Electrical Power and Energy Systems 80 (2016) 150–159

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Power transformers' condition monitoring using neural modeling and the local statistical approach to fault diagnosis

G. Rigatos [a,*], P. Siano [b]

[a] Unit of Industrial Automation, Industrial Systems Institute, 26504 Rion Patras, Greece
[b] Department of Industrial Engineering, University of Salerno, 84084 Fisciano, Italy

ARTICLE INFO

Article history:
Received 7 October 2014
Received in revised form 29 December 2015
Accepted 6 January 2016

Keywords:
Power transformer
Thermal condition monitoring
Neural-fuzzy networks
Fault diagnosis
Local statistical approach

ABSTRACT

On-line monitoring of electric power transformers can provide a clear indication of their status and ageing behavior. This paper proposes neural modeling and the local statistical approach to fault diagnosis for the detection of incipient faults in power transformers. The method can detect transformer failures at their early stages and consequently can deter critical conditions for the power grid. A neural-fuzzy network is used to model the thermal condition of the power transformer in fault-free operation (the thermal condition is associated to a temperature variable known as hot-spot temperature). The output of the neural-fuzzy network is compared to measurements from the power transformer and the obtained residuals undergo statistical processing according to a fault detection and isolation algorithm. If a fault threshold (that is optimally defined according to detection theory) is exceeded, then deviation from normal operation can be detected at its early stages and an alarm can be launched. In several cases fault isolation can be also performed, i.e. the sources of fault in the power transformer model can be also identified. The performance of the proposed methodology is tested through simulation experiments.

Introduction

Power transformers are among the most expensive equipment of the electric power transmission and distribution system and their condition monitoring is important for the uninterrupted and reliable functioning of the power grid. Transformer life management has been a topic of intensive research during the last years because of the need for operating the electric power grid under more harsh conditions and because of the increased demand for electric energy. According to an IEEE survey, oil immersed transformer failure rate per year is 0.00625. Therefore, in a fleet of 100 transformers, ten will have problem in the next 16 years. According to an international survey conducted by CIGRE, typical failure rates for power transformers are in the range of 1–2% per year for the large power transformers (operating voltages up to 300 kV). Load growth has contributed to an increase of the transformer's Hot Spot Temperature (HST), i.e. of a parameter that is directly associated to the ageing of the transformer and to the probability of failures of the transformer's components. The average HST a few decades ago was 50 °C, while under present operating conditions it is around 73 °C [1,2].

Transformers operating beyond their ratings exhibit the following symptoms: (i) increase in temperature of windings, insulation and oil, (ii) increase in leakage flux density outside the core, causing additional eddy current heating in the metallic parts, (iii) moisture and gas content increases with the increase in temperature, (iv) bushings, tap-changers, and cables are exposed to higher stresses, (v) deterioration of the windings insulation appears due to higher thermal stresses [3–5]. Obviously, there is a significant safety and environmental risk of operating aged transformer units close to their loading limits without surveillance and assessment. On the other hand, on-line monitoring of power transformers can provide a clear indication of their status and ageing behavior. Analysis of critical parameters collected from power transformers allows avoidance of irreversible failures and permits preventive maintenance.

During the last years research efforts have been carried out to develop thermal models of improved accuracy for power transformers [6–9]. The load-current profile, the top-oil temperature profile and the weather conditions (ambient temperature, solar heating, wind speed, rain conditions, etc.) are among the parameters that influence the transformer's thermal behavior. As mentioned, significant indications about the thermal condition of a power transformer and the associated failure risks can be obtained through monitoring the transformer's Hot Spot Temperature (HST). A deviation of HST from the anticipated temperature profile is

* Corresponding author.
E-mail addresses: grigat@ieee.org (G. Rigatos), psiano@unisa.it (P. Siano).

http://dx.doi.org/10.1016/j.ijepes.2016.01.019
0142-0615/

probably an indication of ageing of the transformer or in some cases of pre-failure situations. Analytical, as well as numerical (neural/fuzzy) models for HST have been developed [10–12]. These models associate the hot spot temperature to inputs such as: the ambient temperature, the top oil temperature rise over the ambient temperature, and the maximum winding hot-spot rise over the top-oil temperature.

The development of a model of the power transformer's thermal behavior in the fault-free condition and the comparison of the outputs of such a model with online measurements of the real transformer output enables to implement fault detection and isolation (FDI). A statistical FDI method that can be used to find incipient failures in the transformer's components is the so-called Local Statistical Approach to change detection [13–17]. The proposed statistical fault diagnosis method can point out the existence of a fault through the processing of the residuals sequence, where the residuals are defined as the differences between the measured and the estimated HST values at every sampling instant. The proposed FDI method transforms the complex detection problem into the problem of monitoring the mean of a Gaussian vector. The local statistical approach consists of two stages: (i) the global χ^2 test which indicates the existence of a change in some parameters of the transformer's model, (ii) the diagnostics tests (sensitivity or min–max) which isolate the parameter affected by the change [13–17]. The local statistical approach is suitable for detecting incipient faults in the power transformer, thus enabling preventive maintenance.

The concept of the proposed FDI technique is as follows: the thermal profile of the fault-free power transformer system is learned by the neural-fuzzy network. At each time instant the neural network's output is compared to the real Hot-Spot Temperature of the power transformer. The difference between the real condition of the power transformer and the output of the neural network forms a residual. The statistical processing of a sufficiently large number of residuals through the aforementioned FDI method provides an index-variable that is compared against a fault threshold and which can give early indication about deviation of the transformer from the normal operating conditions. Therefore alarm launching can be activated at the early stages of power transformer failure, and repair measures can be taken. Under certain conditions (detectability of changes) the proposed FDI method enables also fault isolation, i.e. it makes possible to identify the source of fault within the power transformer model [20].

The current paper elaborates on and extends the results of [21]. The structure of the paper is as follows: in Section 'Transformers in the electric power grid' an overview of the main types of electric power transformers is given and their significance for the electric power grid is explained. The main types of failures in power transformers is overviewed. In Section 'Analytical thermal model of electric power transformers' the thermal model of oil-immersed electric power transformers is analyzed and the significance of the hot-spot temperature for condition monitoring and secure operation of power transformers is explained. In Section 'Neuro-fuzzy modeling of power transformers thermal condition' neuro-fuzzy modeling is proposed for describing the variations of the hot-spot temperature in electric power transformers as well as its dependency on parameters such as the top-oil temperature and the load current. In Section 'Fault diagnosis for electric power transformers' a systematic method is proposed for fault detection and isolation (FDI) in power transformers, through the monitoring of the variations of the hot-spot temperature. The considered FDI method is the *Local Statistical Approach* to fault diagnosis and is based on the generalized likelihood ratio criterion for change detection. It is explained that the method is suitable for incipient faults diagnosis and preventive condition monitoring. In Section 'Simulation tests' simulation experiments are performed to evaluate the efficiency of the proposed fault diagnosis method in detecting and isolating faults in power transformers. It is shown that despite the nonlinearities of the thermal model of the power transformer, the success rate of the proposed fault diagnosis method is remarkably high. Finally, in Section 'Conclusion' concluding remarks are stated.

Transformers in the electric power grid

Condition monitoring of transformers within the smart grid

Power transformers are the most expensive and strategic components of a power system. One can distinguish between several classes of power transformers using two major classification criteria. The first criterion has to do with the insulating material used in the transformer (e.g. oil-immersed, gas-immersed and dry-type transformers). The second criterion has to do with the incoming and outgoing voltage levels of the power transformers and their role in the electric power grid (e.g. power transformers in generation stations, power transformers in the transmission system, power transformers in the distribution system, distribution substation transformers or distribution network transformers) [1–5].

Fault detection and isolation (FDI) for power transformers aims at continuously assessing the transformer's condition through the monitoring of associated critical parameters and at determining if the transformer is on the verge of a failure (this can be due to an internal fault or due to aging). To implement FDI it is necessary to develop a model of the transformer's functioning that associates its internal state to environmental conditions thus (i) enabling the detection of incipient failures (ii) prohibiting the erroneous interpretation of the monitored critical parameters and (iii) avoiding the launch of false alarms (for example if the transformer is operating in a heat wave, its oil temperature could be expected to be unusually high, but the transformer's FDI system should ascribe the temperature rise to the environmental conditions rather than a transformer problem) [22,23].

Reasons for failures in electric power transformers

Common failures in power transformers are (see Fig. 1):

(i) *Insulation breakdown in windings*. As the transformer ages the windings insulation is weakened to the point that it can no longer sustain the mechanical stresses due to a fault (e.g. in case of a short circuit). Turn-to-turn insulation

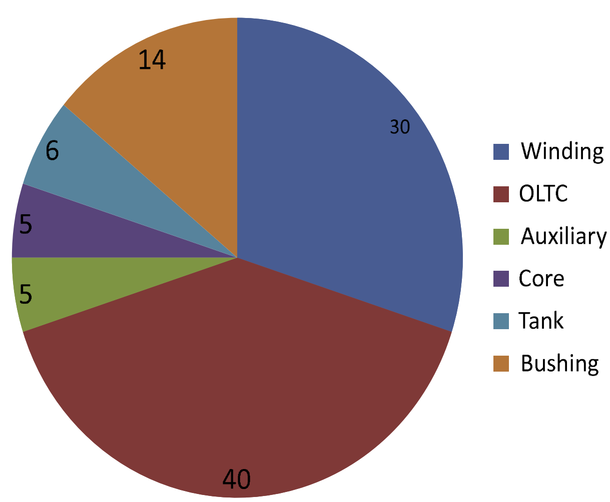


Fig. 1. Frequency of faults in components of electric power transformers [1].

suffers a dielectric failure. Winding failures cause in turn degradation of the overall insulation system, including the transformer's oil. Ageing of transformer oil is characterized by partial discharge (gas evolution starts from the oil) and thermal degradation (the raise of the transformer's temperature and the associated thermal stresses accelerate the oil's decomposition).

(ii) *The on-load tap changers (OLTC) failures*. On-load tap changers are used to change the tapping connection of the transformer's windings while the transformer is energized. The tap changers suffer from ageing. The insulating oil inside the tap changer becomes dirty due to switching arcs, which leads to weakened insulation properties. Another effect of the switching arcs is the wear of the arcing contacts. An additional ageing mechanism is the so-called long-term effect on the changeover selector. This effect starts with the formation of a thin layer of oil. The increased contact resistance due to the oil film layer can cause creation of hard and porous carbon material at places where the load current flows.

(iii) *Bushings failures*. An electrical bushing is an insulating structure including a through conductor or providing a central passage for such a conductor aiming at transmitting electric power in or out of the transformer. Oil-paper insulation is widely used in power transformer bushings. However, prolonged exposure to extreme electrical, thermal, mechanical and environmental stresses can deteriorate the insulation's condition and can break the cellulose bonds of the paper. This can lead to the formation of gas byproducts and bubbles and in turn can result in partial discharge and conducting tracts ending at shorting out one or more layers of the bushings.

The cost savings from performing preventive maintenance for power transformers (e.g. due to deterring critical conditions in the power grid and subsequent cascading events) can be significantly more important than the monitoring cost itself. Condition monitoring for power transformers is helpful in many aspects such as planning of maintenance schedules, obtaining knowledge of the health of equipment, estimating the remaining life of the equipment, finding areas of further improvement, and refining product specifications. Established methods for preventive maintenance of power transformers are shown in Fig. 2.

Analytical thermal model of electric power transformers

Thermal modeling of power transformers

The stages for obtaining an analytical model of the power transformer's thermal behavior are as follows [11]:

- Calculate at each time step the ultimate top oil temperature rise in the transformer from the load current at that instant, using:

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

  where $\Delta\Theta_{TO,U}$ is ultimate top oil temperature (TOT) rise [°C], $\Delta\Theta_{TO,R}$ is the rated TOT rise over ambient [°C], I_L is the load current normalized to rated current [p.u.], q is an empirically derived exponent to approximately account for effects of change of resistance with change in load, R is the ratio of rated-load loss to no-load loss at applicable tap position.
- Calculate the increment in the TOT from the ultimate top oil rise and the ambient temperature at each time step using the differential equation:

$$\tau_{TO}\frac{d\Theta_{TO}}{dt} = [\Delta\Theta_{TO,U} + \Theta_A] - \Theta_{TO} \tag{2}$$

  where Θ_{TO} is the TOT [°C], τ_{TO} is the top oil rise time constant, and Θ_A is the ambient temperature [°C].
- Calculate the ultimate hot spot temperature rise using:

$$\Delta\Theta_{HS,U} = \Delta\Theta_{HS,R} I_L^{2\beta} \tag{3}$$

  where β is an empirically derived exponent, dependent on the cooling method, $\Delta\Theta_{HS,U}$ is the ultimate HST rise over top oil (for a given load current) [°C], $\Delta\Theta_{HS,R}$ is the rated HST rise over top oil (for rated load current) [°C].
- Calculate the increment in the HST rise, using the differential equation:

$$\tau_{HS}\left\{\frac{d\Delta\Theta_{HS}}{dt}\right\} = \Delta\Theta_{HS,U} - \Delta\Theta_{HS} \tag{4}$$

  where Θ_{HS} is the hot spot winding temperature [°C], $\Delta\Theta_{HS}$ is the HST rise above top oil [°C], and τ_{HS} is the hot spot rise time constant [h].

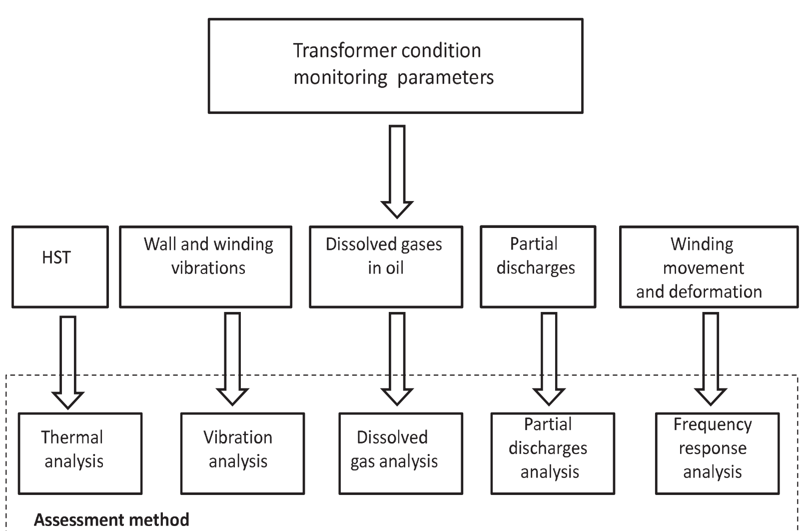


Fig. 2. Condition monitoring methods for power transformers and the associated processed parameters.

- Finally, add the TOT to the hot spot temperature rise to get the HST, using:

$$\Theta_{HS} = \Theta_{TO} + \Delta\Theta_{HS} \tag{5}$$

The model of Eqs. (1)–(5), named top-oil rise model, is based on some simplifying assumptions and its accuracy can deteriorate due to parameter variations. As a result, in order to protect power transformers, conservative safety factors have been introduced that prevent the transformer's overheating. Consequently, the calculated maximum power transfer may be 20–30% less or worse than the real transformer capability.

HST as an ageing and failure indication for power transformers

There are two types of aging in power transformers. The first one is the so-called intransitive aging and indicates the degradation of the transformers' components and particularly the degradation of the insulating material to withstand the designed stresses, such as electrical, mechanical, thermal and physical. The second one is the so-called transitive aging and denotes the rapid degradation of the transformer's components and particularly of the windings insulation due to abnormal operating conditions. The sustainable high Hot-Spot Temperature results in transitive transformer ageing.

Transitive ageing of electric power transformers can be detected by monitoring the transformer's Hot-Spot Temperature. The increase of the transformer HST accelerates the end of the transformer lifetime and vice versa. The relationship between the HST and the transformer life consumption is governed by the Arhennius reaction rate theory (IEEE Standard C57.91-1995) which states that

$$\begin{aligned} &\text{remaining life} = Ae^{B/T}, \quad \text{or} \\ &\text{per unit life} = Ae^{B/\Theta_{HS}+273} \end{aligned} \tag{6}$$

where A and B are empirical constants. The A and B constants are based on materials characteristics of the insulation and they are determined such that per unit life is unity at HST of 110 °C. Indicative values of A and B are 9.8×10^{-18} and 15×10^3. According to the previous formula one can calculate the lost lifetime of the transformer as a function of the HST. Thus, one sees that transitive ageing causes acceleration of the transformer's end of life which is mainly due to the increase of the HST.

Neuro-fuzzy modeling of power transformers thermal condition

The approach followed in this paper, for extracting a neuro-fuzzy model of HST variations, results in improved modeling of the transformer's thermal behavior. Neurofuzzy models have been also used in identification and fault diagnosis for nonlinear systems [24]. In the sequel, fuzzy rules of the Takagi–Sugeno type will be considered. These have the form:

$$\begin{aligned} &R_l: \ IF\ x_1\ is\ A_1^l\ AND\ x_2\ is\ A_2^l\ AND\ \cdots\ AND\ x_n\ is\ A_n^l \\ &\quad THEN\ \bar{y}^l = \sum_{i=1}^{n} w_i^l x_i + b^l \quad l = 1, 2, \ldots, L \end{aligned} \tag{7}$$

where R^l is the l-th rule, $x = [x_1, x_2, \ldots, x_n]^T$ is the input (antecedent) variable, $\bar{y}^l$ is the output (consequent) variable, and w_i^l, b^l are the parameters of the local linear models. The above model is a Takagi–Sugeno model of order 1. Setting $w_i^l = 0$ results in the zero order Takagi–Sugeno model [20]. The output of the Takagi–Sugeno model is given by the weighted average of the rules consequents (Fig. 3):

$$\hat{y} = \frac{\sum_{l=1}^{L} \bar{y}^l \prod_{i=1}^{n} \mu_{A_i^l}(x_i)}{\sum_{l=1}^{L} \prod_{i=1}^{n} \mu_{A_i^l}(x_i)} \tag{8}$$

where $\mu_{A_i^l}(x_i): R \to [0, 1]$ is the membership function of the fuzzy set A_i^l in the antecedent part of the rule R^l. In the case of a zero order TS system the output of the l-th local model is $\bar{y}^l = b^l$, while in the case of a first order TS system the output of the l-th local model is given by $\bar{y}^l = \sum_{l=1}^{L} w_i^l x_i + b^l$.

If the numerically extracted neural-fuzzy model does not approximate efficiently the monitored physical system then a refinement of the partitioning of the patterns space may be required, and improved placement of the centers of the Gaussian activation functions can be attempted [18,19]. The individual steps of data-driven fuzzy modeling for nonlinear function approximation are discussed in [20,25]. These stages are demonstrated in Fig. 4.

Fault diagnosis for electric power transformers

The global χ^2 test for fault detection

As shown in Fig. 5 the proposed method is based on the definition of the residual e_i described as the difference between the output from the power transformer y_i^0 and the output from a neural model y_i [26]. The neural/fuzzy model is used to simulate the power transformer in a fault-free state and has been extracted from input/output data when the power transformer operates normally. To perform also fault isolation, the real power system has been simulated by using the so-called *exact model*. In order to have the neural/fuzzy model and the exact model with the same number of parameters, the exact model can be also represented by a neural/fuzzy model extracted from input/output data of the power transformer. Therefore, when the power transformer is affected by slight parameters variations which can finally lead to failure, the output of the exact model will differ from the output of the neural model. In other words, while the neural model simulates the power transformer in an undistorted stable state, the exact model simulates the real power transformer in all conditions and is extracted from the real power transformer data (Fig. 5).

The partial derivative of the residual square is:

$$H(\theta, \hat{y}_i) = \frac{1}{2}\frac{\partial e_i^2}{\partial \theta} = e_i \frac{\partial \hat{y}_i}{\partial \theta} \tag{9}$$

where θ is the vector of model's parameters. The vector H having as elements the above $H(\theta, \hat{y}_i)$ is called primary residual. In the case of neuro-fuzzy models the gradient of the output with respect to the consequent parameters w_i^l is given by

$$\frac{\partial \hat{y}}{\partial w_i^l} = \frac{x_i \mu_{R^l}(x)}{\sum_{l=1}^{L} \mu_{R^l}(x)} \tag{10}$$

The gradient with respect to the center c_i^l is

$$\frac{\partial \hat{y}}{\partial c_i^l} = \sum_{l=1}^{L} \frac{\bar{y}^l \frac{2(x_i - c_i^l)}{v_i^l} \mu_{R^l}(x_i) \left[\sum_{j=1}^{L} \mu_{R^j}(x_i) - \mu_{R^l}(x_i)\right]}{\left[\sum_{l=1}^{L} \mu_{R^l}(x_i)\right]^2} \tag{11}$$

The gradient with respect to the spread v_i^l is

$$\frac{\partial \hat{y}}{\partial v_i^l} = \sum_{l=1}^{L} \frac{\bar{y}^l \frac{2(x_i - c_i^l)^2}{v_i^{l^3}} \mu_{R^l}(x_i) \left[\sum_{j=1}^{L} \mu_{R^j}(x_i) - \mu_{R^l}(x_i)\right]}{\left[\sum_{l=1}^{L} \mu_{R^l}(x_i)\right]^2} \tag{12}$$

Next, having calculated the partial derivatives of Eqs. (10)–(12), the rows of the Jacobian matrix J are found by

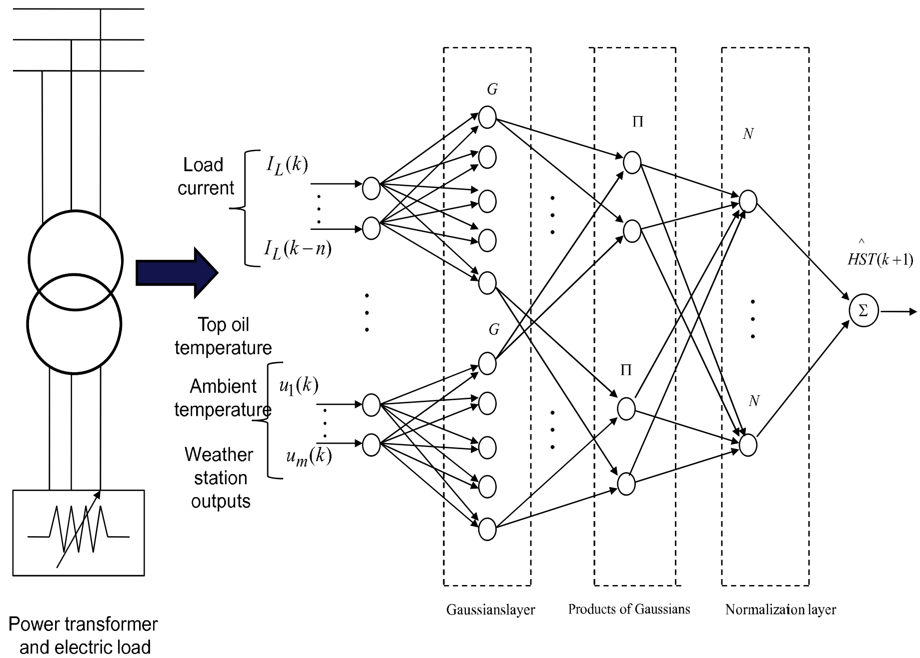


Fig. 3. Inputs/outputs configuration of the neural model of the power transformer thermal dynamics.

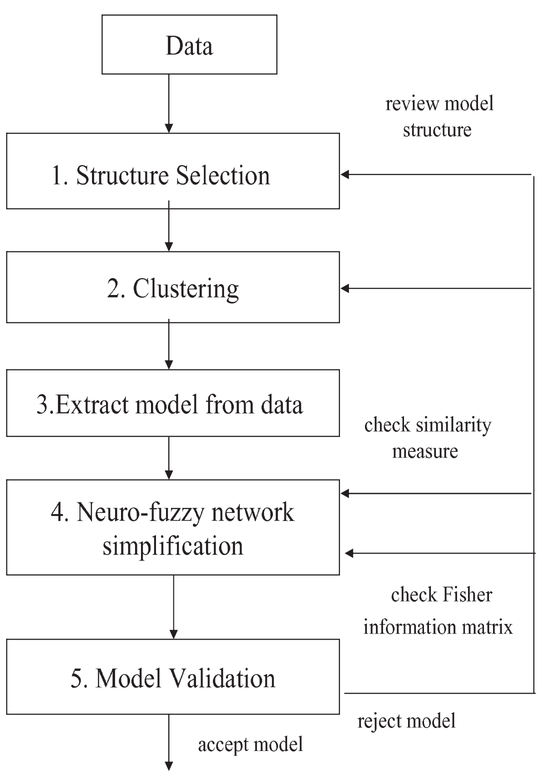


Fig. 4. General scheme of data-driven neural-fuzzy modeling.

$$J(\theta_0, y_k) = \left.\frac{\partial \hat{y}_k(\theta)}{\partial \theta}\right|_{\theta=\theta_0} \tag{13}$$

The problem of change detection with the χ^2 test consists of monitoring a change in the mean of the Gaussian variable which for the one-dimensional parameter vector θ is formulated as

$$X = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_k \frac{\partial \hat{y}_k}{\partial \theta} \sim N(\mu, \sigma^2) \tag{14}$$

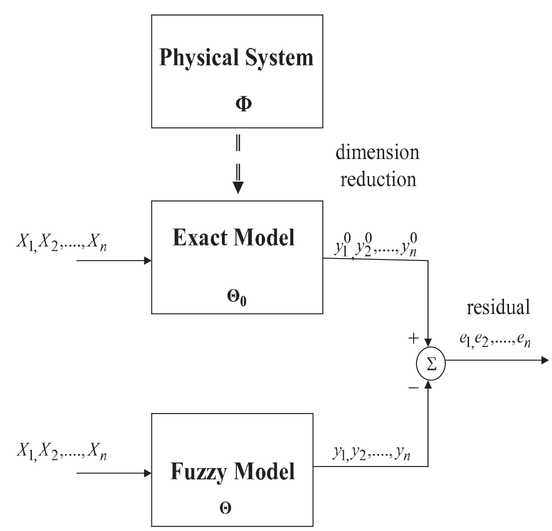


Fig. 5. Residual between the exact model and the neuro-fuzzy model.

where $\hat{y}_k$ is the output of the neural (exact) model generated by the input pattern x_k, e_k is the associated residual and θ is the vector of the model's parameters. For a multivariable parameter vector θ should hold $X \sim (M\eta, \Sigma)$. In order to decide if the power transformer is in fault-free operating condition for a given set of data of N measurements, one defines as θ_* the value of the parameters vector μ minimizing the root mean square error (RMSE) of the HST neural estimate. The notation is introduced only for the convenience of problem formulation, and its actual value does not need to be known. Then the model validation problem amounts to make a decision between the two hypotheses:

$$\begin{aligned} H_0 &: \ \theta_* = \theta_0 \\ H_1 &: \ \theta_* = \theta_0 + \frac{1}{\sqrt{N}} \delta\theta \end{aligned} \tag{15}$$

where $\delta\theta \neq 0$. It is known from the central limit theorem that for a large data sample, the normalized residual given by Eq. (14)

asymptotically follows a Gaussian distribution when $N \to \infty$ [13,27]. More specifically, the hypothesis that has to be tested is:

$$\begin{aligned} &H_0: \ X \sim N(0, S) \\ &H_1: \ X \sim N(M\eta, S) \end{aligned} \tag{16}$$

where M is the sensitivity matrix (see Eq. (17)), η is the parameters' vector and S is the covariance matrix (see Eq. (18)). The product $M\eta$ denotes the new center of the monitored Gaussian variable X, after a change on the system's parameter θ. The sensitivity matrix M of $\frac{1}{\sqrt{N}}X$ is defined as the mean value of the partial derivative with respect to θ of the primary residual defined in Eq. (9), i.e. $E\{\frac{\partial}{\partial\theta}H(\theta, y_k)\}$ and is approximated by [13,27]:

$$M(\theta_0) \simeq \frac{\partial}{\partial\theta}\frac{1}{N}\sum_{k=1}^{N}H(\theta_0, y_k) \simeq \frac{1}{N}J^TJ \tag{17}$$

The covariance matrix S is defined as $E\left\{H(\theta, y_k)H^T(\theta, y_{k+m})\right\}$, $m = 0, \pm 1, \ldots$ and is approximated by [13,27]:

$$\begin{aligned} S =\simeq \sum_{k=1}^{N}\left[H(\theta_0, y_k)H^T(\theta_0, y_k)\right] + \sum_{m=1}^{I}\frac{1}{N-m}\sum_{k=1}^{N-m}\Big[H(\theta_0, y_k)H^T(\theta_0, y_{k+m}) \\ + H(\theta_0, y_{k+m})H^T(\theta_0, y_k)\Big] \end{aligned} \tag{18}$$

where an acceptable value for I is 3. The tool to decide about the existence of a change (fault) is the likelihood ratio

$$\begin{aligned} &s(X) = ln\frac{p_{\theta_1(x)}}{p_{\theta_0(x)}}, \quad \text{where} \\ &p_{\theta_1}(X) = e^{[X-\mu(X)]^T S^{-1}[X-\mu(X)]} \end{aligned} \tag{19}$$

and $p_{\theta_0}(X) = e^{X^T S^{-1} X}$. The center of the Gaussian distribution of the changed system is denoted as $\mu(X) = M\eta$ where η is the parameters vector. The *Generalized Likelihood Ratio* (GLR) is calculated by maximizing the likelihood ratio with respect to η [13,27]. This means that the most likely case of parameter change is taken into account. This gives the global χ^2 test t:

$$t = X^TS^{-1}M(M^TS^{-1}M)^{-1}M^TS^{-1}X \tag{20}$$

Since X asymptotically follows a Gaussian distribution, the statistics defined in Eq. (20) follows a χ^2 distribution with n degrees of freedom. Mapping the change detection problem to this χ^2 distribution enables the choice of the change threshold. Assume that the desired probability of false alarm is α then the change threshold λ should be chosen from the relation $\int_\lambda^\infty \chi_n^2(s)ds = \alpha$, where $\chi_n^2(s)$ is the probability density function (p.d.f.) of a variable that follows the χ^2 distribution with n degrees of freedom.

Fault isolation with the sensitivity method

Fault isolation is needed to identify the source of faults in the electric power transformer model. A first approach to change isolation is to focus only on a subset of the parameters while considering that the rest of the parameters remain unchanged [13,27]. The parameters vector η can be written as $\eta = [\phi, \psi]^T$, where ϕ contains those parameters to be subject to the isolation test, while ψ contains those parameters to be excluded from the isolation test. M_ϕ contains the columns of the sensitivity matrix M which are associated with the parameters subject to the isolation test. Similarly M_ψ contains the columns of M that are associated with the parameters to be excluded from the sensitivity test.

Assume that among the parameters η, it is only the subset ϕ that is suspected to have undergone a change. Thus η is restricted to $\eta = [\phi, 0]^T$. The associated columns of the sensitivity matrix are given by M_ϕ and the mean of the Gaussian to be monitored is $\mu = M_\phi\phi$, i.e. $\mu = MA\phi$, $A = [0, I]^T$. Matrix A is used to select the parameters that will be subject to the fault isolation test. The rows of A correspond to the total set of parameters while the columns of A correspond only to the parameters selected for the test. Thus the fault diagnosis (χ^2) test of Eq. (20) can be restated as:

$$t_\phi = X^TS^{-1}M_\phi\left(M_\phi^TS^{-1}M_\phi\right)^{-1}M_\phi^TS^{-1}X \tag{21}$$

The min–max test

In this approach the aim is to find statistics that will be able to detect a change on the part ϕ of the parameters vector η and which will be robust to a change in the non observed part ψ [13,27]. Assume the vector partition $\eta = [\phi, \psi]^T$. The following notation is used:

$$M^TS^{-1}M = \begin{pmatrix} I_{\varphi\varphi} & I_{\varphi\psi} \\ I_{\psi\varphi} & I_{\psi\psi} \end{pmatrix} \tag{22}$$

$$\gamma = \begin{pmatrix} \varphi \\ \psi \end{pmatrix}^T \cdot \begin{pmatrix} I_{\varphi\varphi} & I_{\varphi\psi} \\ I_{\psi\varphi} & I_{\psi\psi} \end{pmatrix} \cdot \begin{pmatrix} \varphi \\ \psi \end{pmatrix} \tag{23}$$

where S is the previously defined covariance matrix. The min–max test aims to minimize the non-centrality parameter γ with respect to the parameters that are not suspected for change. The minimum of γ with respect to ψ is given for:

$$\psi^* = \arg\min_\psi \gamma = \varphi^T\left(I_{\varphi\varphi} - I_{\varphi\psi}I_{\psi\psi}^{-1}I_{\psi\varphi}\right)\varphi \tag{24}$$

and is found to be

$$\begin{aligned} \gamma^* = \min_\psi \gamma = \varphi^T\left(I_{\varphi\varphi} - I_{\varphi\psi}I_{\psi\psi}^{-1}\ I_{\psi\varphi}\right)\varphi = \begin{pmatrix} \varphi \\ -I_{\psi\psi}^{-1}I_{\psi\varphi}\varphi \end{pmatrix}^T \\ \begin{pmatrix} I_{\varphi\varphi} & I_{\varphi\psi} \\ I_{\psi\varphi} & I_{\psi\psi} \end{pmatrix}\begin{pmatrix} \varphi \\ -I_{\psi\psi}^{-1}I_{\psi\varphi}\varphi \end{pmatrix} \end{aligned} \tag{25}$$

which results in

$$\gamma^* = \varphi^T\left\{\left[I, -I_{\varphi\psi}I_{\psi\psi}^{-1}\right]M^T\Sigma^{-1}\right\} \cdot \Sigma^{-1}\left\{\Sigma^{-1}M\left[I, -I_{\varphi\psi}I_{\psi\psi}^{-1}\right]\right\}\varphi \tag{26}$$

The following linear transformation of the observations is considered:

$$X_\phi^* = \left[I, -I_{\varphi\psi}I_{\psi\psi}^{-1}\right]M^T\Sigma^{-1}X \tag{27}$$

The transformed variable X_ϕ^* follows a Gaussian distribution $N\left(\mu_\phi^*, I_\phi^*\right)$ with mean $\mu_\varphi^* = I_\varphi^*\varphi$ and with covariance $I_\varphi^* = I_{\varphi\varphi} - I_{\varphi\psi}I_{\psi\psi}^{-1}I_{\psi\varphi}$. The max–min test decides between the hypotheses:

$$H_0^*: \ \mu^* = 0 \text{ and } H_1^*: \ \mu^* = I_\varphi^*\varphi \tag{28}$$

and is described by:

$$\tau_\varphi^* = X_\varphi^{*\,T}I_\varphi^{*\,-1}X_\varphi^* \tag{29}$$

The stages of fault detection and isolation (FDI) with the use of the Local Statistical Approach are given in Table 1:

Simulation tests

First a neural network has been used for modeling the thermal dynamics of power transformers. Next fault diagnosis is performed with the use of the local statistical approach. A temperature signal is recorded from the power transformer with the use of suitable sensors. At a first stage the sequence of residuals is obtained, that is the differences between the monitored output of the transformer

Table 1
Stages of the local statistical approach for FDI.

1. Generate the residuals partial derivative given by Eq. (9)
2. Calculate the Jacobian matrix J given by Eq. (13)
3. Calculate the sensitivity matrix M given by Eq. (17)
4. Calculate the covariance matrix S given by Eq. (18)
5. Apply the χ^2 test for change detection of Eq. (20)
6. Apply the change isolation tests of Eq. (21) or Eq. (29)

and the output of the fault-free reference model which is provided by the neural network. Next, hypothesis testing about parametric change (failure) in the system is performed through the computation of the likelihood ratio for the sequence of the residuals, as described in Section 'Fault diagnosis for electric power transformers'. Actually, this computation is carried by considering the Taylor series expansion of the likelihood ratio. Using that neither the parameters subject to change are known, nor the magnitude of this change has been quantified, maximization of the likelihood ratio is performed with respect to the parameters vector (so as to capture the worst case fault). This maximization results into the Generalized Likelihood Ratio. Thus, one arrives at a transformed residuals signal which follows the χ^2 distribution of Eq. (20) and equivalently one computes an optimal criterion for parametric change detection according to the properties of the χ^2 distribution.

Following the experimentation procedure that was described in Fig. 3, a neuro-fuzzy model was used to estimate with high accuracy the winding HST of a laboratory prototype mineral-oil-immersed power transformer. The transformer main characteristics are resumed in Table Table 2. A measurement station has been set up consisting of thermocouples that were monitoring (a) the Hot Spot Temperature of the medium and voltage windings and (b) the Top Oil Temperature. The Hot Spot Temperature could have been also measured with optical fiber sensors. The manufacturer's specifications give, the most probable hot-spot position. A hall effect current transducer has been used in order to measure the load current.

Table 2
Ratings of the modeled power transformer.

Nameplate rating	25 kVA
Vprimary/Vsecondary	10 kV/400 V
Iron losses	195 W
Copper losses (full load)	776 W
Top oil temperature rise at full load	73.1 °C
Weight of core and coil assembly	136 Kg
Weight of oil	62 Kg
Density	757 Kg/m^3
Total weight	310 Kg
Length, width and height of tank	$64 \times 16 \times 80$ cm
Type of cooling	ONAN
Factory/year	MACE/87

Neuro-fuzzy modeling of the power transformer has been carried out. The electric power transformer was modeled with the use of a neuro-fuzzy network, having as output the estimated hot-spot temperature $HST(k)$ and as inputs past values of the top-oil temperature, e.g. $\Theta_{TO}(k-1), \Theta_{TO}(k-2)$ and past values of the load current. e.g. $I_L(k-1)$. Neural models with the same output, such $HST(k)$ and a larger number of inputs, i.e. including more past values of the top-oil temperature and of the load current, could be also considered. The transformer's model has been identified considering both a neural network with Hermite polynomial basis functions and a neuro-fuzzy network of the Takagi–Sugeno type. In the first case, a neural network with Hermite basis functions was used to model the variations of the power transformer HST. As shown in Fig. 6(a) thanks to the inherent multi-frequency characteristics of the Hermite polynomial basis functions, such a neural model can capture with increased accuracy spikes and abrupt changes in the HST profile [28,29]. In the second case, the neural/fuzzy TSK model of Fig. 3 was used to estimate the variations of HST. The obtained approximation is shown in Fig. 6 (b). To decide where the basis functions should be placed, the input space was segmented using the input dimension (grid) partition [20]. The RMSE of training the Hermite and the TSK model was of the order of 4×10^{-3}.

The size of the training set was 300. The LMS (Least Mean Square) algorithm was used for the adaptation of the linear weights $w_1^{(l)}$ [18,19]. This neurofuzzy model actually stands for a rule base that consists of 64 rules (3 input variables partitioned in 4 fuzzy subsets each). All fuzzy sets (Gaussian activation functions) were assumed to have the same spread.

The TSK fuzzy model of the transformer that was extracted from real power transformer HST, TOT and load current data consisted initially of 64 rules. It comprised 64 linear parameters (weights) and 12 nonlinear parameters (centers of fuzzy sets). The size of the model was reduced to 27 rules after omitting those rules which

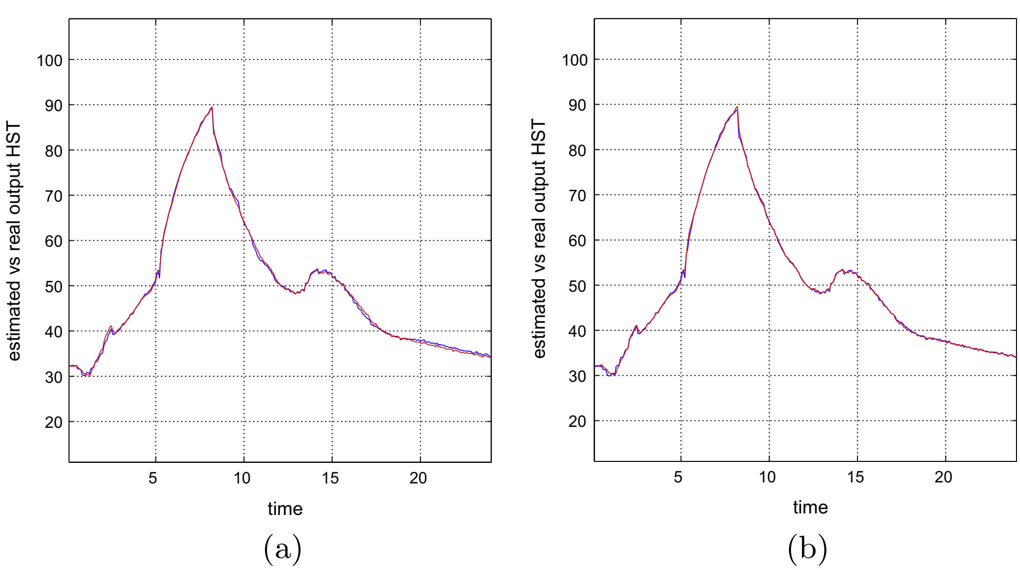


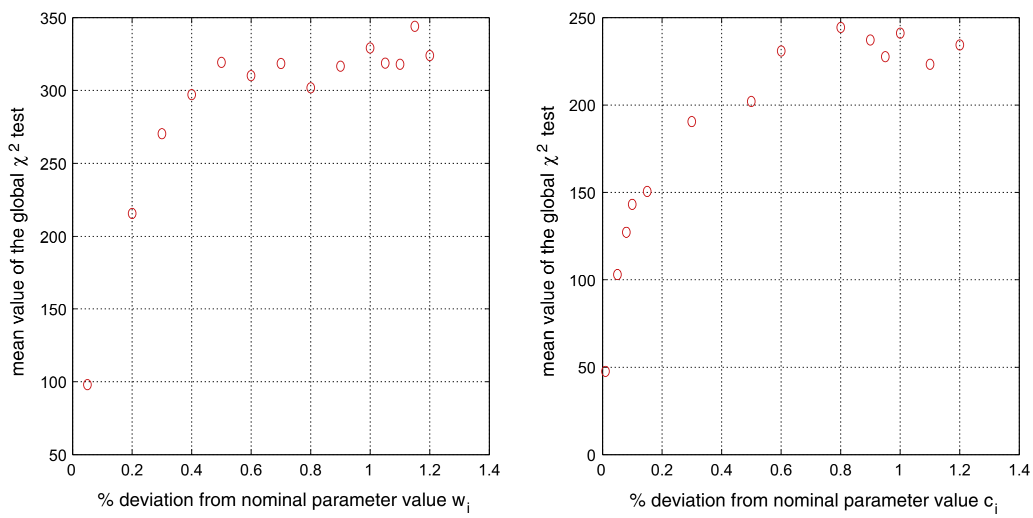
Fig. 6. Approximation of the Hot Spot Temperature of the electric power transformer (red line) (a) by a neural network with Hermite polynomial basis functions (blue-line) (b) by a fuzzy TSK network (blue-line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

received small activation from the existing data set. The parameters set in the new TSK fuzzy model consisted of 39 parameters (27 linear parameters which were the output layer weights and 12 nonlinear parameters which were the centers of the fuzzy sets). This means that by applying the local statistical approach to FDI and the χ^2 change detection test to the considered electric power transformer model, the fault threshold should be equal to 39.

The numerical tests confirmed theory. In case that no fault was assumed for the power transformer the mean value of the χ^2 test over a number of trials was found to be equal to 38.713. Such a value was anticipated according to the theoretical analysis of the χ^2 test. For slight deviations of the parameters of the power transformer from their nominal (fault-free) values, the global χ^2 test was capable of giving a clear indication about the existence of a fault. Thus for changes which varied between 0.01% and 1% of the nominal parameter's value (either for a linear or a nonlinear parameter) the score of the χ^2 test deviated significantly from the fault threshold (which as mentioned before was set equal to 39). As shown in Fig. 7, a small fault (deviation from the nominal parameter's value) suffices to generate an output of the χ^2 test that exceeds several times the value of the fault threshold.

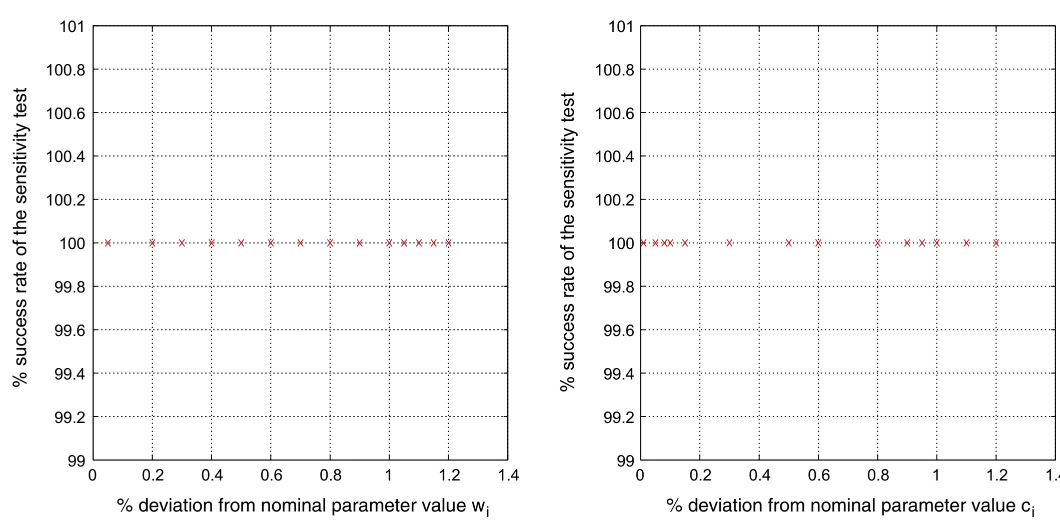
As far as fault isolation is concerned, the numerical results showed that the sensitivity method for fault isolation was very efficient in distinguishing the parameter subject to fault among all 39 parameters in the power transformer's parameter set. The sensitivity fault isolation test was performed for both a linear parameter (weight w_1) and for a nonlinear parameter (weight $c_1^{(1)}$) of the transformer's model. In Fig. 8 it can be observed that the success rate for the aforementioned fault magnitudes attained the value of 100%.

Finally, fault isolation tests for detecting changes in both linear (weight w_3) and nonlinear parameters (center $c_1^{(23)}$) of the power



(a) (b)

Fig. 7. (a) Score of the global χ^2 for changes in a linear parameter of the electric power transformer model, ranging between 0.01% and 1.0% of the nominal value. (b) Score of the global χ^2 for changes in a nonlinear parameter of the electric power transformer model, ranging between 0.01% and 1.0% of the nominal value.



(a) (b)

Fig. 8. Success rate of the fault isolation test (sensitivity method) (a) for changes in a linear parameter of the electric power transformer model, ranging between 0.01% and 1.0% of the nominal value (b) for changes in the a nonlinear parameter of the electric power transformer model, ranging between 0.01% and 1.0% of the nominal value.

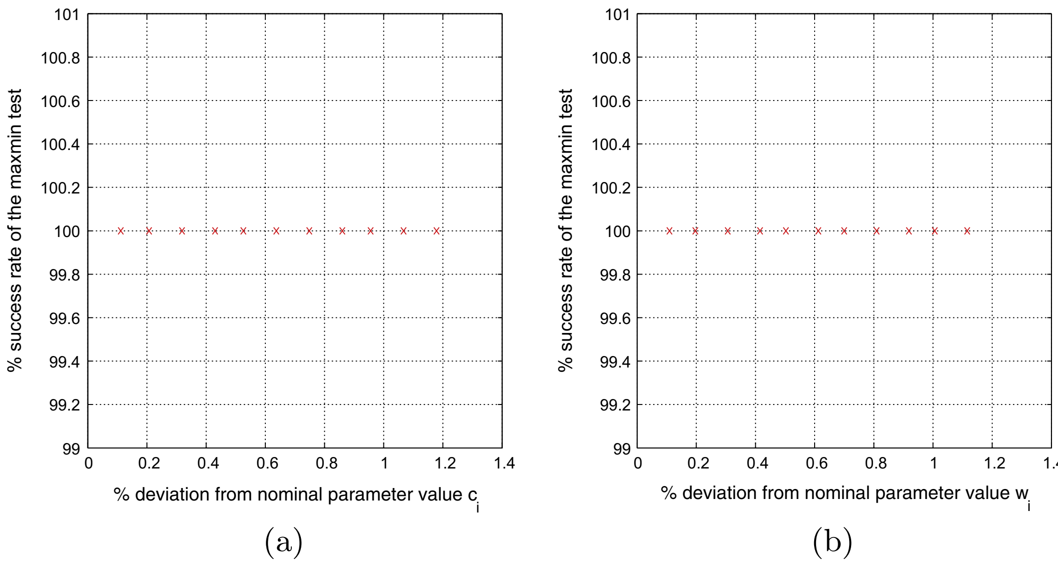


Fig. 9. Success rate of the fault isolation test (max–min method) (a) for changes in a linear parameter of the electric power transformer model, ranging between 0.01% and 1.0% of the nominal value (b) for changes in the a nonlinear parameter of the electric power transformer model, ranging between 0.01% and 1.0% of the nominal value.

transformer model were also performed with the use of the max–min method. The associated results are depicted in Fig. 9. It can be observed that the max–min fault isolation method succeeded also a high success rate in finding the faulty parameter in the transformer's thermal model.

Measurement noise has been considered in the development and testing of the fault diagnosis method. In the fault-free case the output of the power transformer and the output of its fault-free model (that is provided by the neural network) should be different only by the amount of the measurement noise affecting the sensors which measure the hot-spot temperature in the transformer. The sequence of the outputs difference (residuals) follows a Gaussian distribution centered at zero, while its weighted square follows a χ^2 distribution. The same statistical properties hold for the modified residuals sequence which is based on the Taylor series expansion of likelihood ratio and Eq. (14). On the other hand, in the case of parametric change there will be a shift to a non-zero-value for the mean value of the residuals signal and for the modified residuals signal. This is explained in Eq. (16). This shift is due to the structural (parametric) change in the transformer and not due to the existence of measurement noise.

It is also noteworthy that the proposed fault diagnosis method (local statistical approach) is suitable for the detection of incipient faults, that is small parametric changes and minimal deviation of the components of the power transformer from their nominal values. This is because,as noted above the local statistical approach is based on the concept of Taylor series expansion of the likelihood ratio that is computed with the sequence of measurements obtained from the monitored system. The Taylor expansion round the point (vector) of the nominal values of the system's parameters remains valid only for small parametric changes. Therefore, the mathematical developments which lead final to the χ^2 statistical change detection test of Eq. (20) hold if the magnitude of parametric changes is small. In most dynamical systems faults are driftwise instead of stepwise. This means that they progress slowly until they reach a point in which the damage is not-reversible and the system enters a critical condition. Thus, by being able to detect small parametric changes (incipient faults) one can initiate restoration measures for the system at a stage where the failure is still manageable and reversible. The benefits from such an early change detection approach are easy to understand.

Conclusion

A new method for condition monitoring and early fault diagnosis of electric power transformers has been introduced. The paper has proposed (i) to use neural-fuzzy networks for modeling the temperature behavior of an electric power transformer and particularly the variations of a parameter known as Hot Spot Temperature (HST), (ii) to use a statistical fault diagnosis method (Local Statistical Approach to Fault Diagnosis) that enables the early detection of failures (changes) in the transformer's model. A neural/fuzzy model was extracted from data which were obtained from the operation of a real power transformer. Next, the efficiency of the proposed FDI method was tested through the statistical processing of the residuals, i.e. of the sequence of the differences between the measured output of the real and the estimated output of the power transformer. The significance of the obtained results for early fault diagnosis in power transformers and consequently for preventive maintenance of these expensive components of the electric power grid is obvious. By monitoring the variation of the Hot-Spot Temperature (HST) one can have early indications about the occurrence of power transformer failures and can proceed to repair actions before critical conditions emerge. The proposed FDI method can be applied for condition monitoring of more critical components of the electric power grid and can help to maintain the reliable operation of the electric power transmission and distribution system.

References

[1] Metwally IA. Failures monitoring and new trends of power transformers. IEEE Potentials 2011;30(3):36–43.
[2] Arshad M, Islam SM. A novel fuzzy logic technique for power transformer asset management. In: IEEE industry applications conference, conference record of the 2006 41st IAS annual meeting, October 2006, Tampa, Florida, USA. p. 276–86.
[3] Abu-Elanien AEB, Salama MMA. Asset management techniques for transformers. Electr Power Syst Res 2010;80:456–64.
[4] Lin CH, Wub CH, Huang PZ. Grey clustering analysis for incipient fault diagnosis in oil-immersed transformers. Expert Syst Appl 2009;36:1371–9.
[5] Velasquez-Contreras JL, Sanz-Bobi MA, Arellano SG. General asset management model in the context of an electric utility: application to power transformers. Electr Power Syst Res 2011;81:2015–37.

[6] Jardini JA, Brittes JL, Magrini LC, Bini MA, Yasuoka J. Power transformer temperature evaluation for overloading conditions. IEEE Trans Power Deliv 2005;20(1):179–84.
[7] Susa D, Nordman H. Dynamic thermal modelling of power transformers. IEEE Trans Power Deliv 2005;20(1):197–204.
[8] Susa D, Nordman H. Dynamic thermal modelling of power transformers – further development: Part I. IEEE Trans Power Deliv 2006;21(4):1961–70.
[9] Susa D, Nordman H. Dynamic thermal modelling of power transformers – further development: Part II. IEEE Trans Power Deliv 2006;21(4):1971–80.
[10] Galdi V, Ippolito L, Piccolo A, Vaccaro A. Neural diagnostic system for transformer thermal overload protection. IEE Proc Power Appl 2000;147(5):415–21.
[11] Ippolito L, Siano P. Identification of Tagaki–Sugeno–Kang fuzzy model for power transformers predictive overload system. IEE Proc – Gener Transm Distrib 2004;151(5):582–9.
[12] Piccolo A, Siano P, Rigatos G. An adaptive framework for power components dynamic loadability. In: Rigatos GG, editor. Modelling and automation of intelligent industrial systems. IGI Publications; 2010.
[13] Basseville M, Nikiforov I. Detection of abrupt changes. Prentice Hall; 1993.
[14] Benveniste A, Basseville M, Moustakides G. The asymptotic local approach to change detection and model validation. IEEE Trans Autom Contr 1987;32(7):583–92.
[15] Basseville M, Benveniste A, Zhang Q. Surveilliance d' installations industrielles: démarche générale et conception de l' algorithmique, IRISA Publication Interne No 1010; 1996.
[16] Zhang Q. Fault detection and isolation with nonlinear black-box models. In: Proc SYSID 1997, Kitakyushu, Japan; 1997.
[17] Zhang Q, Basseville M, Benveniste A. Fault detection and isolation in nonlinear dynamic systems: a combined input-output and local approach. Automatica 1998;34(11):1359–73.
[18] Harris CJ, Hang X, Gon Q. Adaptive modelling, estimation and fusion from data: a neurofuzzy approach. Springer; 2002.
[19] Wang LX. A course in fuzzy systems and control. Prentice-Hall; 1998.
[20] Rigatos G, Zhang Q. Fuzzy model validation using the local statistical approach. Fuzzy Sets Syst 2009;60(7):882–904.
[21] Rigatos G, Siano P, Piccolo A. Incipient fault detection for electric power transformers using neural modeling and the local statistical approach to fault diagnosis. In: IEEE SAS 2012, 2012 IEEE sensors applications symposium, University of Brescia, Italy; February 2012.
[22] Catterson VM, Davidson EM, McArthur SDJ. Embedded intelligence for electrical network operation and control. IEEE Intell Syst 2011;26(2):38–45.
[23] Catterson VM, Davidson EM, McArthur SDJ. Agents for active network management and condition monitoring in the smart grid. In: Proc of the first intl workshop on agent technologies for energy systems, ATES, 2010, Toronto, Canada; 2010.
[24] Namburu SM, Azam MS, Luo Jianhui, Choi Kihoon, Pattipati KR. Data-driven modeling, fault diagnosis and optimal sensor selection for HVAC chillers. IEEE Trans Automat Sci Eng 2007;4(3):469–73.
[25] Rigatos GG, Siano P, Piccolo A. A neural network-based approach for early detection of cascading events in electric power systems. IET J Gener Transm Distrib 2009;3(7):650–65.
[26] Isermann R. Fault-diagnosis systems: an introduction from fault detection to fault tolerance. Springer; 2006.
[27] Zhang Q, Basseville M, Benveniste A. Early warning of slight changes in systems. Automatica 1994;30(1):95–113 [Special Issue on Statistical Methods in Signal Processing and Control].
[28] Zhang Q, Benveniste A. Wavelet networks. IEEE Trans Neural Netw 1993;3(6):869–98.
[29] Rigatos GG, Tzafestas SG. Neural structures using the eigenstates of the quantum harmonic oscillator. Open Syst Inform Dyn 2006;13(1).