at the top of the high-voltage and low-voltage windings, as shown in Figure 8. The aggregation of the values obtained from these two sensors provides a measure of the actual value of the transformer's hot-spot temperature (Θ_H). The third temperature sensor (S3) was inserted at the top of the tank and provides the actual value of the top oil temperature (Θ_{TO}). The hall-effect current sensor (S4) provides the actual value of the load current (K). Figure 8 shows the location of these sensors in the experimental transformer used in this work.

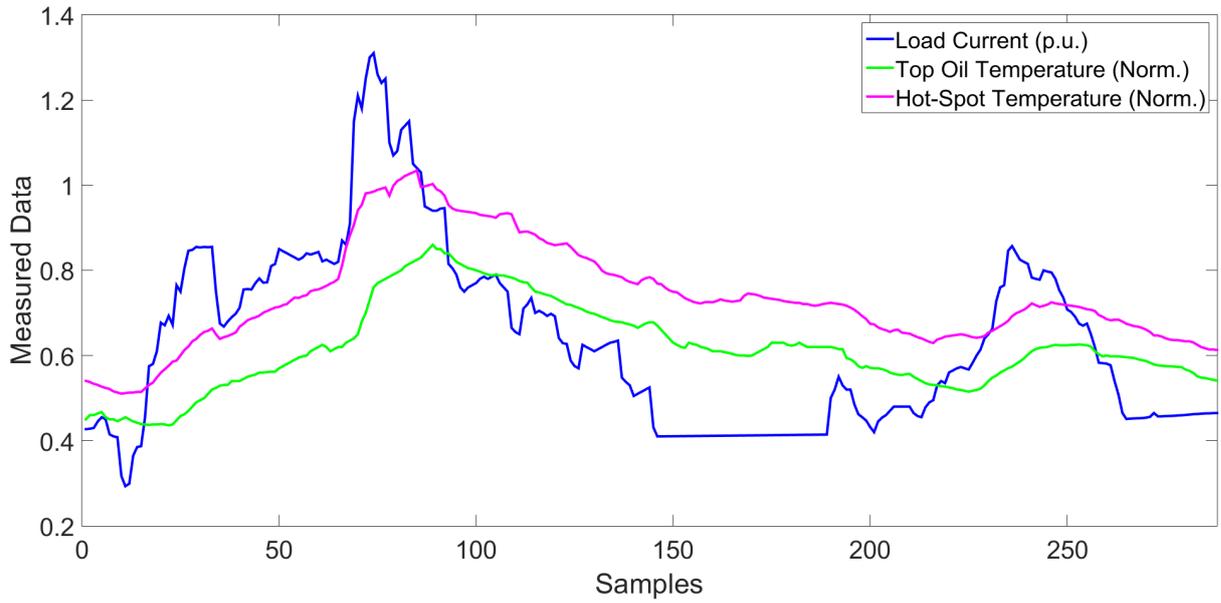
Figure 8: Sensor's location in the experimental transformer



Source: ALVES, K. S. (2020).

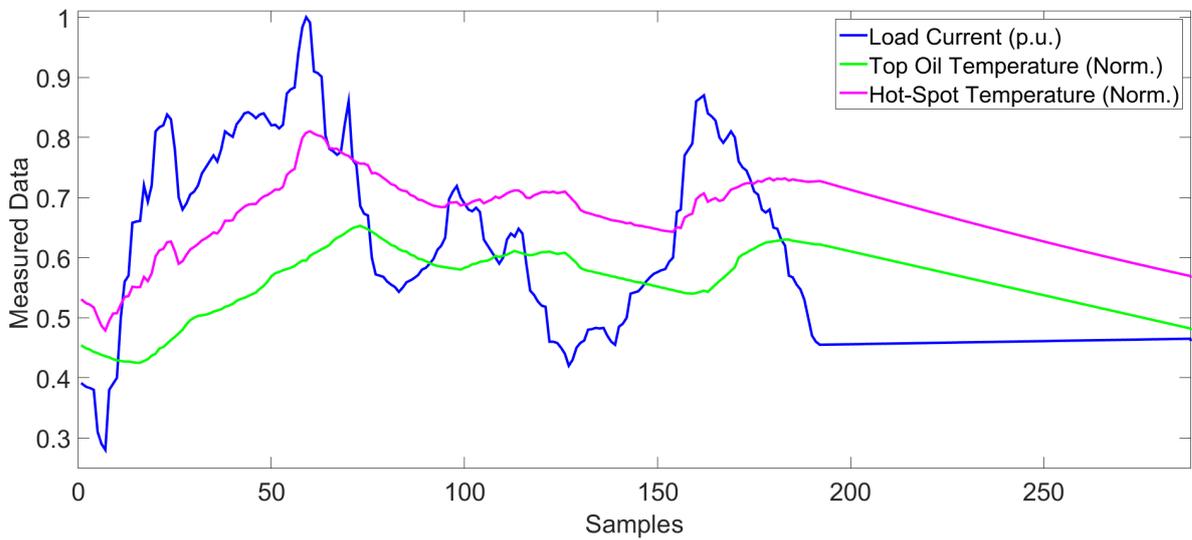
In our experiments, two data sets composed of the records of the temperatures and load current acquired from each sensor in an interval of 24h with a 5-min sample rate were used to evaluate the proposed models. Figure 9 shows the learning data set adopted during the training phase. To cover relevant operating conditions, two different load conditions were considered for model evaluation during the testing phase : *i*) Data set 1: without overload and *ii*) Data set 2: with overload. Figures 10 and 11 shows the behavior of the hot-spot and top oil temperatures for a given load current for these two data sets.

Figure 9: Learning data set



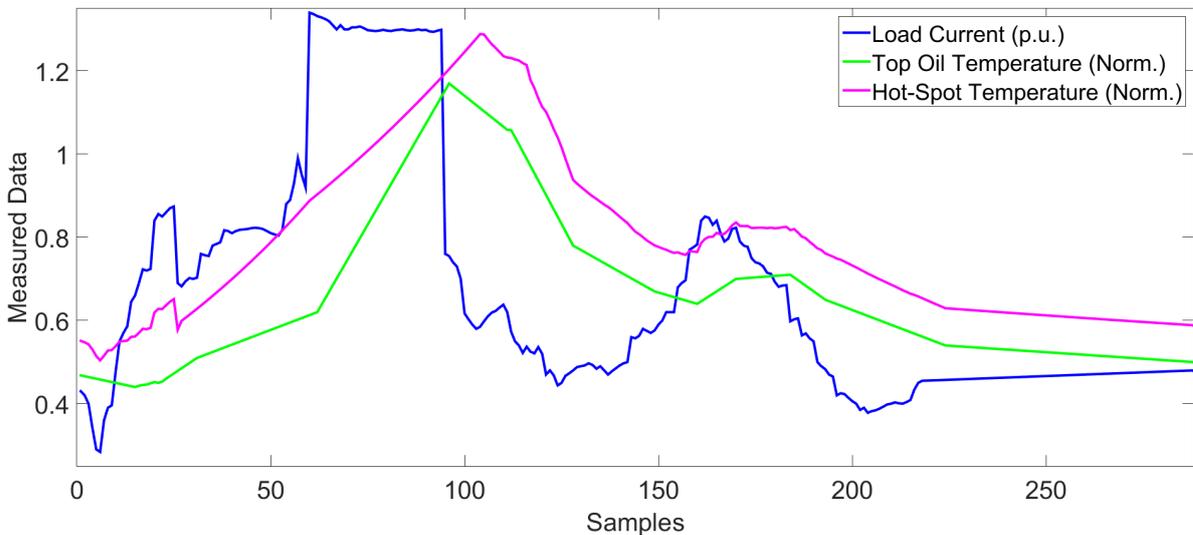
Source: Author (2020).

Figure 10: Data set 1: without overload condition



Source: Author (2020).

Figure 11: Data set 2: with overload condition



Source: Author (2020).

Therefore, the purpose of the proposed models is to estimate the hot-spot temperature from the load current and the top oil temperature. Different studies (GOMIDE, F., 2008, VACCARO, A., 2000), experimental trial and error tests indicate that the relevant model inputs for this case are the load current (K), the top oil temperature (Θ_{TO}) and one step delayed load current ($q^{-1}K$, where q^{-1} is the delay operator). This choice has shown to reduce the model sensitivity concerning fluctuations in the thermal parameters, which can vary considerably from one transformer to another (BIRATTARI, M., 2005).

To prove the efficiency of evolving models in the estimation of the hot-spot temperature of power transformers, they were also compared with other non-evolving (fixed structure) models described in the literature. These models include the deterministic model based on IEEE Standard C57.91-2011 (IEEE-DM) described in Chapter 2, a model based on a Multilayer Perceptron Neural Network (MLP) and a model based on an Adaptive Neurofuzzy Inference System (ANFIS) (JANG, J., 1993).

In the deterministic modeling (IEEE-DM) the experimental transformer characteristic parameters used in this work were the following:

$$R = 4, \Delta\Theta_{H,R} = 5^\circ \text{ C}, \Delta\Theta_{TO,R} = 54^\circ \text{ C}, \Theta_{H,R} = 80^\circ \text{ C}$$

$$\Theta_{A,R} = 21^\circ \text{ C}, q = 0.8, m = 0.8, \tau_{TO} = 3 \text{ h}, \tau_H = 0.1 \text{ h}$$

The MLP neural network was implemented with a single hidden layer with 4 neurons trained with the backpropagation algorithm. The ANFIS model was implemented with

four fuzzy sets for each input variable and four fuzzy rules generated by means of the fuzzy c-means clustering procedure (BEZDEK, J. C., 1981). The additional parameters adopted for eMG, SM-eMG and ESM-eMG are: $w = 15$ and $\Sigma_{init} = 2 \times 10^{-2}I_3$. Furthermore, the SM-eMG has $\bar{\gamma} = 0.0270$ for the data set 1 and $\bar{\gamma} = 0.0150$ for the data set 2. And ESM-eMG has $\bar{\gamma} = 0.0290$ for the data set 1 and $\bar{\gamma} = 0.0360$ for the data set 2.

Tables 9 and 10 show the results obtained for all models implemented in this work.

Table 9: Performance of thermal modeling of real power transformers, operation without overload condition

Model	RMSE	NDEI	MAE	Rules
IEEE-DM (IEEE, 2012)	1.0245	16.1089	0.7524	-
MLP (STORK, D. G., 2012)	0.0467	0.7336	0.0343	4
ePL-KRLS (BALLINI, R., 2018)	0.0137	0.2153	0.0101	1
ANFIS (JANG, J., 1993)	0.0124	0.1952	0.0091	4
eMG (GOMIDE, F., 2011)	0.0113	0.1791	0.0067	3
ESM-eMG	0.0105	0.1663	0.0074	3
SM-ePL-KRLS (ALVES, K. S., 2020)	0.0103	0.1624	0.0076	1
ESM-ePL-KRLS (ALVES, K. S., 2020)	0.0103	0.1616	0.0075	1
SM-eMG	0.0096	0.1521	0.0062	3

Source: Author (2020).

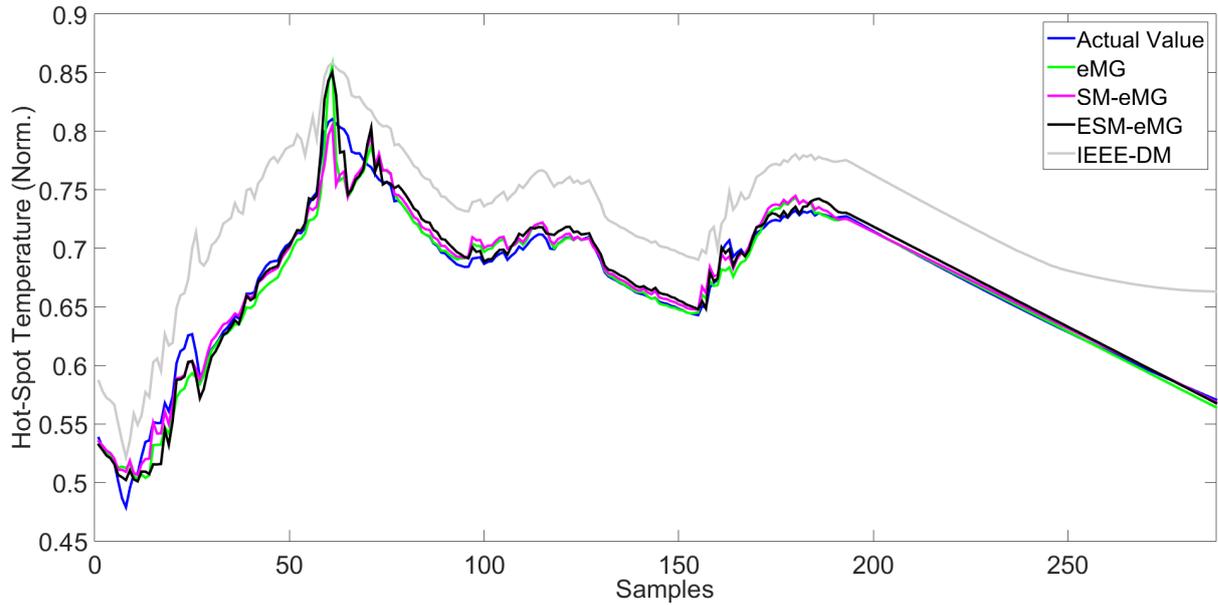
Table 10: Performance of thermal modeling of real power transformers, operation with overload condition

Model	RMSE	NDEI	MAE	Rules
IEEE-DM (IEEE, 2012)	0.4005	1.9446	0.2769	-
ANFIS (JANG, J., 1993)	0.0481	0.2340	0.0333	4
ePL-KRLS (BALLINI, R., 2018)	0.0330	0.1615	0.0242	2
eMG (GOMIDE, F., 2011)	0.0330	0.1615	0.0196	3
MLP (STORK, D. G., 2012)	0.0317	0.1539	0.0219	4
ESM-eMG	0.0283	0.1385	0.0203	3
SM-ePL-KRLS (ALVES, K. S., 2020)	0.0265	0.1293	0.0188	1
ESM-ePL-KRLS (ALVES, K. S., 2020)	0.0252	0.1231	0.0186	1
SM-eMG	0.0245	0.1197	0.0155	3

Source: Author (2020).

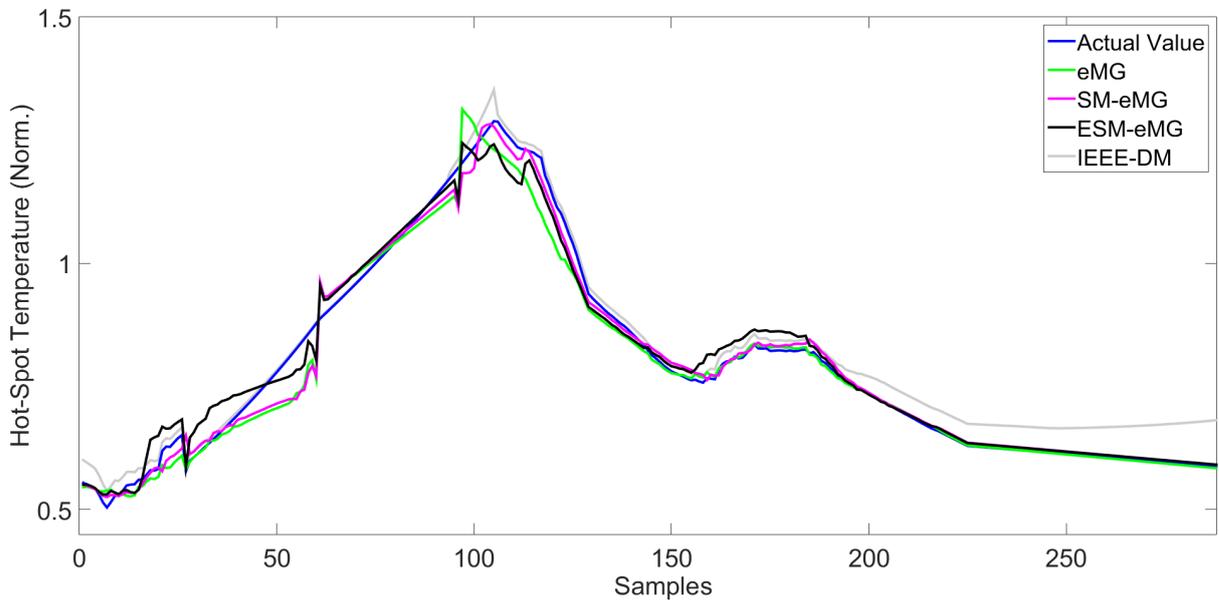
Figures 12 and 13, presents graphical depictions of the results obtained from SM-eMG and ESM-eMG models for the two data sets and compared with the deterministic model described in Chapter 2, which is the most used model in practice for the prediction of the hot-spot temperature, as mentioned earlier.

Figure 12: Hot-spot estimation, operation without overload condition



Source: Author (2020).

Figure 13: Hot-spot estimation, operation with overload condition



Source: Author (2020).

Tables 11 and 12 present the results of the MGN tests considering the proposed models in relation to the original eMG model. Since the p -value is lower than 0.05, the null hypothesis is rejected, assuming that the models have different accuracy. In Table 11, the SM-eMG x eMG test returned a p -value greater than the established limit α_{MGN} .

Thus, this is the only occurrence where the null hypothesis was accepted, assuming that the compared models have equal accuracy.

Table 11: Results of the MGN test, operation without overload condition

Model 1 x Model 2	MGN	<i>p</i> -value
ESM-eMG x eMG	3.2414	0.0004
SM-eMG x eMG	0.6436	0.2613

Source: Author (2020).

Table 12: Results of the MGN test, operation with overload condition

Model 1 x Model 2	MGN	<i>p</i> -value
ESM-eMG x eMG	23.0633	0.0000
SM-eMG x eMG	26.8007	0.0000

Source: Author (2020).

As can be seen in presented Tables and Figures, the proposed models reached the best results in general, with a competitive number of rules when compared with the original eMG model. These results show that the proposals yield improved performance if compared with previous techniques discussed in (GOMIDE, F., 2011) and (ALVES, K. S., 2020).

In particular, for the case where the data contains an overload condition the SM-eMG and ESM-eMG models show statistically significant evidence of much superior performance if compared to the other models. This is an important result because the determination of a precise real-time transformer overload rating is critical to increasing system operation margins (VACCARO, A., 2000).

It is interesting pondering that in order to real implementation in power transformers, a few simple modifications are necessary, for instance, adding temperature sensors in the locations specified to thermal modeling. The acquisition and management of the data from these sensors are a way to control and extend the time life of power transformers. As discussed in (ALVES, K. S., 2020), these modifications include the installation of a hall effect sensor to measure the load current and the insertion of only one temperature fiber-optical based sensor in the transformer's inspection cover to measure the top oil temperature. Many transformers in operation already have such sensors installed, which makes the proposed approach highly applicable to the real-world problem of estimating the hot-spot temperature of power transformers. These modifications are suitable for being non-invasive and low-cost implementations.

5 CONCLUSION

This work has introduced two innovative models to deal with the estimation of hot-spot temperature in power transformers: the so-called SM-eMG and ESM-eMG. The models were evaluated using synthetic time series forecasting and nonlinear system identification problems, and actual data for thermal modeling power transformers, which has two load conditions considered: with and without an overload condition.

To evaluate and compare the results were considered error metrics and the number of rules, as well as different evolving and non-evolving approaches. To support the effectiveness of the proposed models compared to eMG, the results of the MGN statistical tests were presented.

The obtained results showed that the SM-eMG, followed by ESM-eMG, reached the best results with a competitive number of rules in comparison with the original eMG model. These results suggest both as interesting options to integrate a decision support tool in the operational management of the electrical system, to control the load current and prolong the lifetime of the power transformers, in particular in the presence of overload conditions.

5.1 FUTURE WORKS

As future work, we plan to integrate one of the proposed models into a decision support tool to control the load current and extend the lifetime of the power transformers. Moreover, implement a system to update the initial parameters of the model in use, SM-eMG or ESM-eMG. Also, we plan to investigate the usefulness of interval type-2 fuzzy logic system (FLS) to handle the presence of uncertainty in the data sets.

REFERENCES

- HELL, M. B.; OLIVEIRA, F. L. C.; AGUIAR, E. P.; ALVES, K. S. **An enhanced set-membership evolving participatory learning with kernel recursive least squares applied to thermal modeling of power transformers.** *Electric Power Systems Research*, 184:106334, 2020.
- VIEIRA, R.; GOMIDE, F.; BALLINI, R. **Kernel Evolving Participatory Fuzzy Modeling for Time Series Forecasting.** In *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–9. IEEE, 2018.
- BEZDEK, J. C. *Pattern Recognition with Fuzzy Objective Function Algorithm.* Kluwer Academic Publishers, 1981.
- VILLACCI, D.; BONTEMPI, G.; VACCARO, A.; BIRATTARI, M. **The role of learning methods in the dynamic assessment of power components loading capability.** *IEEE Transactions on Industrial Electronics*, 52(1):280–289, 2005.
- SOUZA, L.; LEMOS, A.; CAMINHAS, W.; BOAVENTURA, W. **Thermal modeling of power transformers using evolving fuzzy systems.** *Engineering Applications of Artificial Intelligence*, 25(5):980–988, 2012.
- CHIU, S. Fuzzy model identification based on cluster estimation. *Journal of Intelligent & fuzzy systems*, 2(3):267–278, 1994.
- ANGELOV, P. P.; FILEV, D. **An approach to online identification of Takagi-Sugeno fuzzy models.** *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(1):484–498, Feb 2004. ISSN 1941-0492.
- KASABOV, N.; FILEV, D. **Evolving intelligent systems: methods, learning, & applications.** In *2006 International Symposium on Evolving Fuzzy Systems*, pages 8–18. IEEE, 2006.
- YAGER, R.; FILEV, D. **Approximate Clustering via the Mountain Method.** *IEEE Transactions on Systems, Man and Cybernetics*, 24(8):1279–1284, Aug 1994.
- HELL, M.; COSTA, P.; GOMIDE, F. **Recurrent Neurofuzzy Network in Thermal Modeling of Power Transformers.** *IEEE Transactions on Power Delivery*, 22(2): 904–910, April 2007.
- HELL, M.; COSTA, P.; GOMIDE, F. **Participatory Learning in Power Transformers Thermal Modeling.** *IEEE Transactions on Power Delivery*, 23(4):2058–2067, Oct 2008.

- LEMOS, A.; CAMINHAS, W.; GOMIDE, F. **Multivariable Gaussian Evolving Fuzzy Modeling System**. *IEEE Transactions on Fuzzy Systems*, 19(1):91–104, Feb 2011.
- HAYKIN, S. *Adaptive Filter Theory*. Upper Saddle River, NJ, USA:Prentice-Hall, 1996.
- IEEE. **IEEE Guide for Loading Mineral-Oil-Immersed Transformers and Step-Voltage Regulators**. *IEEE Std C57.91-2011 (Revision of IEEE Std C57.91-1995)*, pages 1–123, March 2012.
- IPPOLITO, L. **An Adaptive Fuzzy Approach to Predictive Overload Protection Systems for Power Transformers**. *Automatika: Journal of Automation, Measurement, Electronics, Computing and Communications*, 45(3-4):169–178, 2004.
- MIKHA-BEYRANVAND, M.; FAIZ, J.; REZAEALAM, B.; REZAEI-ZARE, A.; JAFARBOLAND, M. **Thermal analysis of power transformers under unbalanced supply voltage**. *IET Electric Power Applications*, 13(4):503–512, 2019.
- JANG, J. **ANFIS: adaptive-network-based fuzzy inference system**. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3):665–685, 1993.
- LI, Y.; WANG, Y.; JIANG, T. **Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation**. *AEU-International Journal of Electronics and Communications*, 70(7):895–902, 2016.
- CLARKE, P.; LAMARE, R. C. **Low-complexity reduced-rank linear interference suppression based on set-membership joint iterative optimization for DS-CDMA systems**. *IEEE Transactions on Vehicular Technology*, 60(9):4324–4337, 2011.
- TENBOHLEN, S.; SCHMIDT, N.; BREUER, C.; KHANDAN, S.; LEBRETON, R. **Investigation of Thermal Behavior of an Oil-Directed Cooled Transformer Winding**. *IEEE Transactions on Power Delivery*, 33(3):1091–1098, 2018.
- HELL, M.; BALLINI, R.; GOMIDE, F.; LIMA, E. Evolving fuzzy modeling using participatory learning. In D. Filev P. Angelov and N. Kasabov, editors, *Evolving Intelligent Systems: Methodology and Applications*. Wiley-Interscience/IEEE Press, New York/Piscataway, 2010.
- LUGHOFER, E. D. **FLEXFIS: A Robust Incremental Learning Approach for Evolving Takagi–Sugeno Fuzzy Models**. *IEEE Transactions on Fuzzy Systems*, 16(6):1393–1410, Dec 2008. ISSN 1941-0034.
- DIEBOLD, F. X.; MARIANO, R. S. **Comparing Predictive Accuracy**. *Journal of Business & Economic Statistics*, 20(1):134–144, 2002.

- KASABOV, N. K.; QUN, S. **DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction.** *IEEE Transactions on Fuzzy Systems*, 10(2):144–154, April 2002. ISSN 1941-0034.
- AGUIAR, E. P.; NOGUEIRA, F. M. A.; VELLASCO, M. M. B. R.; RIBEIRO, M. V. **Set-membership type-1 fuzzy logic system applied to fault classification in a switch machine.** *IEEE Transactions on Intelligent Transportation Systems*, 18(10): 2703–2712, 2017.
- RUBIO, J. **SOFMLS: Online Self-Organizing Fuzzy Modified Least-Squares Network.** *IEEE Transactions on Fuzzy Systems*, 17(6):1296–1309, Dec 2009. ISSN 1941-0034.
- RONG, H.; SUNDARARAJAN, N.; HUANG, G.; SARATCHANDRAN, P. **Sequential adaptive fuzzy inference system (SAFIS) for nonlinear system identification and prediction.** *Fuzzy sets and systems*, 157(9):1260–1275, 2006.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern classification.* John Wiley & Sons, 2012.
- GALDI, V.; IPPOLITO, L.; PICCOLO, A.; VACCARO, A. **Neural diagnostic system for transformer thermal overload protection.** *IEEE Proceedings - Electric Power Applications*, 147(5):415–421, Sep. 2000.
- AZAR, A. T.; VAIDYANATHAN, S. *Computational intelligence applications in modeling and control.* Springer, 2015.
- DAPONTE, P.; GRIMALDI, D.; PICCOLO, A.; VILLACCI, D. **A neural diagnostic system for the monitoring of transformer heating.** *Measurement*, 18(1):35–46, 1996.
- YAGER, R. **A model of participatory learning.** *IEEE transactions on systems, man, and cybernetics*, 20(5):1229–1234, 1990.
- YAGER, R. Participatory learning: A paradigm for more human like learning. In *2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No. 04CH37542)*, volume 1, pages 79–84. IEEE, 2004.
- YOUNG, P. *Recursive estimation and time-series analysis: an introduction.* Springer Science & Business Media, 2012.
- ANGELOV, P.; ZHOU, X. **Evolving Fuzzy Systems from Data Streams in Real-Time.** In *2006 International Symposium on Evolving Fuzzy Systems*, pages 29–35, Sep. 2006.

APPENDIX A – Publications

This Appendix presents the list of publications related to this work.

A.1 Journal

1. ROCHA, M. V. G., ALVES, K. S. T. R., HELL, M. B., OLIVEIRA, F. L. C., AGUIAR, E. P., Power Transformers Thermal Modeling using an Enhanced Set-Membership Multivariable Gaussian Evolving Fuzzy System. *Electric Power Systems Research*, 2020. - Under review (Submitted on May 22, 2020).

A.2 Congress

1. ROCHA, M. V. G., ALVES, K. S. T. R., HELL, M. B., OLIVEIRA, F. L. C., AGUIAR, E. P., Modelagem Térmica de Transformadores de Potência Baseada em um Sistema Fuzzy Evolutivo Set-Membership Gaussiano Multivariado. *In: CONGRESSO BRASILEIRO DE AUTOMÁTICA*, 23., 2020. **Anais da Sociedade Brasileira de Automática**, v. 2, n. 1, 2020.
2. SOUZA, D. P. M., CHRISTO, E. S., ROCHA, M. V. G., ALVES, K. S. T. R., HELL, FERNANDES, T. E., M. B., OLIVEIRA, F. L. C., AGUIAR, E. P., Modelagem Térmica de Transformadores de Potência Baseada em SODA e Sistema de Inferência Fuzzy Otimizado por Enxame de Partículas. *In: CONGRESSO BRASILEIRO DE AUTOMÁTICA*, 23., 2020. **Anais da Sociedade Brasileira de Automática**, v. 2, n. 1, 2020.