Application of Classifiers for On-line Monitoring of Transformer Winding Axial Displacement by Electromagnetic Non-destructive Testing

M. A. Hejazi\textsuperscript{a}; G. B. Gharehpetian\textsuperscript{a}; G. R. Moradi\textsuperscript{a}; M. Mohammadi\textsuperscript{b}; H. A. Alehoseini\textsuperscript{a}

\textsuperscript{a} Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran \textsuperscript{b} School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran

Online publication date: 19 February 2011


To link to this Article DOI: 10.1080/15325008.2010.528537

URL: http://dx.doi.org/10.1080/15325008.2010.528537

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Application of Classifiers for On-line Monitoring of Transformer Winding Axial Displacement by Electromagnetic Non-destructive Testing

M. A. HEJAZI,1 G. B. GHAREHPETIAN,1 G. R. MORADI,1 M. MOHAMMADI,2 and H. A. ALEHOSEINI1

1Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran
2School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran

Abstract Transformer winding on-line monitoring using electromagnetic non-destructive testing has been suggested in this article. As a test object, a simplified model of transformer has been used. The winding axial displacement can be modeled on the test object. The scattering parameters of the test object can be measured and stored in a database. To detect the axial displacement, two indices have been defined using the magnitude and phase of scattering parameters. The $k$-nearest neighbor and decision tree classifiers have been used for the detection of the winding axial displacement and its value. The accuracy of the $k$-nearest neighbor method to find the axial displacement value has been improved by using a proposed $k$-nearest neighbor regression algorithm. The comparison of the average error of two classifiers shows the superiority of the $k$-nearest neighbor regression over the decision tree classifier.

Keywords transformer winding axial displacement, on-line monitoring, electromagnetic non-destructive testing, scattering parameters, $k$-nearest neighbor, decision tree

1. Introduction

Due to short circuits, mechanical forces can mainly cause axial displacement and/or radial deformation of transformer windings. These mechanical damages may not result in an immediate failure of the transformer, but the ability of transformer to resist against future dielectric and mechanical stresses may be highly decreased [1, 2].

There are several transformer mechanical damages monitoring and diagnostic methods. Each method has its advantages and disadvantages [3, 4]. In recent years, several off-line methods, such as the short-circuit test (SC) [1], low-voltage impulse (LVI) [5], and frequency response analysis (FRA) [3], have been proposed for the detection of winding displacement.

In the SC test method, the SC reactance is measured while the transformer is disconnected. In this method, the sensitivity of the reactance to the winding displacement
is very low, and the type and location of the mechanical damage in the winding cannot be determined [3].

In the FRA method, experimental approaches of comparison are time based, type based, and construction based. A prerequisite of all three methods is the independency of the measurement from the setup to create reproducible results. The FRA method can be used off-line and on-line [6, 7]:

- In the off-line method, for which the transformer is out of operation, the transformer is switched on and off on the high-voltage (HV) side. Thereby, the transformer is usually disconnected from the power network on the low-voltage (LV) side [6]. The well-known FRA method has been carried out off-line. The off-line methods will not meet all the needs of the transformer monitoring system.

- In the on-line method, the frequency response should be measured during the operation of the transformer. The stochastic transient over-voltages caused by switching and the lightning can be used to determine the transfer function. This method does not require a switching of the transformer and has the benefit of continuous monitoring of the transformer winding. Many factors affect this method, such as the response of arresters and different power system topologies. The measurement timing depends on the time of occurrence of the over-voltage transients [8]. This method is in the research phase and has not been used for any transformer.

Compared to the off-line methods, an on-line method has the advantage of stationary installation and, hence, an improved reproducibility of the test. The merit of an on-line monitoring method over off-line methods is the prediction of the important fault before its occurrence. On-line monitoring is very valuable because it not only decreases the direct damages to the equipment and possible injuries of people, but it can also reduce the collateral damages of a transformer failure which is due to several reasons, such as indirect damages caused by loss of energy (for example, destruction due to a stop of the process in chemical plants), loss of production capacity of a power station, and penalties due to undelivered electrical energy. Especially for aged transformers, and generally at strategic positions in the electrical network, on-line monitoring is necessary and valuable, because major failures costs for outages, repair, and associated damages can be prevented [9].

In this article, a new on-line monitoring method has been proposed that is based on the measurement of scattering parameters of the winding. This method, as in the FRA method, is based on the comparison of results. The simulations have shown that the scattering parameter of the winding can be used as an index for the detection of winding axial displacement [10]. In this study, the scattering parameter of a simplified model of a transformer has been measured using a network analyzer. In the article, two indexes have been suggested for the detection of the axial displacement of the winding as well. The $k$-nearest neighbor ($k$-NN) and decision tree (DT) classifiers [11, 12] have been used for the estimation of the winding axial displacement value. Also, these classifiers have been compared with each other.

2. Proposed Diagnosis Method Based on Electromagnetic Non-destructive Testing (NDT)

Electromagnetic NDT techniques have been widely used for many fault detection applications, such as dielectric property measurement of materials, thickness measurement of
Transformer On-line Monitoring Using S Parameters

Dielectric slabs, and surface-breaking crack detection in metals [13, 14]. This technique is contactless and based on the measurement of the reflection coefficient (magnitude and phase), which is the ratio of the reflected wave from the material under test to the transmitted wave. The aim of this method is retrieving information (detection, location, shape, and size evaluation) through the near-field measurement of the reflection coefficient of an open-ended rectangular waveguide. In [15], a semi-empirical signal processing technique based on the use of ANN was used for data treatment.

In the above-mentioned method, a single antenna excited in the TE01 mode is displaced over the material under test, but in some cases, such as in large structures (for example, in a power transformer with unknown position of the mechanical damage and in a harsh environment), it is impossible to access the surface of the device under test directly so it is impossible to scan it and detect the fault with the above-mentioned method. A new idea for detecting mechanical damage is using multiple antennas and measuring the scattering parameter, which is the generalized form of the reflection coefficient for an \( N \)-port microwave network (Figure 1). The power can be injected or absorbed in all terminals. As a result, there are \( N \) incoming waves and \( N \) outgoing waves [16].

The \( N \) incoming waves are usually designated by a vector of \( N \) complex quantities \( V_N^- \), and the \( N \) outgoing waves are designated by a vector of \( N \) complex quantities \( V_N^+ \). The relationship between these two vectors can be expressed by Eq. (1):

\[
\begin{bmatrix}
V_1^- \\
V_2^- \\
\vdots \\
V_N^-
\end{bmatrix} =
\begin{bmatrix}
S_{11} & S_{12} & \cdots & S_{1N} \\
S_{21} & S_{22} & \cdots & S_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
S_{N1} & S_{N2} & \cdots & S_{NN}
\end{bmatrix}
\begin{bmatrix}
V_1^+ \\
V_2^+ \\
\vdots \\
V_N^+
\end{bmatrix}.
\]

(1)

In Eq. (1), the elements of this frequency-dependent matrix, which are termed “S parameters,” determine the behavior of the network. In this research, there is only one rectangular aperture. As a result, the \( S \)-matrix dimension is \( 1 \times 1 \). The single element of the \( S \)-matrix, which is equal to the reflection coefficient, is easily determined by the following equation:

\[
\Gamma = S_{11} = \frac{V_1^-}{V_1^+}.
\]

(2)

This can be written in the following simple form:

\[
\Gamma = S_{11} = \sqrt{\frac{P_{\text{ref}}}{P_{\text{in}}}}.
\]

(3)

where \( P_{\text{in}} \) is the transmitting power and \( P_{\text{ref}} \) is the receiving power of the antenna.

The principle of this method is based on the measurement of the magnitude and phase of scattering parameters measured by several antennas. In this article, the idea is the application of the antenna in both transmitting and receiving modes. The reflected wave from the inside of the transformer can be received by the antenna. These antennas can be placed inside or outside the tank. If the antenna is placed outside the tank, then it needs a dielectric window, which forms a robust electrical aperture for sending and receiving very high-frequency electromagnetic waves. Figure 2 shows an externally mounted UHF sensor on the tank of a transformer [17].
Figure 1. $N$-port network.

Figure 2. Two views of an externally mounted UHF sensor on the tank of a transformer [13].
This UHF sensor has been used for partial discharge detection inside the transformer, and the antenna works in the receiving mode. Therefore, the antenna used for the detection of partial discharge can have another application, which is finding the axial displacement of the winding. In this article, the data of the normal condition (scattering parameter) should be measured and stored in a database as a base case. Then, the new set of data can be measured in specified intervals and compared with the base case scattering parameter, which is defined as the fingerprint of the transformer. In the last step, an expert system can be used to determine the occurrence, the type, and the value of the axial displacement.

The proposed method is a new method to determine axial displacements, which is based on comparisons like the FRA method that compares the measured transfer function with a reference one [3]. It should be noted that the proposed method can be used on-line without disconnecting the transformer. In this article, the comparison approach of the scattering parameters is time based. This technique is also the most accurate in the FRA results comparisons [3]. In the time-based approach, the measured parameters are compared with the latest measurement results. In the proposed method, the time-based comparison can be accurately used, because this method can be used on-line and the intervals between measurements can be decreased to have more information from the same transformer over a period of time.

3. Measurement Setup

The simplified model of a transformer used in this work is shown in Figure 3. The dimensions of the model are approximately 5% of a real transformer. The model is based on an ideal metallic cylinder in a metallic box. The inner metallic cylinder represents the LV winding, while the outer metallic cylinder represents the HV winding.

The antenna used in the model is a rectangular aperture with the dimensions of a standard X-band waveguide (WR90). The excitation frequency should be between the first and second cut-off frequencies (6.6 and 14.7 GHz for the model) in order to have a single mode of propagation. From the electromagnetic wave point of view, if the length of a hole is less than 0.1 of the wavelength, then the wave cannot enter the hole. The wavelength at the middle of the band is 31 mm, which is greater than one-tenth of the distance between the HV winding disks in the reduced scales (approximately 6 mm × 0.05 = 0.3 mm). As a result, considering the dimensions of the HV winding of a power transformer and the frequency of the electromagnetic wave, the inner parts (core and LV winding) can be neglected, and only the outer surface of HV winding should be modeled, as shown in Figure 3. Figure 4 is a photo of the used model.

The scattering parameter of the model has been measured using a network analyzer. The excitation frequency has been changed from 7 to 12 GHz in 250-MHz steps. It should be mentioned that in a power transformer, the dimensions are 20 times larger, i.e., in the UHF frequency band.

The cylinder has been shifted axial in 0.1-mm steps and can have 94 different positions. In each position, the magnitude and the phase of scattering parameters have been measured and stored. The feature vector used for the classification contains both the magnitude and the phase at \( n = 201 \) frequencies (points). Figures 5 and 6 show the magnitude and the phase of the scattering parameters over the frequency band or the reference position (normal condition) and for four other positions. The cylinder has been shifted in steps of 2.5% of the transformer height. Figures 7 and 8 show the magnitude and phase of scattering parameters for steps of 0.25% of the transformer height.
Figure 3. Simplified model of transformer.

Figure 4. Measurement setup and network analyzer.
Figure 5. Magnitude of scattering parameters over frequency band for reference position and for four other cylinder positions.

Figure 6. Phase of scattering parameters over frequency band for reference position and for four other cylinder positions.

Figure 7. Magnitude of scattering parameters for steps of 0.25% of the transformer height.
It is obvious that the proposed method has a good sensitivity to displacements higher than 0.25% of the test object height. In the FRA method, the axial displacement sensitivity is limited to 1.2% of the winding height [4].

4. Detection of Winding Axial Displacement Using Suggested Indices

In this article, the following two indices have been used for fault detection:

\[
MED(x) = \sqrt{\frac{\sum_{i=1}^{n}(|S_i(x)| - |S_i(0)|)^2}{n}},
\]

\[
PED(x) = \sqrt{\frac{\sum_{i=1}^{n}(\angle S_i(x) - \angle S_i(0))^2}{n}}.
\]

In the above-mentioned equations, \(|S_i(x)|\) and \(\angle S_i(x)\) are the amplitude and phase of the scattering parameters for the position \(x\); and at frequency \(i\), \(|S_i(0)|\) and \(\angle S_i(0)\) are the amplitude and phase of the scattering parameters for the reference position at frequency \(i\). \(MED\) is the magnitude Euclidean Distance (ED), and \(PED\) is the phase EDs of \(S\) parameters at \(n\) different frequencies measured at position \(x\) from the reference position. Figures 9 and 10 show \(MED\) and \(PED\) versus displacement (\(x\)) for 201 frequency points (from 7 to 12 GHz). It is assumed that the central position of the cylinder is the reference position and that the cylinder has a positive or negative (upward or downward, respectively) axial displacement. The increase of these indices is proportional to the increase of the axial displacement. As a result, they can be used as an index to detect the axial displacement and estimate the displacement value, approximately. These indices cannot discriminate between the upward and downward displacement, i.e., \(MED\) and \(PED\) are even functions of the displacement.
5. Axial Displacement and Its Value Determination Using Classifiers

In the previous section, it has been shown that the proposed method can detect the axial displacement. The next step is determining the displacement value, which can be important for utility engineers. The suggested method of this article is based on a pattern recognition system that compares the new measured $S$ parameters with the information in the database by using the $k$-NN and DT classifiers.

5.1. $k$-NN Classifier

The nearest neighbor classifier (NNC) is a simple, popular, and efficient classification algorithm, which assumes that the class label of an unknown pattern is the same as that
of its nearest neighbor. When the number of neighbors is more than one, it extends to the k-NN classifier [18–20].

The data is divided into two sets: the training set and the testing set. The training set is stored in memory and is used for classification, while the testing set is used to determine the classification accuracy. The performance of classifier is determined according to the classification accuracy [18].

The training samples are denoted by $n$-dimensional numeric attributes. The $k$-NN algorithm searches the $k$ training samples that are adjacent to the unknown sample given for the classification process. The closeness is measured in terms of EDs and can be calculated as follows:

$$ED(Z, Y) = \sqrt{\sum_{i=1}^{n} (z_i - y_i)^2}. \quad (6)$$

The unknown sample is classified according to the most-occurring class among the $k$-NNs. The $k$ values are chosen as odd numbers in the case of binary classification, so as to avoid the tied votes [21]. The best choice of $k$ is dependent upon the data; larger values of $k$ usually decrease the effect of noise on the classification but make the boundaries between classes ambiguous [22].

In [23], for instance, the distance between the unlabeled sample and each of its $k$ nearest training samples is computed and then used as a weight to influence the global decision process. Other variants make use of fuzzy set and Dempster–Shafer theories to better exploit information conveyed by the set of $k$ nearest training samples [24, 25].

The $k$-NN algorithm is described as follows [26]:

(i) Training algorithm:

Store the pattern $p_i (1 \leq i \leq n)$ with the class label $l(p_i)$ into the training database.

(ii) Classification algorithm:

Step 1: Give a query pattern $p_q$ to be classified.

Step 2: Compute $ED(p_q, p_i)$, which is the $ED$ between the query pattern $p_q$ and each pattern $p_i$ in the training database.

Step 3: Select the $k$-NNs \{$p_{i1}, p_{i2}, \ldots, p_{ik}\}$ according to the $ED$.

Step 4: According to Eqs. (5) and (6), compute $l(p_q, p_j)$, which is the membership degree of the query pattern $p_q$ belonging to each class $c_j (c_j \in \Omega)$, as follows:

$$l(p_q, c_j) = \frac{\sum_{i=1, \ldots, c} \delta(c_j, l(p_i))}{k}, \quad (7)$$

where $p_i \in \{p_{i1}, p_{i2}, \ldots, p_{ik}\}$ is one of the nearest neighbors of the query pattern $p_q$ and $\delta$ is a Kronecker delta function, defined as follows:

$$\delta(a, b) = \begin{cases} 1 & a = b \\ 0 & \text{otherwise} \end{cases}. \quad (8)$$

Step 5: According to Eq. (7), the class with the largest membership degree is assigned to the query pattern $p_q$:

$$l(p_q) = \{c_j \mid l(p_q, c_j) = \max[l(p_q, c_i)], \quad c_i \in \Omega\}. \quad (9)$$
The k-NN error varies with the value of k when the number of training patterns is fixed. Therefore, the combination of multiple k-NNs with different k may improve the accuracy of the whole system [26].

5.2. k-NN Regression

The same method can be used for regression by simply averaging of the values of its k-NNs. It can be useful to weight the contributions of the neighbors so that the nearer neighbors contribute more to the average than do the more distant ones. The proposed method can be summarized as follows:

(i) Training algorithm:
Store the pattern \( p_i \) (1 \( \leq \) i \( \leq \) n) with the class label \( l(p_i) \) into the training database.

(ii) Classification algorithm:
Step 1: Give a query pattern \( p_q \) to be classified.
Step 2: Compute \( ED(p_q, p_i) \), which is the \( ED \) between the query pattern \( p_q \) and each pattern \( p_i \) in the training database.
Step 3: Select the k-NN matrix of patterns \( p_i \in \{ p_{i_1}, p_{i_2}, \ldots, p_{i_k} \} \) and its relating class label matrix \( l(p_i) = [l(p_{i_1}), l(p_{i_2}), \ldots, l(p_{i_k})] \) according to the \( ED \).
Step 4: Compute the weighting matrix according to Eq. (10):

\[
W = \left[ 1/ED(p_q, p_{i_1}), 1/ED(p_q, p_{i_2}), \ldots, 1/ED(p_q, p_{i_k}) \right]
= [W_1, W_2, \ldots, W_k].
\] (10)

Step 5: Compute the class label of the query pattern by Eq. (11):

\[
l(p_q) = \frac{\sum_{j=1}^{k} w_j l(p_{i_j})}{\sum_{j=1}^{k} w_j}.
\] (11)

5.3. DT Classifier

The DT is a non-parametric learning technique that is able to produce classifiers about a given problem in order to deduce information for new unobserved cases. The DT has the hierarchical form of a tree structured upside down [27]. A DT model employs a recursive divide-and-conquer strategy to divide the dataset into partitions so that all of the records in a partition have the same class label. Finding the best split point and performing the split are the main tasks in the DT [28].

The type of DT developed in this research is the classification and regression tree (CART) introduced by Breiman et al. [12]. The DT has several advantages to other methods of classification. Compared with neural networks, DTs are less computationally intensive, relevant attributes are selected automatically, and classifier output is a set of if–then conditional tests of the features [12]. Usually, a common set of features is used jointly in a single decision step. An alternative approach is to use a multi-stage or sequential hierarchical decision scheme. Classification trees offer an effective implementation of
such hierarchical classifiers. A DT classifier has a simple form that can be compactly stored and that efficiently classifies new data. DT classifiers can perform automatic feature selection and complexity reduction [12].

5.4. Splitting Index

A splitting index is used to evaluate the goodness of the alternative splits for an attribute. Several splitting indices have been used. The criterion for division usually relies on Gini’s index, as seen in CART [12]. The Gini index is defined in what follows.

Suppose $S$ is a set of $s$ samples. These samples have $m$ different classes ($C_i$, $i = 1, \ldots, m$). According to the differences of classes, $S$ can be divided into an $m$ subset ($S_i$, $i = 1, \ldots, m$). Suppose $S_i$ is the sample set, which belongs to class $C_i$, and $s_i$ is the sample number of set $S_i$; then the Gini index of set $S$ is

$$gini(S) = 1 - \sum_{i=1}^{m} P_i^2,$$

(12)

where $P_i$ is the probability that any sample belongs to $C_i$, estimated with $s_i/s$ [30].

5.5. Feature Selection

Feature selection, i.e., selecting a subset of the features available for describing the data before applying the learning algorithm, is a common technique for simplifying or speeding up computations [31]. The layer relation of a tree helps to make an explanation for data, to eliminate redundant data and noise, and to select classification features. In 1999, Borak and Strahler used a DT to subtract the classification feature from a large amount of data [32].

5.6. Detection of Winding Displacement Value Using DT

The proposed algorithm is summarized in the following.

Step 1: The $S$ parameters, measured in different axial positions of the model, are stored in a database. This data is used as a reference feature vector.

Step 2: The $S$ parameters are measured in new unknown axial positions.

Step 3: The axial position of the winding is estimated using the DT classification method.

To classify the $S$ parameters as phasor quantities, it is possible to use the magnitude, the phase, or both in the feature vector. The best results have been obtained in the case of the magnitude and the phase selection in the feature vector.

6. Results

6.1. $k$-NN method

The scattering parameter magnitudes have been selected as pattern vectors $p_i$ ($1 \leq i \leq n$). The cylinder has been shifted axial in 94 different positions. Considering $k$-NN method, the classes should be defined. The cylinder has been shifted axial in 0.1-mm steps. Every $N$ step can be considered as a unit displacement class. For example, for $N = 10$, the number of positions in each class is 10 and the total number of classes is also 10; for
Table 1
Accuracy ($A$) for different numbers of neighbors ($k$) and steps in each class ($N$)

<table>
<thead>
<tr>
<th>$k$</th>
<th>$N$</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>10</td>
<td>80.851</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>72.34</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>65.957</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>60.638</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>48.936</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>31.915</td>
</tr>
</tbody>
</table>

$N = 5$, the number of positions in each class is 5 and the total number of classes is 19. The results for different numbers of $k$ and $N$ are summarized in Table 1. In this table, the accuracy ($A$) is defined as follows:

$$A = \frac{\text{Number of correct classified query patterns}}{\text{Total number of query patterns}} \times 100. \quad (13)$$

It is obvious that the accuracy can be improved by decreasing $k$ and increasing $N$. Errors occur mostly because of the marginal positions that may be detected in the adjacent class.

In the second algorithm, the patterns $p_i$ ($1 \leq i \leq n$) are the scattering parameters, and every step is defined as a unique class $l(p_i)$. The $k$-NN regression algorithm has been used to estimate the unknown position of the winding based on the measured scattering parameters. Table 2 lists the $A$ parameter for different numbers of $k$. In this table, $E$ is the acceptable difference between the estimated position and the actual position. $E$ represents the ability of the method in the detection of the displacement.

6.2. DT Method

The scattering parameters magnitude and phase have been selected as feature vectors. A part of the DT obtained from 75 training samples is shown in Figure 11.

Table 2
Accuracy ($A$) for different numbers of neighbors ($k$) and acceptable differences between estimated and real position ($E$)

<table>
<thead>
<tr>
<th>$E$ (mm)</th>
<th>$k = 1$</th>
<th>$k = 3$</th>
<th>$k = 5$</th>
<th>$k = 7$</th>
<th>$k = 9$</th>
<th>$k = 11$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>61.70213</td>
<td>53.19149</td>
<td>44.68085</td>
<td>42.55319</td>
<td>38.29787</td>
<td>36.17021</td>
</tr>
<tr>
<td>0.2</td>
<td>87.23404</td>
<td>76.59574</td>
<td>69.14894</td>
<td>59.57447</td>
<td>52.76596</td>
<td>63.82979</td>
</tr>
<tr>
<td>0.3</td>
<td>93.16702</td>
<td>88.29787</td>
<td>82.97872</td>
<td>78.7234</td>
<td>75.53191</td>
<td>74.46809</td>
</tr>
<tr>
<td>0.4</td>
<td>94.68085</td>
<td>93.16702</td>
<td>90.42553</td>
<td>91.48936</td>
<td>87.23404</td>
<td>81.91489</td>
</tr>
<tr>
<td>0.5</td>
<td>95.74468</td>
<td>95.74468</td>
<td>93.61702</td>
<td>93.61702</td>
<td>91.48936</td>
<td>82.97872</td>
</tr>
<tr>
<td>0.6</td>
<td>96.80851</td>
<td>95.74468</td>
<td>97.87234</td>
<td>97.87234</td>
<td>92.55319</td>
<td>87.23404</td>
</tr>
<tr>
<td>0.7</td>
<td>97.87234</td>
<td>98.93617</td>
<td>97.87234</td>
<td>98.93617</td>
<td>95.74468</td>
<td>89.3617</td>
</tr>
<tr>
<td>0.8</td>
<td>98.93617</td>
<td>98.93617</td>
<td>100</td>
<td>98.93617</td>
<td>95.74468</td>
<td>91.48936</td>
</tr>
<tr>
<td>0.9</td>
<td>98.93617</td>
<td>98.93617</td>
<td>100</td>
<td>98.93617</td>
<td>97.87234</td>
<td>91.48936</td>
</tr>
<tr>
<td>1</td>
<td>98.93617</td>
<td>98.93617</td>
<td>100</td>
<td>98.93617</td>
<td>97.87234</td>
<td>93.61702</td>
</tr>
</tbody>
</table>
Figure 11. Part of DT, \((M(f)\) and \(P(f)\) show magnitude and phase of scattering parameter at frequency \(f\), respectively.

The most important feature is the phase at \(f = 8.925 \text{GHz} \) (\(P(8.925)\), the first node of Figure 11). In the second level, the magnitudes at 10.05 GHz and 7 GHz are important (\(M(7)\) and \(M(10.05)\) in Figure 11). This DT has been tested by 19 test samples. In Table 3, the comparison of the DT test results with \(k\)-NN regression has been presented.

### Table 3
Comparison of DT test results with \(k\)-NN regression

\((d\) and \(d_e\) are displacement and estimated displacement, respectively)\\

<table>
<thead>
<tr>
<th>(k)</th>
<th>(k = 3)</th>
<th>(k = 5)</th>
<th>(k = 7)</th>
<th>(k = 9)</th>
<th>(k = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d) (mm)</td>
<td>(d_e) (mm)</td>
<td>Error (%)</td>
<td>(d) (mm)</td>
<td>(d_e) (mm)</td>
<td>Error (%)</td>
</tr>
<tr>
<td>-4.3</td>
<td>-3.8</td>
<td>0.625</td>
<td>-4.5</td>
<td>0.25</td>
<td>-4.4</td>
</tr>
<tr>
<td>-3.9</td>
<td>-3.5</td>
<td>0.5</td>
<td>-4.0</td>
<td>0.125</td>
<td>-4.0</td>
</tr>
<tr>
<td>-3.3</td>
<td>-2.5</td>
<td>1</td>
<td>-3.6</td>
<td>0.375</td>
<td>-3.6</td>
</tr>
<tr>
<td>-2.9</td>
<td>-3</td>
<td>0.125</td>
<td>-2.9</td>
<td>0</td>
<td>-2.8</td>
</tr>
<tr>
<td>-2.3</td>
<td>-2.8</td>
<td>0.625</td>
<td>-2.5</td>
<td>0.25</td>
<td>-2.4</td>
</tr>
<tr>
<td>-1.9</td>
<td>-1.5</td>
<td>0.5</td>
<td>-1.9</td>
<td>0</td>
<td>-1.8</td>
</tr>
<tr>
<td>-1.3</td>
<td>-1.2</td>
<td>0.125</td>
<td>-1.4</td>
<td>0.125</td>
<td>-1.5</td>
</tr>
<tr>
<td>-0.9</td>
<td>-1</td>
<td>0.125</td>
<td>-1.1</td>
<td>0.25</td>
<td>-0.7</td>
</tr>
<tr>
<td>-0.3</td>
<td>0</td>
<td>0.375</td>
<td>-0.4</td>
<td>0.125</td>
<td>-0.3</td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
<td>0.125</td>
<td>-0.1</td>
<td>0.25</td>
<td>-0.1</td>
</tr>
<tr>
<td>0.7</td>
<td>0.2</td>
<td>0.625</td>
<td>1.0</td>
<td>0.375</td>
<td>0.9</td>
</tr>
<tr>
<td>1.1</td>
<td>1.2</td>
<td>0.125</td>
<td>1.0</td>
<td>0.125</td>
<td>1.0</td>
</tr>
<tr>
<td>1.7</td>
<td>1.5</td>
<td>0.25</td>
<td>1.5</td>
<td>0.25</td>
<td>1.3</td>
</tr>
<tr>
<td>2.1</td>
<td>2.2</td>
<td>0.125</td>
<td>2.1</td>
<td>0</td>
<td>1.9</td>
</tr>
<tr>
<td>2.7</td>
<td>2.5</td>
<td>0.25</td>
<td>2.8</td>
<td>0.125</td>
<td>2.7</td>
</tr>
<tr>
<td>3.1</td>
<td>3</td>
<td>0.125</td>
<td>3.2</td>
<td>0.125</td>
<td>3.2</td>
</tr>
<tr>
<td>3.7</td>
<td>3.8</td>
<td>0.125</td>
<td>3.7</td>
<td>0</td>
<td>3.7</td>
</tr>
<tr>
<td>4.1</td>
<td>3.9</td>
<td>0.25</td>
<td>4.2</td>
<td>0.125</td>
<td>4.1</td>
</tr>
<tr>
<td>4.7</td>
<td>4.6</td>
<td>0.125</td>
<td>4.4</td>
<td>0.375</td>
<td>4.3</td>
</tr>
</tbody>
</table>
Table 4
Comparison of average error percent of the two methods (DT with the $k$-NN regression)

<table>
<thead>
<tr>
<th>Method</th>
<th>k-NN</th>
<th>$k = 3$</th>
<th>$k = 5$</th>
<th>$k = 7$</th>
<th>$k = 9$</th>
<th>$k = 11$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average error (%)</td>
<td>0.322</td>
<td>0.171</td>
<td>0.184</td>
<td>0.230</td>
<td>0.309</td>
<td>0.322</td>
</tr>
</tbody>
</table>

The percent of error is calculated by the following equation:

$$\text{Error(\%)} = \frac{\text{Absolute Error}}{\text{Winding Height}} \times 100. \quad (14)$$

It is obvious that the DT can discriminate between the symmetrical upward and downward displacements. In [22], to determine the axial displacement extent using FRA, the winding has been moved in three steps. Each step was equal to 4% of the total winding height. In this article, the winding has been moved in 94 steps. Each step is equal to 0.1% of the total winding height.

The average error percents have been compared in Table 4. The results show that the $k$-NN with $k = 3$ has the minimum error and is superior to the DT method for the detection of the axial displacement value.

As it can be seen in Tables 3 and 4, the DT method in the worst case has a maximum error of 1% (of the winding height), and $k$-NN in the worst case has a maximum error of 1.5% for $k = 11$; thus, both methods are more accurate than the FRA method, which has an error of 4% [33]. The minimum detectable axial displacement in the proposed method is 0.125% of the transformer winding height but is 0.5% of the transformer winding height the FRA method [34]. The ratio of the excitation frequency to the winding dimensions in the proposed method is much higher than the same ratio in the FRA method. This higher ratio results in higher accuracy, which can be seen in the both proposed method results.

7. Conclusion

In this article, the on-line monitoring of the transformer winding has been investigated using scattering parameters. To show the capability of the method, the scattering parameters of a simplified model of transformer have been measured by a network analyzer in different axial positions of the test object.

The proposed method has the following merits:

- in this method, there is no electrical connection to the windings; therefore, neither the HV nor LV windings of the transformer should be disconnected from the network;
- this method can be used for off-line and on-line applications;
- transformers can be monitored in the specific intervals or continuously;
- the amount of the axial displacement can be determined;
- the sensitivity to displacements is higher than the FRA method because the scattering parameters are measured at higher frequencies;
- the detection accuracy of the axial winding displacement is dependent to the transmitter frequency, i.e., decreasing the wavelength enhances the sensitivity to the displacement;
• this method cannot be affected by the power factor of the load and loading conditions; and
• antenna movement is not necessary.

Two indices have been defined by using the magnitude and phase of the scattering parameters to detect the axial displacement of the windings. The results show that these indices can be used to detect the winding displacement and present information about the approximate amount of the displacement.

DT and $k$-NN regression have been used for the exact estimation of the displacement. Both of them can discriminate between upward and downward displacements. The results show that the $k$-NN with $k = 3$ has the minimum error and is superior to the DT method for the detection of the axial displacement value. Both methods are more accurate than the FRA method.

References


