TRANSFER FUNCTION ESTIMATION AND AI APPLICATION FOR TRANSFORMER FRA INTERPRETATION

A thesis submitted to

The University of Manchester

for the degree of

Doctor of Philosophy

in the Faculty of Science and Engineering

2019

Xiaozhou Mao

Department of Electrical and Electronic Engineering

Blank page

Contents

List of Tables and Figures

Abstract

Winding fault takes up a considerable proportion among all the fault types of transformer. Frequency response measurement of transformer has been developed as an effective tool to detect the mechanical integrity of windings. In Frequency Response Analysis (FRA) technique, the diagnosis measurement should be compared with the reference measurement, and winding displacement/deformation may be suggested by the occurrence of difference. However, no IEEE or IEC standard has been published regarding the interpretation of frequency responses. Also, a large amount of frequency responses have been accumulated by the utilities over many years. Utilities may or may not know transformer design information, such as the winding construction types. Different winding construction types are susceptible to different modes of mechanical deformations, and the same asset management method can be applied to transformers with same winding construction types.

Two methods are proposed to derive the mathematical expression for the FRA trace, in the format of a transfer function. Using this mathematical expression data can be generated in the same format, no matter how or by which FRA measurement device the original FRA trace is produced; and the same data format is of necessity for applying numerical indices and AI techniques to interpret FRA. Feature Extraction Method divides the frequency range into several regions, and complex poles and zeros are extracted to form a feature transfer function, the difference between which and the measured data is then corrected by real poles and zeros and the constant. It is validated by 48 frequency responses of eight 400/275/13 kV autotransformers. Extreme Points Identification Method detects extreme points and iterates to optimise the transfer function's parameters. The Feature Extraction Method is good at describing the subtle feature of frequency responses, whilst the Extreme Points Identification Method ensures the simplistic expression to be identified, which is physically achievable by filter design principle. Both methods have been successfully applied for the diagnosis of faulty transformer winding.

Two methods are proposed to identify the winding construction types using frequency responses. The supervised machine learning method, Support Vector Machines (SVM), is utilised to build an identification model, using FRA traces of transformers with known winding type. After testing and validating, the SVM model is then applied to FRA traces with unknown winding type information. A set of data from the UK's National Grid FRA database, was used to demonstrate and verify the SVM model. The proposed method can successfully identify winding types including multiple layer, plain disc, interleaved disc and single helical windings. The unsupervised machine learning method, Hierarchical Clustering, is utilised to cluster frequency responses according to the similarity and dissimilarity. Once the frequency responses are in the same cluster, it is in default to think the windings should share the same winding construction type. It was applied to UK's National Grid FRA database so the frequency responses of transformers with unknown winding type can be identified by being clustered into a group together with a frequency response with known winding type.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

Copyright Statement

(i) Copyright in text of this thesis rests with the author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the author and lodged in the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without permission (in writing) of the author.

(ii) The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

(iii) Further information on the conditions under which disclosures and exploitation may take place is available from the Head of Department of Electrical and Electronic Engineering.

Acknowledgment

I would like to express my sincere gratitude to my supervisor Prof Zhongdong Wang for her continuous encouragement and support to my PhD study. I respect her for her intellectual curiosity, meticulous academic attitude, and cares about her students. She is my role of model for life.

I would like to thank Prof Paul Jarman, Mr Andrew Fieldsend-Roxborough and Dr Gordon Wilson from National Grid for the financial and technical support on the FRA project.

I would like to thank Prof Zanji Wang from Tsinghua University, during and after his short academic visit to Manchester, for providing the technical help in the verification of my proposed method for transfer function estimation. His kind support and valuable advices towards my research are highly appreciated.

I would like to thank Dr Hongzhi Ding from Doble Engineering for providing the helpful information regarding the investigated faulty transformers.

Also, I would like to thank my friends and colleagues for all the peer support they provided during my PhD study.

Last but not least, I would like to thank my family for their moral and financial support, encouragement and constant love for me, especially my father and mother.

Blank page

Chapter 1 Introduction

1.1 Introduction

Secure and stable electricity supply depends substantially on the reliability of equipment used in a power system. Among all the electrical equipment, transformers are one of the most important and expensive apparatus. The lifetime of a transformer is expected to be around 40 years [1], and longer operation time of transformers can be acquired by proper asset management. Indeed, the in-service time of many transformers in the UK has already exceeded the estimated lifetime. Condition monitoring and diagnostics of transformers are of importance to avoid failures or outages which usually cause significant economic losses.

In fact, besides ageing, various types of transformer faults have impacts on the lifetime of a transformer. Transformer faults can be generally classified into thermal, mechanical and electrical faults according to the fault causes, or faults in winding, tap changer, insulation systems, etc. according to the fault locations. Mechanical fault of transformer windings is referred here as the winding movement, i.e. winding displacement or deformation. According to a survey by CIGRE, among all the failures for large power transformers with on-load tap changers, the percentage of winding failure was 19%, while for power transformers without on-load tap changers, the percentage of winding failure was 26% [2]. Figure 1-1 illustrates the percentage of transformer fault in each category [3, 4]. It can be seen that winding mechanical failure takes up a considerable percentage, 30.56%, among all the transformer fault types for transformers aged 15 to 25 years [4].

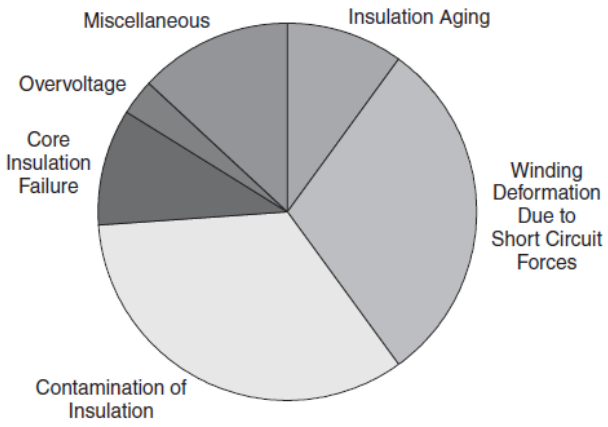


Figure 1-1 Failure Percentages of Transformers Aged 15-25 Years [3]

Frequency Response Analysis (FRA) has been developed as an effective and sensitive technique to verify the winding integrity of transformers. Different frequency regions of

the FRA traces are dominated by different factors, such as transformer core, interaction between windings, winding properties and measurement equipment. Frequency response should be measured and recorded at the transformer factory, which can be used as a reference/baseline for future diagnosis. The FRA technique compares the reference and diagnostic frequency responses, and the deviation between them can be interpreted to firstly identify if there is a mechanical fault or not, and secondly if possible, to ascertain the type, severity, and location of winding fault. It is noticeable that different winding construction types may be susceptible to different winding faults due to their difference in physical structure. However, no IEEE or IEC standard has been published regarding the interpretation of frequency responses.

This PhD study has two main objectives. One is to develop a mathematic representation of FRA curves in the format of a transfer function. The FRA traces then can be expressed mathematically in the same format, no matter how or by which FRA measurement device the FRA curve is produced. This will ultimately help apply numerical indices and AI techniques for FRA interpretation. The other is to identify the winding construction types using frequency response, for more accurate interpretation of frequency responses and more efficient asset management.

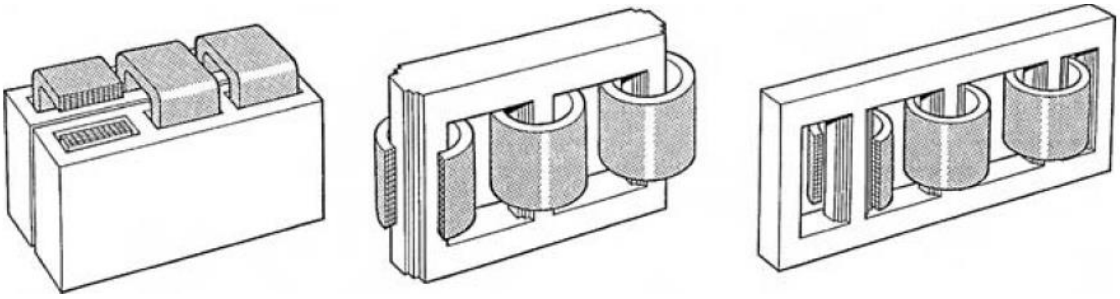
1.1.1 Basic Understanding of Transformer Construction

A transformer is an electrical apparatus which transfers electrical energy between two or more circuits through inductive coupling conductors, i.e., the coils (also called windings), without electrical connection between them. The changing current in the primary winding of the transformer generates a changing magnetic field, which in turn induces a changing electromotive force (emf) in the secondary winding. The voltage ratio of primary and secondary windings is decided by the ratio of turns of conductors in primary and secondary winding.

Transformers can be categorised according to power capacity, duty, frequency range, voltage class, cooling type, etc. The basic components of power transformers generally include core, windings, insulation systems, tap changer, bushing, oil conservator, breather, Buchholz relay and other accessories.

A transformer core provides the main path for magnetic flux to flow when a transformer is in operation. There are two types of transformers in terms of core construction: core-form

and shell-form transformers. Their main difference lies in the way how windings are wrapped. As shown in Figure 1-2, windings surround the core in a core-form transformer, whilst in a shell-form transformer the core surrounds the windings. The core-form transformers are used widely in Europe and the UK. In Figure 1-2 (b), a three-phase three-limb core-form transformer is illustrated. In this arrangement, the cross section of yokes at the top and bottom is designed the same as the cross section of the limbs, thus a return path of the flux is not needed. For large transformers, because of the limitation on its height during transportation, the cross section of the yokes is therefore reduced to about half of the cross section of the limbs. Consequently, a three-phase five-limb core-form transformer should be used to provide flux return path, as shown in Figure 1-2 (c).



(a) Shell-Form Transformer (b) Three-Limb Core-Form Transformer (c) Five-Limb Core-Form Transformer
Figure 1-2 Core and Shell Type Transformers [1]

Insulation materials used in an oil immersed power transformer include insulating oil, cellulose paper and pressboard, and wood-based laminates. The transformer's durability mainly depends on the insulation quality and the design of insulation system.

Tap changers are used to regulate the secondary voltage which is influenced and varied by load. The tap changers are normally connected to the high voltage windings because it is easier to access than the low voltage windings. Another advantage is that the current is less and this leads to better durability of the tap changers. Depending on whether the transformer is in operation when adjusting turn ratios, tap changers can be classified into two categories: on load tap changer and off circuit tap changer.

A transformer bushing is an insulated device. It provides a safe passageway for an electrical conductor, which carries current, to go through the grounded transformer tank. The conductor can be built inside the bushing, or the bushing can be constructed allowing the conductor to be drawn through it. Porcelain, oil paper and/or resin are commonly used insulation types for bushings. An energised conductor, which is close to the material at earth potential, produces an electric field. The bushing can control the shape and strength of the electric field, and it reduces the electrical stresses on the insulating materials.

Windings are one of the most important components in a transformer; and the choice of winding construction type is greatly influenced by the core type, the transformer manufacturer's historic experience, the transformer's voltage and power ratings. In general, for the same power rating, a higher voltage winding prefers to use the winding construction type which gives a larger winding series capacitance due to the requirement to withstand a stringent Basic Insulation Level (BIL) level.

In shell-form transformers, sandwich type windings are used, and the cross-sectional top view of Figure 1-2 (a) is shown in Figure 1-3. The high voltage winding sections are located between low voltage winding sections, hence the name of sandwich type. The sandwich type windings have high short-circuit durability and high mechanical strength.

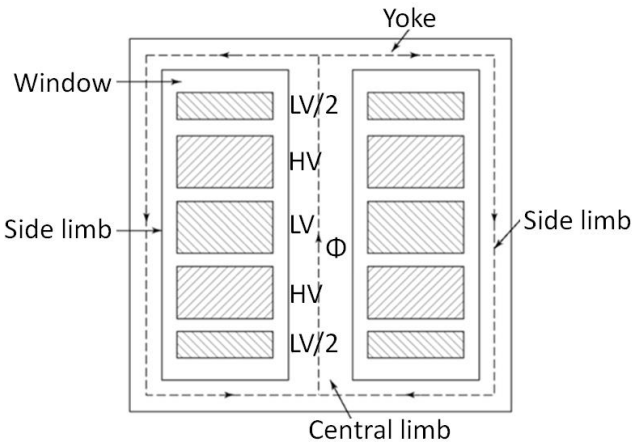


Figure 1-3 Sandwich Winding of Shell-Form Transformer [5]

In core-form transformers, concentric coils are utilised and windings can be generally categorised as single helical type, multilayer type, and disc type [6]. The Single Helical winding owns the smallest winding series capacitance and can be used when the voltage is low. Plain Disc windings and Intershielded Disc windings have been applied for a wide range of voltage levels (normally for a voltage level <= 132 kV but sometimes 275 kV). Interleaved Disc windings are suitable for higher BIL level (normally for a voltage level >= 275 kV). For historical reasons some manufacturers use Multiple Layer windings' design rather than disc-type windings, hence Multiple Layer windings are used for all high and low voltage windings in a transformer. It is also applicable for voltages from 33 kV right to 400 kV, or even up to 800 kV [7]. Figure 1-4, Figure 1-5, Figure 1-6 and Figure 1-7 are the schematic figures of the windings of core-form transformers, illustrating how a winding is geometrically arranged. The turn numbers given in those figures are electrical turn number representing electrical continuity, e.g. turn 1 is connected with turn 2 although they can be geometrically/physically apart.

Figure 1-4 shows a 3D cross-sectional view of Single Helical winding structure. It rolls its first turn from the top and ends its last turn at the bottom using a former. This type of winding is only suitable for low voltage, e.g. a tertiary winding in a transmission transformer. It can also be used for a generator step-up transformer. This is due to the fact that the generator step-up transformer has low voltage and thus high current in the primary winding. Considering the limitation on current density of the conductor material, the cross section of the primary winding's conductor should be large enough for a high current, making the helical winding suitable to be used for low voltage primary winding [8].

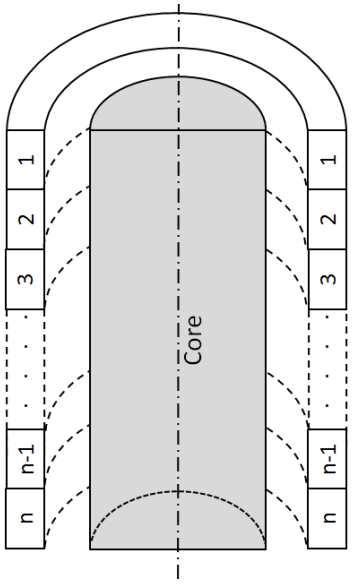
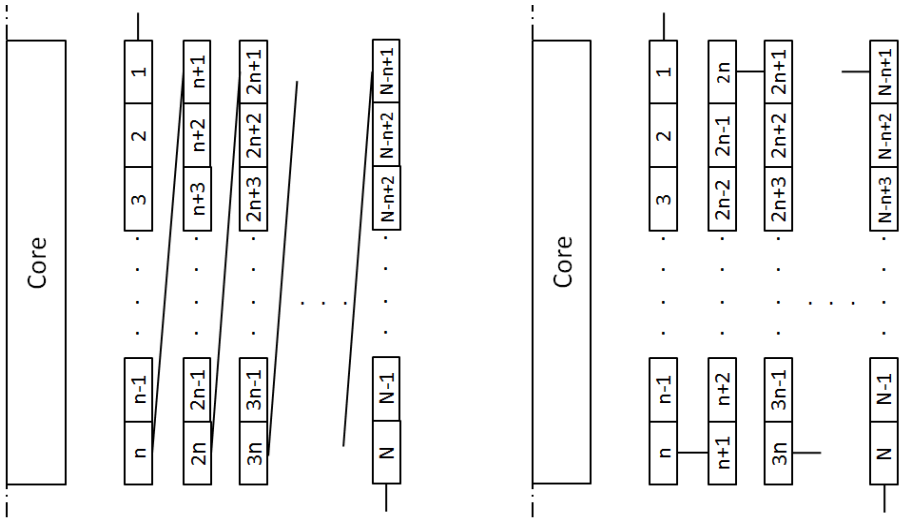


Figure 1-4 Single Helical Type Winding

For convenience, when illustrating the structure of the following winding construction types, the half 2D cross section of an axial symmetric arrangement, centred to the centre line of the core, will be used instead of the 3D one.

The difference between the two interlayer connections for Multiple Layer winding is illustrated in Figure 1-5. There are oil ducts and solid insulations between the layers. It is not practical to link the turn at the bottom part of a layer to the turn at the top part of next layer due to the limitation of space, as shown in Figure 1-5 (a). The construction of linkage in Figure 1-5 (b) is the more preferable design. For transformer high voltage windings, when Multiple Layer type winding is utilised, the voltage drop is large between the adjacent layers and this needs a higher insulation level. The Multiple Layer type winding tends to have less mechanical strength, which needs some special design consideration.



(a) Inter Layer Connection 1 (b) Inter Layer Connection 2

Figure 1-5 Multiple Layer Type Winding

Plain Disc type, Interleaved Disc type and Intershielded Disc type windings are widely used disc-type windings for core-form transformers in the UK. The disc type winding is convolved in such a way that the horizontal discs are wound one disc after another disc, as shown in Figure 1-6, while for Multiple Layer type windings the vertical layers are built in order. Compared with Multiple Layer winding, the Plain Disc winding design benefits the high voltage transformer applications due to less insulation requirements.

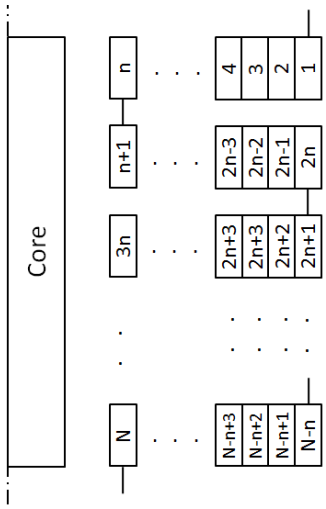


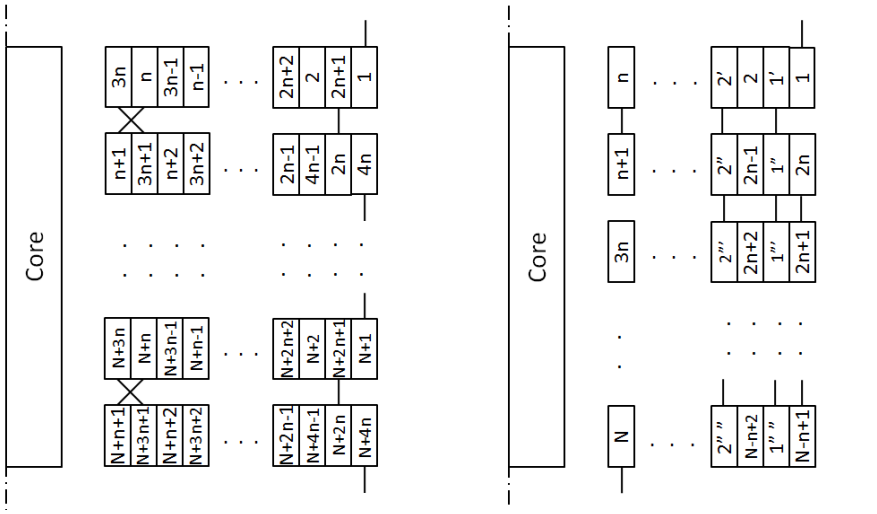
Figure 1-6 Plain Disc Type Winding

Higher level of insulation ability is needed when the voltage rating of transformer goes up, and this can be simply represented as BIL. The BIL is usually tested by lightning impulse voltage tests. As we know that lightning impulse is composed of high frequency components, the transformer winding cannot be regarded as an inductive component any longer and the winding capacitance, series and ground, will come to play. The initial impulse voltage distribution coefficient α is defined as:

$$\alpha = \sqrt{\frac{C_g}{C_s}} \qquad \text{Equation 1-1}$$

where C_g is the winding capacitance to ground, and C_s is the winding series capacitance. In this sense α becomes an important parameter for us to estimate the degree of non-linear distribution of the impulse voltage along the winding [9].

On the other hand, the design practice for the volts per turn under the AC voltage tends to remain unchanged, whilst the turn number should be increased with the increase of the voltage rating. With a Plain Disc winding, the increase of turn number will simply reduce the absolute value of winding series capacitance and increase the winding capacitance to ground, which results in more non-linear distribution of lightning impulse voltage along the winding. This is the main reason why Plain Disc windings at higher voltage failed the BIL test in factory during 1950s. Interleaved Disc type was thus introduced by English Electric, UK, in the late 1950s [10]. Its series capacitance is higher than that of the Plain Disc winding type. Interleaved Disc winding is labour intensive as there are many joints and cross-overs needed to be hand-made. To save the labour cost, another method to increase series capacitance is used as the invention of the Intershielded Disc type winding in Japan by Morita, Ohta, and Kurita in 1970s [11-13]. The Intershielded Disc winding has an lower level of series capacitance than Interleaved Disc winding [9]. Shielded turns added between winding turns are able to make the voltage distributed more uniformly along a winding, and therefore a lightning surge impact can be tolerated. The arrangements of the two disc windings are shown in Figure 1-7.



(a) Interleaved Disc Type Winding (b) Intershielded Disc Type Winding

Figure 1-7 Disc Type Windings

The variations in series capacitance, ground capacitance and inductance of different winding construction types will lead to the difference of their RRA characteristics. This is

why it is important to understand the transformer design and manufacture development history. The voltage and power ratings of a transformer will give us clues on the winding equivalent parameters, more importantly knowing the manufacturer related winding technology will also help us to estimate the range of the C_s, C_g and L values.

1.1.2 Transformer Condition Assessment Methods

Periodical condition assessments of in-service transformers, either by online or off-line tests, are used for asset management. For different types of transformer faults, different methods and tools have different sensitivity, and unfortunately there is no "one fit for all" tool at the moment for fault detection and it is unlikely in the near future. Hence the power utilities use a suite of methods/tools to conduct health assessment. Dissolved Gas Analysis (DGA), Partial Discharge Analysis (PDA), and radio interference analysis are commonly used online testing methods. Frequency Response Analysis (FRA), Capacitance and Power/Dissipation Factor (C&PF/C&DF) test, Polarisation and Depolarisation Current (PDC), and Frequency Domain Spectroscopy (FDS) are some other diagnostic techniques intended for offline tests.

Dissolved Gas Analysis (DGA) is the most frequently used method to detect incipient electrical or thermal faults, and it has a tremendous advantage of either conducting laboratory based tests on sampled oil or using on-line DGA monitors [14]. Research suggested that the percentage of the gases extracted from the oil can help identify types of fault, according to Roger's ratio [15]. Recently the industry practice tends to emphasise simplicity, so the key gas method prevails. As an example, if a large quantity of hydrogen (H_2) gas is detected, a partial discharge type of fault may exist in the transformer. In the same logic, ethylene (C_2H_4) may represent a thermal fault while carbon monoxide and dioxide (CO and CO_2) may represent a cellulose related fault.

The Partial Discharge Analysis (PDA) technique can be used to recognise the PD phenomenon [16]. PD as defined by IEC 60270, is a localized discharge in a solid or liquid dielectric insulation system under high-voltage field stress [17]. If unusual values such as gas-in-oil values are detected, PD measurement methods can be used for further diagnosis. The multi-terminal PD measurement, Ultra High Frequency (UHF) PD measurement method, and the acoustic PD measurement are based on the detection of the electric currents, electromagnetic wave, acoustic radiation of the PD phenomenon, respectively. Electrical PD measurement is the most sensitive method to directly detect the discharge

signal [18]. The localisation of a PD source can be identified by the UHF method according to the difference of time of flight measured by different sensors. Research has also been conducted recently to locate PD source through UHF signal amplitude strength attenuation [19]. However, there is a limitation on the amount of the UHF sensors to be used, which is normally 3, through the oil filling valves. Thus the acoustic localisation method can be used; piezo-electric acoustic sensors are installed on the transformer tank [20].

As for offline tests, capacitance and power factor of insulation can be measured between windings or between winding and ground at power frequency, for the purpose of identifying the ageing and degradation of insulating material [21].

In Frequency Domain Spectroscopy (FDS) test, the complex capacitance/permittivity and dielectric power/dissipation factor are measured at different frequencies, normally from 1 mHz to 1 kHz. Oil/paper conductivity, temperature, moisture in the insulation, insulation ageing, acids, etc. are the important influencing factors of FDS. The FDS properties can be used to assess the ageing and/or dryness states of a transformer oil-paper insulation system. Typically, in the frequency region 0.01 Hz to 10 Hz, the dielectric response is primarily affected by the oil properties, while the frequency region below 0.01 Hz and above 10Hz is mainly impacted by the properties of solid insulation i.e. paper and pressboard. The FDS measurement is time-costing at low frequencies [22].

Polarisation and Depolarisation Current (PDC) measurement is carried out in the time domain to determine the conductivity, moisture content, and ageing status of transformer insulation system. A step voltage is applied to the insulation sample, and afterwards, the sample is short circuited. During this process, the polarization (charging) and depolarization (discharging) currents are recorded. Normally, larger currents indicate higher conductivity, higher moisture content and worse ageing state of the transformer insulation materials [23, 24].

Frequency Response Analysis (FRA) is an effective way to assess transformer mechanical condition, particularly on the condition of transformer windings in terms of winding displacement and deformation. It is capable to discover winding mechanical faults without intruding inside a transformer. For frequency response measurements, the transformer winding is usually stimulated by a variable frequency sinusoidal signal at the injection terminal point and the response signal is received and measured at another winding terminal point. The so called frequency response of transformer winding is described in the

frequency domain by the magnitude ratio and phase difference between the injected and received signals. The frequency range is normally from a few Hz to a few MHz.

The frequency response measurement is normally carried out by manufacturers before dispatching the transformer to the site first, as the reference for future diagnostic FRA measurements, which are conducted normally after the transportation and installation of a transformer, or after a system short circuit fault current passes through the transformer, or just during a periodical maintenance outage. The FRA trace is influenced by the collective effect of the electrical parameters of the winding equivalent circuit, i.e. the capacitance and the inductance, which are determined by the physical dimensions and structure of windings. The alteration in the reference and diagnostic frequency response traces are mainly caused by the changes in the electrical parameters of the windings, and thus winding displacement/deformation may be indicated.Shifting/appearance/disappearance of resonance/antiresonance is regarded as the key feature for FRA to diagnose the winding displacement/deformation. In general, shifting of higher resonant frequency indicates a smaller geometrical change in a winding or part of the winding; and shifting of lower resonant frequencies (discard the influence of core magnetization) indicates a larger geometrical change of winding or the relative change between windings' positions. Not only the deformation, displacement or damage of the transformer winding, the FRA technique can also detect the core defects, the contact resistance, the residual magnetization, the floating shield, etc. However, the difficulty in interpreting the FRA is to correlate the changes in the FRA traces to the change of physical parameters of the winding's equivalent electrical circuit. No standard has been developed to guide the identification of winding mechanical faults, or to quantify severity of winding movement.

1.1.3 Short-Circuit Withstand Capability and Winding Movement

When the power transformers are in operation, the current-carrying windings are surrounded by magnetic flux. Consequently, electromagnetic forces are induced on the windings, which lead to mechanical stress. The forces can be transferred through the windings to other structural components of the transformer.

The short-circuit withstand capability of transformers refers to their ability to survive the external short circuit event, i.e., system disturbances. There are different kinds of short circuits in the power system, such as single line to earth short circuit, double phase short circuit, and three phase short circuit. When a short circuit fault happens, the current goes

through an unexpected pathway with very low impedance, thus a high level of short circuit current flows in the power network. The short circuit current can be dozens time larger than the current in normal operation. Since the electromagnetic force is proportional to the squared current, the winding need to stand hundreds time larger induced mechanical strength compared with that in the normal condition. Huge mechanical forces can be caused in milliseconds. Though the system protection procedure generally interferes promptly, damages can still be caused [25]. With the increase in voltage and power ratings, the short-circuit withstand capability of the transformers have become more demanding at the transformer design stage.

The induced large mechanical force may lead to the deformation, displacement or damage of transformer winding. There are two types of winding deformation. One is the radial deformation and the other is the axial displacement.

The leakage magnetic flux going axially through the core window will generate a radial force, which may lead to the radial deformation. The inner (LV) winding withstands the inward radial force, while the outer (HV) winding withstands the outward radial force, as shown in Figure 1-8 [25]. This can be explained by the left-hand rule, considering the fact that the current directions are different in the HV and LV windings, while they are located in the same magnetic field. The outward radial force on HV winding is resisted by the tension of the conductor along with the friction force between turns. As a result of the outward radial force, the tensile tangential mechanical stress is developed. The conductor material, copper, has comparatively high material yield strength, which means the HV winding does not normally fail. When the tensile tangential mechanical stress exceeds the conductor material yield strength, it can lead to the damaged insulation, local axial slackness, and cracker of the winding. For the LV winding, the compressive tangential mechanical stress is developed. There are two failure patterns of compressive-stressed winding. One is the forced buckling and the other the free buckling, depending on the stiffness of the winding conductor and the winding support structure. When the stiffness of the support structure is stronger, the forced buckling occurs as shown in Figure 1-9 (a) [26]. Otherwise, when the winding conductor is stronger, the free buckling occurs. The winding can move inwards or outwards for the free buckling, and the outwards direction is illustrated in Figure 1-9 (b) [26].

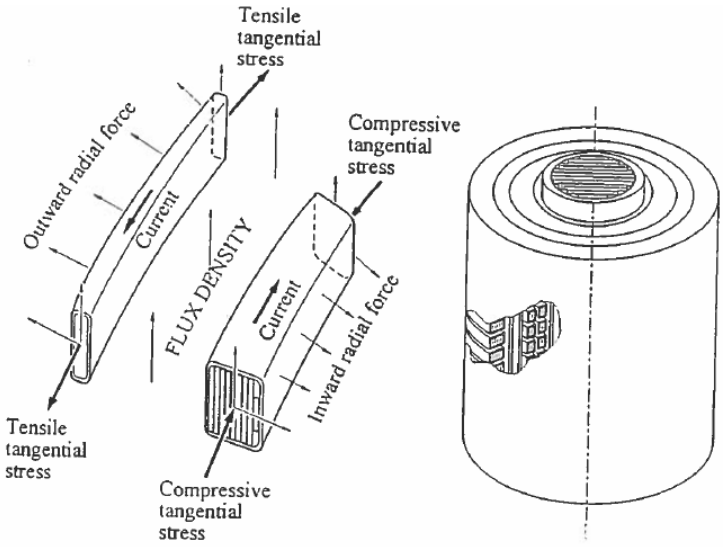
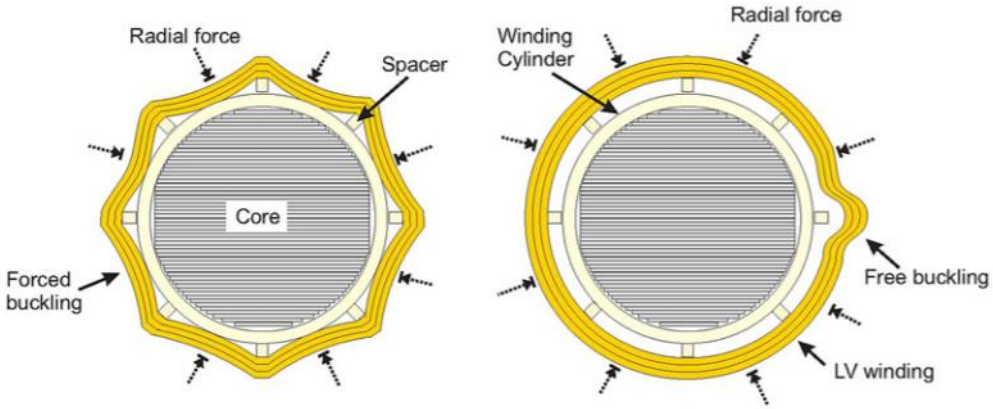


Figure 1-8 Radial Forces and Mechanical Stresses [25]

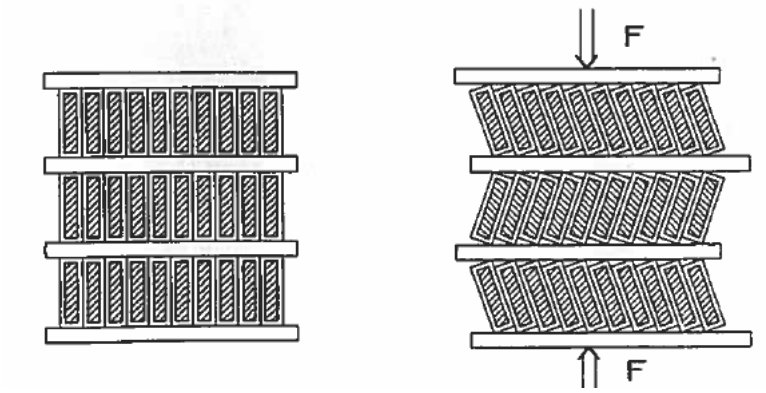


(a) Forced Buckling (b) Free Buckling

Figure 1-9 Radial Deformation [26]

The axial force consists of two parts. The first part is the axial compression force. It is generated on parallel conductors which carry currents in the same direction, in the magnetic flux caused by the conductors themselves. The second part is the axial expansion force. The conductors at the bottom and top of the winding withstand axial expansion forces in the opposite direction, which direct axially to the winding centre. The expansion force is caused by the radial leakage magnetic flux, which is induced by other conductors on the bottom and top of the winding. Normally there is a displacement between the magnetic centres of the HV and LV winding, which is a magnetic unbalance. A net axial force is caused consequently, which can enlarge the axial displacement between the HV and LV winding further. The axial forces may lead to different failure modes, such as the titling collapse, the telescoping, conductor bending, etc. The conductor titling happens when the adjacent set of conductors shift oppositely, and the cross sections of the normal and tilted conductors are shown in Figure 1-10 [25]. The telescoping only occurs in layer type windings, where some of the turns move axially over the adjacent turns. The

conductor may also bend between the support columns, as shown in Figure 1-11 [27]. Damage to the conductor insulation is one consequence of the axial force related failures.



(a) Normal Conductors (b) Tilted Conductors

Figure 1-10 Conductor Tilting [25]

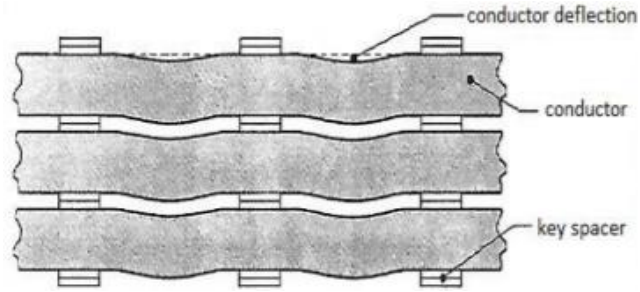


Figure 1-11 Conductor Bending [27]

The physical deformation of transformer windings leads to the variation in electrical parameters of the winding's equivalent circuit. Therefore, the diagnostic frequency response of the faulty windings varies from its reference frequency response.

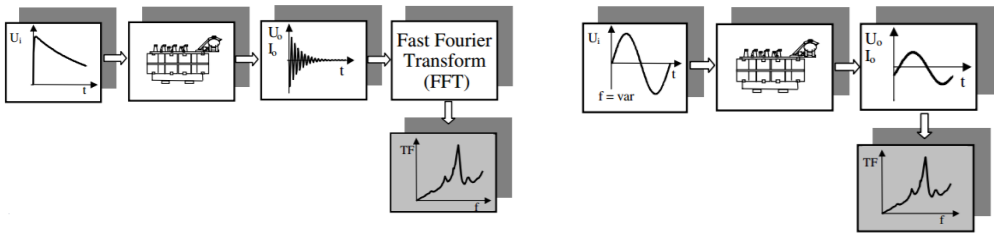
1.2 FRA Measurement and Data Storage

Several FRA measurement connection methods have been developed and applied, of which the sensitivity are different to different types of winding faults. To make the frequency response comparable, ideally the same connection method should be used for the reference and diagnostic FRA measurements. Besides, due to the choice of different measurement equipment, the measured frequency response may be stored in different format. Generally, a unified data format is desired for the convenience of further analysis.

1.2.1 FRA Measurement

The sweep frequency response analysis (SFRA) [28] and the impulse frequency response analysis (IFRA) [29, 30] are two methods of frequency response measurement, which are

illustrated in Figure 1-12. The main difference between them lies in the injected signals. IFRA injects an impulse voltage signal, whilst SFRA injects a sinusoidal signal with varying frequency on a wide frequency range. If IFRA is used, the time domain signals, both injected and responding signals, can be transformed into the frequency domain using Fast Fourier Transform (FFT).



(a) IFRA (b) SFRA

Figure 1-12 Methods of Determination of a Transformer's Transfer Function[31]

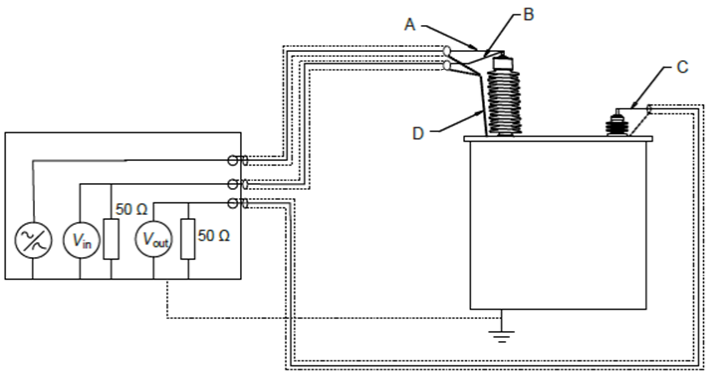
In 1966, W. Lech and L. Tyminski from Poland proposed the Low Voltage Impulse (LVI) method, which was found to be sensitive to detect winding movement [29]. In 1992, P. T. M. Vaessen and E. Hanique from KEMA laboratories, the Netherlands evolved this method into FRA method in their study on power transformers between 100 kVA and 450 MVA [30]. The input signal and the output signal were sampled in the time-domain, and the signals were converted into the frequency domain through FFT. Aliasing was avoided by using a low voltage impulse generator with adjustable output voltage from 0 to 300 V. A relatively high output voltage leads to a better signal to noise ratio. The resolution of both magnitude and phase responses in the frequency domain could be guaranteed. The reproducibility and simplicity in implementation of this method were proven. This technique was also called as 'transfer function' by T. Leibfried and K. Feser at Siemens, Germany in 1999 [32], and by E. Rahimpour at the University of Tehran, Iran [33, 34] and J. Christian and K. Feser at the University of Stuttgart/Siemens, Germany [35-37]. Actually, what they call 'transfer function' is the measurement data of frequency response, which is different from the 'mathematical expression' concept in the following chapters of the PhD thesis.

The SFRA technique was firstly introduced in the year of 1978 by E. P. Dick and C. C. Erven from Kinectrics Inc, Ontario Hydro, Canada [38]. They analysed the sweeping frequency response results of five different transformers and found that the alteration of inter-disc capacitance, inter-winding capacitance, winding inductance, bushing capacitance will lead to the changes in the frequency response traces, which suggested the feasibility of diagnosis of transformer windings by the frequency response method. A simple equivalent

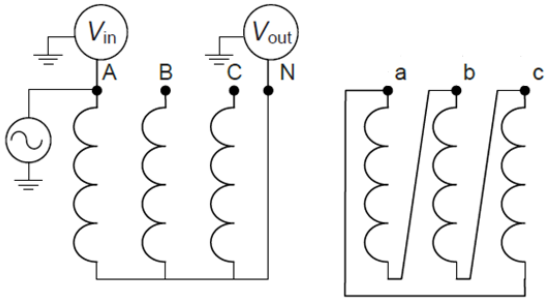
circuit model of the transformer winding was built, but the mutual inductance was not considered in this model. The authors suggested that though the correlation between the parameters of the electrical equivalent circuit and the FRA traces is proved, it still remains difficult to accurately interpret the frequency responses.

In the past decades, SFRA has been widely used for transformer winding deformation diagnosis. FRA technique has good performance in detecting winding mechanical faults, including axial bending displacement, radial deformation, short-circuit turns, bushing faults, and insulation degradation [34, 39-48]. The IEEE standard C57.149 is developed to guide FRA measurement of oil-immersed power transformers [49]. The measurement set up, the measurement process, the storage of data are specified in this standard. However, for the interpretation of the data, the IEEE standard only provides the typical FRA examples to illustrate the failure modes without any further detailed explanations. The IEC standard 60076-18 also focuses on the measurement technique and instrument while no interpretation guidance is provided, although some FRA measurement examples illustrating the influencing factors are included in the annex [50]. The CIGRE brochure WG A2.26 aims to provide the basic information of FRA to the non-experts to standardise test techniques, and to propose for further study on the improvement of FRA interpretation [51].

According to IEC 60076-18, the FRA measurement are usually arranged as shown in Figure 1-13 [50]. An input signal is injected to the winding system through the source lead, which is noted as line A in Figure 1-13 (a). The output signal is measured from a 50 Ω earthed resistors through the response lead C. The reference lead is noted as B while lead D represents the earth connection.



(a) Measurement leads



(b) End-to-End Measurement

Figure 1-13 FRA Measurement Setup [50]

In [52], J.A.S.B. Jayasinghe and Z.D. Wang, from the University of Manchester, UK, studied the influence of connection methods on frequency responses. They carried out sensitivity studies regarding simulated axial winding deformation, forced buckling on the low voltage winding and conductor axial bending on the high voltage winding, using a model of 30 MVA 132/11 kV transformer. Their study concluded that the 'inductive inter-winding measurement' method has the best performance both in the diagnosis of axial and radial displacement. As for the axial bending of conductor, the 'end-to-end measurement' method is the most sensitive method. The authors suggested that 'end-to-end measurement' and 'inductive inter-winding measurement' connection method should be both used, in order to identify most types of winding deformation.

In IEC standard 60076-18, the end-to-end measurement, end-to-end short circuit measurement, capacitive inter-winding measurement, inductive inter-winding measurement are introduced.

The end-to-end measurement is the most common connection method, as shown in Figure Figure 1-13 (b), and it is also called open-circuit measurement in the IEEE C57.149 standard. This measurement should be carried out on one end of a winding to the other end, with all the other terminals floating, and the phases and windings set as apart as possible. Input and output voltages are recorded on the two ends of the winding under test. [49].

The end-to-end short circuit measurement method is made on a single winding as well. However in this method, the other winding should be short-circuited, as shown in Figure 1-14, using a three-phase YNd1 transformer as example. By this configuration, the influence of the core is eliminated, and focus can be put on the bulk leakage inductance between the windings. Apart from the low freuqency region which is contolled by the core, its frequency response at the winding-under-test dominant region is similar to that of end-to-end open-circuit measurement.

The inductive inter-winding measurement, as shown in Figure 1-15 using a three-phase YNd1 transformer as example, is made on two windings from the same phase. The other ends of the two windings should be earthed. This connection method is also called the transfer voltage measurement. All the terminals of other windings should be left floating. In the low frequency region, the frequency response is controlled by the voltage ratio of the two windings. The higher frequency region is not the emphasis of this measurement configuration.

Same as the inductive inter-winding measurement, capacitive inter-winding measurement is performed on two windings from the same phase, while the ends of the two windings are left floating, as shown in Figure 1-16. With the two windings isolated, the frequency response is dominated by the network impedance between them. The magnitude of the frequency response has a rising trend when frequency increases, which is capacitive.

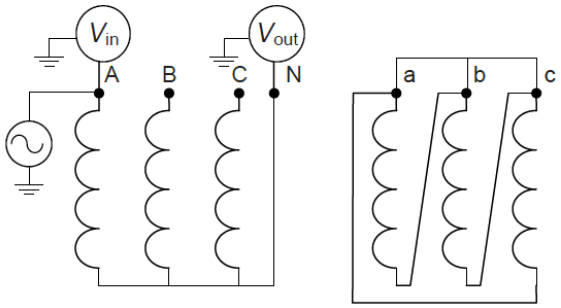


Figure 1-14 End-to-End Short-Circuit Measurement [50]

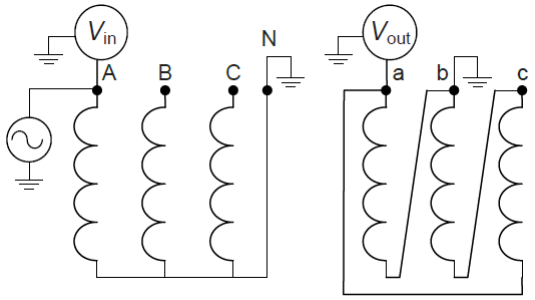


Figure 1-15 Inductive Inter-Winding Measurement [50]

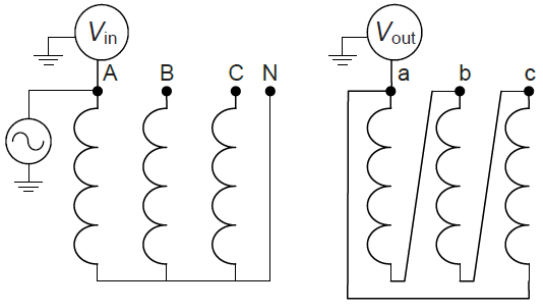


Figure 1-16 Capacitive Inter-Winding Measurement [50]

Among all the connection methods, the end-to-end measurement is recommended by IEC standard 60076-18 to be carried out on each winding of each phase. The other measurement methods can be used as supplementary measurements in diganosis [50].

An example of measured frequency response is plotted in Figure 1-17, both in logrithmic and lineaer frequency scale. Only the magnitude spectrum is given here as the phase spectrum is regarded as redundancy information. The frequency range is from 5 Hz to 200 kHz. This frequency response is measured on a 400/275/13 kV autotransformer, B phase, HV to LV terminals, with the other neutrals earthed.

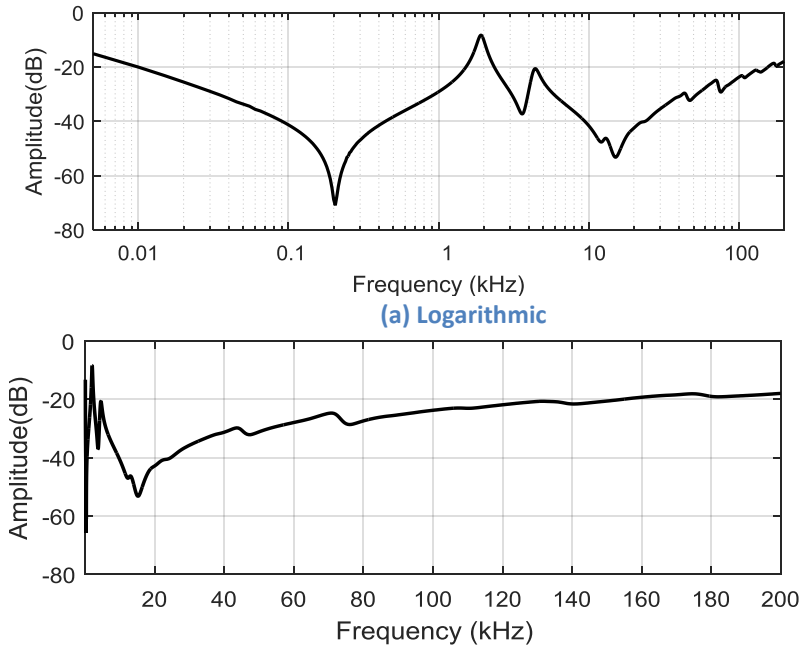


Figure 1-17 Example of Measured Frequency Response

The logarithmic frequency scale plot in Figure 1-17 (a) is useful in showing the overall shape and trend of the frequency repsonse, however it tends to over-emphasise the low and mid- frequency regions. On the other hand, depending on the upper frequency limit, the linear frequency scale plot in Figure 1-17 (b) also squeezes low frequencies in a narrow space in the plot and therefore could neglect the characteritiscs in the low frequency region. Consequently it is recommended to use both linear and logrithmic scales in diagnosis.

Figure 1-18 plots the end-to-end frequency responses of LV windings of two identical single-phase generator transformers from a bank of single-phae units, i.e. the B phase and C phase. [51]. It can be seen that the resonant points of the B phase frequency response deviate from that of the C phase freuqnecy response, both in the frequency location and the

magnitdue. The buckling of the LV winding is shown in Figure 1-19 [51]. This is a good example to demonstrate that the FRA technique is sensitive to detect the mechanical fault in the transformer winding.

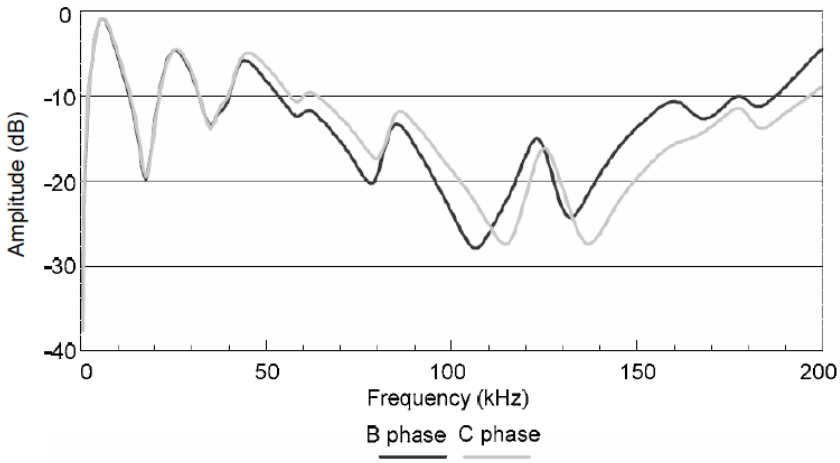


Figure 1-18 Comparison of Frequency Responses from LV Winding of Faulty and Healthy Phases [51]

Figure 1-19 Buckling of Inner LV Winding from B Phase [51]

1.2.2 Data Storage

According to IEC 60076-18 standard [50], the lowest frequency measurement shall be at 20 Hz or lower than 20 Hz. For test objects with a maximum voltage greater than or equal to 72.5 kV, the minimum highest frequency measurement should be 1 MHz. For test objects with a maximum voltage less than 72.5 kV, the minimum highest frequency measurement should be 2 MHz. As suggested, for compatibility and simplicity for all test objects, a highest measurement frequency of at least 2 MHz should be utilised. For

freuqency range below 100 Hz, intervals of measured data points should be smaller than 10 Hz; For freuqency range above 100 Hz, at least 200 measurements shall be made in each decade of frequency, approximately evenly distributed on either a linear or logarithmic scale. If the transformer operator does not require the low frequency information used to diagnose changes in the core, then a lower measurement frequency of not less than 5 kHz may be specified for the measurement. According to the IEEE standard C57.149 [49], the test equipment should be able to sample at least 200 measurement points per decade, either distributed linearly or logarithmically.

For the measuring equipment, IEC 60076-18 standard requires that the signal to noise ratio should be no less than 6 dB, and the dynamic range of the magnitude ratio range should cover -90 dB to 10 dB. The accuracy of the magnitude ratio should be better than ± 0.3 dB from -40 dB to 10 dB, and ± 1 dB from -80 dB to -40 dB, while the accuracy of the phase difference should be ± 1°. For the accuracy of frequency, the error should be less than ± 0.1%.

HP network analyser HP-4195A, Doble M5400 test set, and OMICRON FRAnalyzer and FRANEO 800 are the commonly used FRA measurement devices.

HP network analyser HP-4195A can be used to measure the frequency response from 10 Hz up to 500 MHz, with 0.001 Hz resolution. The dynamic range of magnitude ratio is larger than 100 dB. The phase difference between ± 180° can be measured. For this device, the accuracy of the magnitude ratio and phase difference is ± 0.05 dB and ± 0.3°, at the resolution 0.001 dB and 0.01°. This device was used for the FRA measurement investigated in this PhD study. Five frequency regions were recorded for the investigated FRA measurement, i.e., 5 Hz-2 kHz, 50 Hz-20 kHz, 500 Hz-200 kHz, 5 kHz-2 MHz and 25 kHz-10 MHz. For each frequency region, 400 evenly linearly distributed points are measured.

The Doble M5400 test set provides a frequency response measurement from 10 Hz to 25 MHz. It measures the frequency response at logarithmically spaced intervals of 1.2% of the frequency range specified. The dynamic range of this device is larger than 90 dB, with a ±1 dB accuracy at -80 dB [53].

The OMICRON FRAnalyzer [54] has a frequency measurement range of 10 Hz to 20MHz. its dynamic range is larger than 120dB. When the magnitude ratio is down to -50 dB, the accuracy reaches ±0.1 dB difference. And when the magnitude is between -50dB to -80

dB, the error is within ±1 dB. The distribution of measurement points and the frequency ranges can be customized. For each frequency region, the maximum quantity of measurement data points is 400.

The newest FRA measurement device from OMICRON is FRANEO 800 [55]. Its measurement frequency range is from 1 Hz to 30 MHz. For this device, higher precision measurements can be performed, with an accuracy of ± 0.5 dB down to -100 dB. The low noise ensures that even strong attenuated measurement traces can be measured with high accuracy. Thereby, FRANEO 800 is able to achieve the best dynamic range (> 150 dB) for SFRA.

Due to the diverse choice of measurement devices and customised measurement settings, the stored frequency response data may vary largely in terms of quantity and location of measurements points. In order to compare two sets of frequency response data using numerical indexes, the two sets of data needs to be in the same format.

1.3 Aims and Objectives

Although FRA has been developed for some time for the detection and diagnosis of the mechanical integrity of windings, there has been no IEC or IEEE standard established for the interpretation of the frequency responses.

Numerical indices have been used to indicate the similarities and quantify the alteration between two frequency response traces. Usually a single numerical value is used. Nevertheless, the whole measurement frequency range can be split into several frequency regions and the numerical indices can be extracted for each frequency region separately, as stated in the Chinese standard DL/T 911-2016 [56]. Such practice is based on the knowledge that different frequency regions are controlled by different physical factors.

The underlying idea behind the development of numerical indices is that the type, location and severity of mechanical faults may be indicated through the accumulated experience on the corresponding changes in numerical indices.

However, the values of numerical indices can be varied easily by varying the measured frequency points. For example, in the frequency region controlled by transformer core, the key information is the frequency location and the magnitude of the first antiresonance (trough). A large amount of frequency points on the dropping and rising parts before and

after the first trough can be viewed as redundant, and the number of sampling points can affect the absolute value of the numeric index. These details require careful consideration before the application of numeric indices, especially when setting the “healthy” or “deformed” criteria.

In addition, AI methods have also been applied for the interpretation of FRA measurement data, which require the data to be processed in a unified format.

FRA analysers on the market use different frequency and amplitude resolutions, which may lead to the fact that the FRA traces can have different number of frequency points, thus data stored in a database with different formats, cannot be analysed by numerical indices or auto-processed by AI techniques. In such a scenario, data pre-processing is needed so that the measurement data with different frequency resolution can be reproduced in the same desired format.

The redundant and inconsistent format of the FRA measurement data calls for a solution, i.e. the transfer function estimation method. The transfer function describes the FRA trace mathematically with dozens of parameters, and any changes in the FRA trace can be reflected precisely by the changes in the transfer function parameters.

Noticeably, the boundary frequencies between frequency regions which are dominated by different physical factors are empirical. For example, for 400/275/13 kV transformers, especially the common and series windings, the frequency regions lower than 2 kHz, 2 kHz - 20 kHz, 20 kHz - 1000 kHz are dominated by transformer core, interaction between windings influence, winding structure respectively, and the region higher than 1 MHz is governed by the measurement setup.

The boundary frequencies are varied depending on the voltage ratio, power rating, and construction type of transformers. The difference in the frequency response traces may be resulted from the difference in the construction of transformers, which is the main reason for us not to generalise the accumulated knowledge from one transformer to another. Therefore, it is important to develop fundamental understanding of transformer winding structure, equivalent circuit, winding movement including deformation and displacement types, and FRA technique.

Transformer design information, especially the winding construction type, is the manufacturers’ safe guarded know-how. It is well known that different winding construction types are susceptible to different modes of mechanical deformations. The AI

interpretation techniques, which were developed for the diagnosis of winding fault, tends to be usually applicable only for one type of transformer. Hence knowing the winding design information is helpful for transformer fault diagnosis. Besides, appropriate asset management method can extend the lifetime of transformers, and the same asset management method can be applied to transformers with same winding construction types. Thus the winding design information is also useful for asset management.

Over the years, FRA measurement data have been collected by utilities into their databases of transformer asset, with or without design information. Better interpretation of faults of transformer winding and effective asset management, especially for those transformers without any technical support from the Original Equipment Manufacturer (OEM), calls for the development of non-intrusive winding construction type recognition techniques.

The main objectives of the PhD thesis are:

- Develop transfer function estimation method and make it possible for visualisation of key parameters. The data format of any stored frequency response will be unified and a unique solution of transfer function which is physically achievable will be developed (Chapter 4);
- Take advantage of a large database to develop fundamental understanding of FRA. The influence of the voltage, power ratings, and winding construction types of transformers will be studied. The accumulated experience from one transformer will be analysed to see if it can be generalised to be applicable to another (Chapter 3);
- Apply Support Vector Machine (SVM) and Hierarchical Clustering (HC) methods on a large database to identify winding construction types. The SVM model as a supervised machine learning method, will be trained and tested by FRA data from windings of known types, and it is then applied to other FRA data to identify the winding construction type. On the other hand, the HC method is an unsupervised machine learning method, the frequency response will be processed by the HC method for clustering according to their distances. (Chapter 5 & 6).

1.4 Methodology

The following methodologies are adopted by the PhD thesis.

1. Two transfer function estimation methods have been developed.

1) The first method extracts complex zeros and poles from frequency regions and then combines them together to create a Feature Transfer Function, which corresponds to the resonance and anti-resonance of FRA traces. Then a Difference Transfer Function is estimated to correct the difference between Feature Transfer Function and the measured data. It gives satisfactory match between the measured FRA traces and estimated transfer function. However, the zeros and poles predicted do not appear sequentially, and may be mathematically feasible but not physically achieved by a circuit.

2) Thus the second method is developed, which first initialises a real pole according to the starting low-frequency-range magnitude of FRA trace, and then initialises complex zeros and poles according to the local minimums' and maximums' frequency locations and magnitudes. Then, the mutual influence is eliminated between complex zeros and poles. Afterwards, the shift of resonance and antiresonance are considered. The transfer function developed in this way has both mathematical and physical advantages in terms of achievability.

2. Two winding construction type classification methods have been developed.

1) Support Vector Machines (SVM) method belongs to the category of supervised machine learning method. A model is built, using the frequency responses of known winding construction types as the training data set, and this model can be used to predict the unknown FRA traces' winding construction type.

2) Hierarchical clustering method belongs to the category of unsupervised machine learning method. Frequency responses with similar distance are grouped together. Frequency responses with unknown winding construction type should share the same winding construction type with a frequency response with known winding construction type in the same cluster.

1.5 Outline of Thesis

Chapter 2 Literature Review

This chapter summarises and reviews the relevant research carried out by different researchers. There are three main parts of this chapter. The first part describes the various numerical indices which have been employed to evaluate the frequency responses. The second part compares the functionalities of the three types of models which have been developed for the frequency response analysis, i.e. white box model, grey box model and finite element model. The last part summaries the application of Artificial Intelligence (AI) in the FRA filed, such as parameter estimation for FRA models, identification of winding faults etc.

Chapter 3 FRA Database of UK National Grid Company

This chapter introduces how the frequency responses are stored in the UK National Grid Database. Next, the basic information of transformers from the database are analysed, including their voltage ratio, power rating and winding construction types. Then, the fundamental understanding on the influence of transformer design on the frequency responses, especially the influence from winding construction types, are analysed. Lastly, the frequency responses used in Chapter 4, 5 and 6 are introduced,

Chapter 4 Transfer Function Estimation

In this chapter, two methods to estimate the transfer function are introduced, and for each method one paper is produced.

The first method is based on the MATLAB command 'invefreqs'. The whole frequency ranges are divided into several frequency regions, and the key parameters, complex zeros and poles, are extracted from each frequency region to form a Feature Transfer Function. Afterwards, the difference between the Feature Transfer Function and measured frequency responses are corrected by a Difference Transfer function.

The second method detected the peak and trough on the FRA trace. One peak corresponds to a pair of complex poles whilst one trough corresponds to one pair of complex zeros. The format of the transfer function for each pair of complex zeros and poles and for the real poles is defined, in order to find a unique solution.

Both methods produce satisfactory matching results. However, although the first method gives a well- matching transfer function, it may produces parameters without physical meaning, which means the zeros and poles not appear in sequence and the real part of complex zeros and poles may be positive. The second method uses FRA magnitude data

only, and it is able to find a unique physically achievable solution. However, the second method is not able to match the phase data which ranges from -180° to +180°.

Chapter 5 Identification of Winding Construction Types by Supervised Machine Learning Method

In this chapter, winding construction types are classified using supervised machine learning method Support Vector Machine (SVM). Different winding construction types own different equivalent electrical parameters, and thus different features of FRA traces are produced. An identification model is built according to the frequency responses with known winding construction types. This model consists of several binary classifiers to enable multiclass classification, and the quantity of the binary classifiers is determined by the quantity of winding construction types to be identified. For a FRA trace with unknown winding construction type, its winding construction type is predicted according to the votes of the binary classifiers. Cross validation is used to evaluate the effectiveness of classification model by constructing different SVM identification models using different training and testing data. One of the constructed models is discussed in detail, in which 54 frequency responses from 4 winding construction types are used as training data to build the identification model, and another 54 frequency response are used as testing data to verify the model, achieving 100% accuracy. The support vectors are the most important input frequency responses, which decide the parameters of the SVM model built. Sensitivity studies are carried out to investigate the support vectors.

Chapter 6 Identification of Winding Construction Types by Unsupervised Machine Learning Method

In this chapter, the winding construction types can be identified using unsupervised machine learning method. Hierarchical Clustering is an unsupervised algorithm which groups similar observations together. 28 frequency responses from 5 winding construction types are investigated. The distance between every two frequency responses are calculated and the closest two frequency responses are clustered. The frequency responses are merged together until only group is left. The cut off threshold can be either the distance between frequency responses or the quantity of clusters left. Frequency responses of winding with unknown construction type, clustered together with frequency responses of known winding construction type, should share the same winding type. The frequency responses are clustered on different frequency regions controlled by different influencing factors i.e. the core and the winding structure. It was found that the high voltage winding, low voltage

winding and tertiary winding can be clustered into different groups. In addition, the windings with high series capacitance (Multiple Layer and Interleaved Disc) and the windings with low series capacitance (Plain Disc and Intershielded Disc) can be clustered into different groups.

Chapter 7 Conclusions and Future Work

This chapter concludes the work by this PhD study and suggests the direction for the work to be conducted.

Chapter 2 Literature Review

2.1 Introduction

FRA is a comparative technique. It firstly needs to find a reference for comparison purpose, then identifies the differences between the current measurement result and the reference, and lastly interprets the differences and provides diagnosis. The relevant research work on FRA will be reviewed and summarised in this chapter, as shown in Figure 2-1.

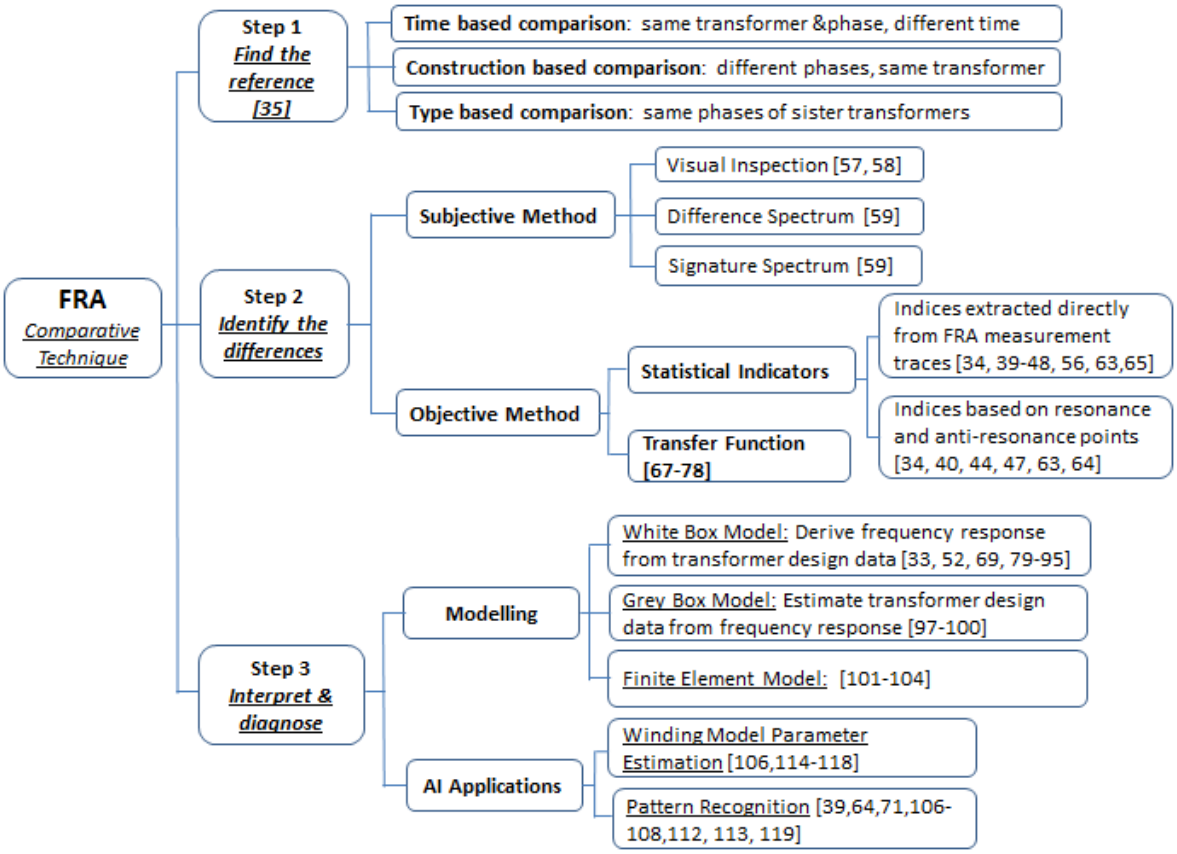


Figure 2-1 Summary of Research on FRA

As stated, the measured frequency response is to be compared with its fingerprint in a database. The fingerprint is also called the baseline/reference measurement. There are mainly three categories of comparison for FRA measurement results, i.e. the time based comparison, the construction based comparison and the type based comparison, as outlined in Figure 2-2. Time based comparison compares two frequency responses of the same transformer same phase from different time; type based comparison compares two frequency responses of the same phase from the diagnosed transformer and its sister

transformer; construction based comparison compares two frequency responses from different phases of the same transformer [35].

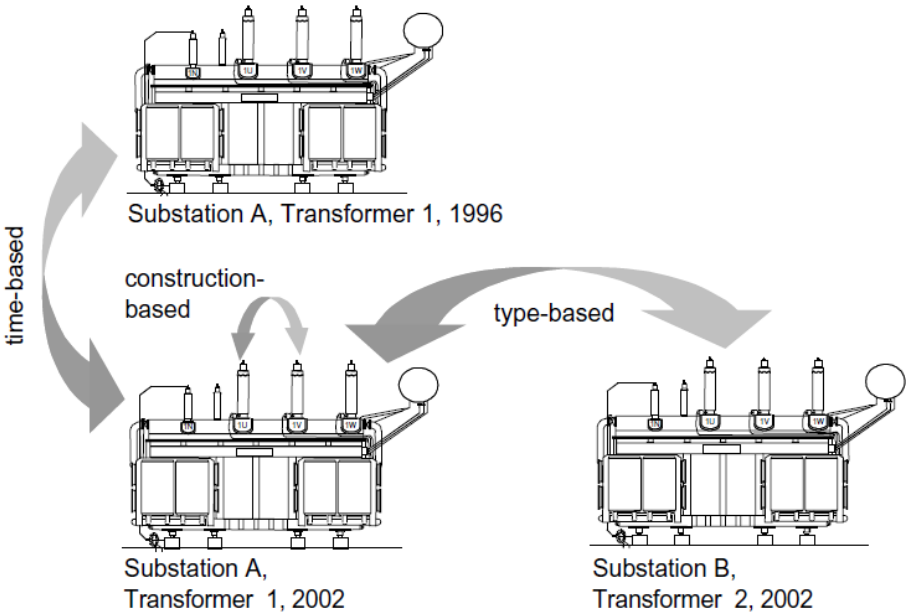


Figure 2-2 FRA Measurement Results in Different Comparison Types [35]

When identifying the differences, the emphasis is usually on the analysis of the magnitude while little attention is paid to the phase spectrum, due to their high level of redundancy to each other. The local maximum of magnitude spectrum is called the peak, while the local minimum of magnitude spectrum is called the trough. Any shifting, disappearing or appearing of peaks or troughs may indicate the existence of winding related faults, mechanical or even electrical. Most of the time although phase spectra are calculated and stored after each transformer FRA measurement, they tend to be ignored and only magnitude spectra are analysed in depth. However, a recent publication emphasizes the importance of phase spectrum in diagnosis [41].

The comparative tools of FRA can be divided into two types i.e., the subjective method and the objective method. The subjective methods rely on the observation of plots, such as the direct comparison of the reference and the diagnosis FRA traces, or any difference plots produced using the reference and the diagnostic traces. The objective method refers to the method uses mathematical parameters to describe the FRA trace or its alteration.

The subjective tools include visual inspection [57, 58], difference spectrum [59] and signature spectrum [59]. It is suggested in [57] that the expert normally observes the shape of the curve and the resonant frequencies to judge the existence of fault. The change in the magnitude of the frequency response, appearance or disappearance of the resonant

frequencies, and the shift of existing resonant frequencies are used as criteria. The difference spectrum is obtained by subtracting the reference FRA trace from the diagnostic FRA trace. Minor alteration can be seen in the high frequency range. For example, the relocation of leads was reflected in frequency range higher than 2 MHz. Major faults can be seen in the low frequency range. For example, High Voltage middle winding grounded to earth and interturn faults at the middle of High Voltage winding are reflected in frequency range lower than 2 kHz [59]. To produce the signature spectrum, at each frequency, the frequency responses of the three phases are averaged, the differences between this mean value and those of the three phase frequency responses are squared and then summed up, and then the square root of the sum against frequency is the so-called signature spectrum. For the healthy transformer, its signature spectrum has a smaller magnitude than that of the faulty transformer. In the example provided by the author, the healthy transformer's signature spectrum lay lower than 20 dB while the faulty one is located around 100 dB [59].

The subjective methods highly rely on the experts' experience. Due to this limitation, objective methods have been the focus of recent research and development. The objective methods use the quantitative numbers to reflect the difference or similarity between the baseline and the diagnostic measurements, and thus it gives a more direct and more accurate conclusion, rather than relying on experience only. The objective tools use numerical indices for the comparison of FRA traces. Those indices are extracted base on comparison of the measured frequency responses. However, the criterion of numerical indices to distinguish 'healthy or faulty' is still a challenge.

To interpret and diagnose the FRA measurement traces, experiments can be carried out for sensitivity study, using manually created winding faults, but simulation is a more economical way for such investigation. The transformer winding white box models are constructed based on the design data of a transformer. Once enough FRA measurement data are collected, either produced by experiments or simulation, experience can be accumulated regarding the changes of frequency response traces caused by particular faults. AI techniques have been applied widely for the recognition of fault patterns of transformer windings through frequency responses. Besides, it has also been investigated how to estimate the transformer design data from frequency responses, by researchers who are keen to know the transformer design data. This is the so-called grey box model. AI techniques have also played an important role in the estimation of transformer design data by the grey box model.

2.2 Numerical Indices for FRA Comparison

Numerical indices, including the single statistical indicator and the transfer function, have been used to quantify the result of frequency response comparison. A statistical indicator is a single value which reflects the similarity or dissimilarity between two frequency response measurements. A transfer function uses a set of parameters to mathematically describe one frequency response measurement, and the two sets of parameters are compared between the reference and diagnostic frequency response measurements.

2.2.1 Statistical Indicators Methods

The statistics indicators can be divided into three types, indices extracted directly from FRA traces, indices based on resonance and anti-resonance points, or indices base on transfer functions.

Table 2-1 summarises the first type of statistical indicators, which have been used for the comparison of FRA traces [60, 61]. X and Y are the magnitude response of the diagnostic and reference FRA data normally in dB scale, f is the vector of frequency samples, df is the sampling interval of frequency, N is the number of sample points, and φ_X and φ_Y are the phase vectors of the FRA data. All the statistic indicators are extracted from magnitude response but the complex distance, CD in Table 2-1, which uses both magnitude and phase vectors. Noticeably, for X and Y, their values may be or may not be in dB, which might lead to a different calculation result.

The correlation coefficient (CC), standard deviation (SD) and absolute sum of logarithmic error (ASLE) are three widely used indicators. The closer the correlation coefficient is to 1, the closer is the similarity of the two groups of data. In [62], the two groups of measured FRA data are viewed same if the value of correlation coefficient of magnitude response is larger than 0.995. The stricter criteria are used in [42], where the correlation coefficient value of 0.9998 for magnitude and 0.95 for phase are set as the boundary for identifying winding displacement. The standard deviation is a quantity describing how much the members of a group differ from those of the other group. In [42], a low value of standard deviation, lower than 1 for magnitude and lower than 10 for phase, is used to indicate that two groups of data resemble each other. If standard deviations are larger, the existence of winding deformation is diagnosed. The absolute sum of logarithmic error is an indicator that reflects the logarithmic difference between two groups of data. For magnitude, a

ASLE value higher than 0.4 for magnitude and higher than 0.8 for phase suggests that a fault has occurred in the winding [42].

Table 2-1 Numerical Indices Extracted from FRA Traces [60, 61]

Abbreviation	Definition	Expression Equation	Reference
ED	Euclidean distance	$ED = \sqrt{\sum_{i=1}^{N}(Y_i - X_i)^2}$	[39, 40]
CD	Complex distance	$CD = \sqrt{\sum_{i=1}^{N} \begin{matrix}(|X_i cos\varphi_{X_i} - Y_i cos\varphi_{Y_i}|^2 \\ +|X_i sin\varphi_{X_i} - Y_i sin\varphi_{Y_i}|^2)\end{matrix}}$	[41]
SD	Standard Deviation	$SD = \sqrt{\frac{1}{N-1}(\sum_{i=1}^{N}(Y_i - X_i)^2)}$	[42]
ID	Integral of difference	$ID = \int (Y(f) - X(f))df$	[34, 43]
IA	Integral of absolute difference	$IA = \int |Y(f) - X(f)|df$	[34]
SDA	standardised difference area	$SDA = \frac{\int |Y(f) - X(f)|df}{\int |X(f)|df}$	[44]
ASLE	Absolute sum of logarithmic error	$ASLE = \frac{1}{N}\sum_{i=1}^{N} |20log_{10}Y_i - 20log_{10}X_i|$	[42, 45]
RMSE	Root mean square error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\frac{(|Y_i| - |X_i|)}{\frac{1}{N}\sum_{i=1}^{N}|X_i|})^2}$	[46]
E σ_e	Expectation Standard deviation	$E(\Delta) = \frac{1}{N}\sum_{i=1}^{N}\Delta_i$ $\Delta_i = \frac{|Y(i)| - |X(i)|}{\frac{1}{N}\sum_{i=1}^{N}|X(i)|}$ $\sigma_e = \sqrt{var(\Delta)} = E(\Delta - E(\Delta))$	[47]
σ	Stochastic spectrum deviation	$\sigma(\%) = \frac{100}{N}\sum_{i=1}^{N}|\frac{(Y_i - X_i)}{X_i}|$	[48]
MAX	Maximum of difference	$MAX(Y_i - X_i)$	[40, 43]

ρ	Correlation factor	$\rho = \frac{\sum_{i=1}^{N} X_i^* Y_i^*}{\sqrt{\sum_{i=1}^{N}(X_i^*)^2 \sum_{i=1}^{N}(Y_i^*)^2}}$ $Y_i^* = \lvert Y_i \rvert - \frac{1}{N}\sum_{i=1}^{N} Y_i$ $X_i^* = \lvert X_i \rvert - \frac{1}{N}\sum_{i=1}^{N} X_i$	[44, 47]
R	Relative Factor	$R = \begin{cases} 10, & 1-\rho < 10^{-10} \\ -\log_{10}(1-\rho), & others \end{cases}$	[56]
CC	Correlation Coefficient	$CC = \frac{\sum_{i=1}^{N} X_i Y_i}{\sqrt{\sum_{i=1}^{N} {X_i}^2 \sum_{i=1}^{N} {Y_i}^2}}$	[42, 45, 63]
SSE	Sum squared error	$SSE = \frac{1}{N}\sum_{i=1}^{N}(Y_i - X_i)^2$	[40, 45]
SSRE	Sum squared ratio error	$SSRE = \frac{1}{N}\sum_{i=1}^{N}(\frac{Y_i}{X_i} - 1)^2$	[39, 40, 45]
SSMMRE	Sum squared max–min ratio error	$SSMMRE = \frac{1}{N}\sum_{i=1}^{N}(\frac{\max(Y_i, X_i)}{\min(Y_i, X_i)} - 1)^2$	[45]

The statistical indicators Correlation Coefficient (CC), Standard Deviation (SD), Expectation (E) and Maximum of Difference (MAX) are used in Chapter 4, to evaluate how well the estimated transfer function matches to the measured FRA data.

Different from the fist type of indicators using all the sampled data points, the second type of indictors focuses only on the resonant points, i.e., the peaks and troughs on the FRA traces. The second type of statistical indicators is summarized in Table 2-2. X and Y correspond to reference and diagnostic FRA measurements, A and f are the magnitude and frequency vectors of resonant points, K is the number of resonant points, and AF is the area below the FRA trace between two neighbour anti-resonance points, as shown in Figure 2-3 [47] which is measured on a simulated 400kVA coil block, w is the weight of resonant points for amplitude and frequency.

Table 2-2 Indices Based on Resonance and Anti-Resonance Points [60, 61]

Abbreviation	Definition	Expression Equation	Reference
MDA	Mean deviation of areas	$MDA = \frac{1}{K}\sum_{i=1}^{K}\lvert AF_{Yi} - AF_{Xi}\rvert$	[47]
MAD	Mean amplitude deviation	$MAD = \frac{1}{K}\sum_{i=1}^{K}\lvert A_{Yi} - A_{Xi}\rvert$	[47]
MFD	Index of amplitude deviation	$MFD = \frac{1}{K}\sum_{i=1}^{K}\lvert f_{Yi} - f_{Xi}\rvert$	[47]
IAD	Index of amplitude deviation	$IAD = \sum_{i=1}^{K}\left\lvert\frac{A_{Yi} - A_{Xi}}{A_{Xi}}\right\rvert$	[44, 63]
IFD	Index of frequency deviation	$IFD = \sum_{i=1}^{K}\left\lvert\frac{f_{Yi} - f_{Xi}}{f_{Xi}}\right\rvert$	[44, 63]
$\boldsymbol{F_a}$	Amplitude function	$F_a = \sum_{i=1}^{K}\frac{A_{Yi}}{A_{Xi}}$	[64]
$\boldsymbol{F_f}$	Frequency function	$F_a = \sum_{i=1}^{K}\frac{f_{Yi}}{f_{Xi}}$	[64]
$\boldsymbol{W_a}$	Weighted amplitude function	$W_a = \sum_{i=1}^{K}\frac{A_{Yi}}{A_{Xi}}w_{ai}$	[34, 40, 44]
$\boldsymbol{W_f}$	Weighted frequency function	$W_f = \sum_{i=1}^{K}\frac{f_{Yi}}{f_{Xi}}w_{fi}$	[34, 40, 44]

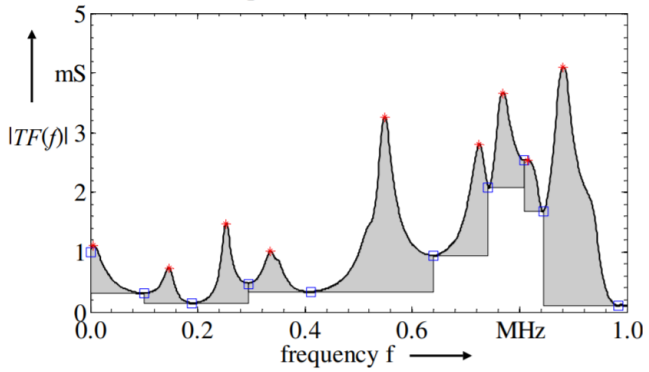


Figure 2-3 Areas between Two neighbour Antiesonance Points [47]

The third type of statistic indicators are used when the transfer function expression of FRA traces is available. This type of indicators will be reviewed in next section regarding transfer function. The other two types of statistic indicators are reviewed in this section.

In 1995, J. Bak-Jensen and B. Bak-Jensen from the University of Aalborg, Denmark introduced the statistical indicator, stochastic spectrum deviation, σ, to evaluate the frequency response, using the magnitude spectrum in dB [48]. Experiments were carried out on a single phase transformer with $20/\sqrt{3}$ kV primary voltage and $110/\sqrt{3}$ kV secondary voltage. In their study, sensitivity test was carried out to investigate the influence on frequency response from insulation oil, the core, the winding/coil insulation. σ of frequency region 100 Hz to 10 kHz and frequency region 10 kHz to 1 MHz was calculated separately. After the insulating grease is removed, σ reached 3.807% and 0.760% for frequency region10 Hz to 10 kHz and 10 kHz to 1 MHz. For different depths of grooves sawed in the core from 1mm to 8 mm, σ reached 0.443% at most, hardly reflecting the influence. When two turns on neighbour winding layer were short circuited, σ could be as large as 12.63%. It is understandable that the spectrum deviation is more sensitive to winding related electrical faults. The change in frequency response was studied during the accelerated ageing procedure leading towards breakdown. After breakdown, σ increased up to about 5% for frequency region 100 Hz to 10 kHz and about 0.3% for frequency region 10 kHz to 1 MHz. It is summarised that successful recognition of failures and ageing problems in transformers relies on measurable variations in characteristic parameters – the winding capacitances, total iron losses and core reluctance.

In 1999, D.K. Xu, C.Z. Fu, and Y.M. Li from Xi'an Jiaotong University, China used the correlation coefficient (CC) and standard deviation (SD) as the inputs of the Artificial Neural Networks (ANN) algorithm for discriminating winding health conditions [65]. The faults were simulated by connecting capacitors between the tap points, and between tap points and earth points, on a 35/10kV three-phase model transformer with 54 tapping points in the primary winding of B phase. The Correlation Coefficient and Standard Deviation on the low, middle, high and the whole frequency ranges, all together 8 parameters, are used as the input of a three layer ANN system with feed forward connections. A binary output suggests either a 'NORMAL' or 'ALARM' diagnosis result for the healthy condition of the transformer. The predication accuracy was 95%, with 24 training patterns and 20 testing patterns. It was suggested that various training patterns should be collected for better identification of faults. According to authors, the method is effective, simple, robust and easy to apply.

In 2004, objective quantitative diagnosis criteria were proposed in the Chinese FRA Standard DL/T 911-2016, using relative factor R [56]. The relative factor R is defined

using the correlation factor ρ, as shown in Table 2-1, to identify the health condition of transformer windings. The relative factors R_{LF} in low frequency region from 1 kHz to 100 kHz, R_{MF} in middle frequency region from 100 kHz to 600 kHz, and R_{HF} in high frequency region from 600 kHz to 1000 kHz are calculated. The winding health conditions can be classified into four different stages according to the values, as shown in Table 2-3. Though the criteria have been approved to be able to identify the radial and axial displacement [47], research also suggested that the Chinses Standard DL/T 911-2016 may fail to identify some major faults, such as the mechanical failure with abnormalities in high frequency region [63].

Table 2-3 Relative Factor and Winding Health Condition [56]

Winding health condition	**R**
Severe deformation	$R_{LF} < 0.6$
Obvious deformation	$0.6 \leq R_{LF} < 1$ or $R_{MF} < 0.6$
Slight deformation	$1.0 \leq R_{LF} < 2$ or $0.6 \leq R_{MF} < 1$
Normal deformation	$2.0 \leq R_{LF}$, $1.0 \leq R_{MF}$ and $0.6 \leq R_{HF}$

In 2005, J.W. Kim from Technical Research Laboratories, Korea, B. Park from SK Teletech, Korea, C.J. Jeong, S.W. Kim and P. Park from Pohang University of Science and Technology, Korea used Sum Squared Error (SSE), Correlation Coefficient (CC), Sum Squared Ratio Error (SSRE), Sum Squared Max-Min Ratio Error (SSMMRE) and Absolute Sum of Logarithmic Error (ASLE) for the fault diagnosis of power transformer [45]. When calculating the indices, only the magnitude response was used. With the aim to distinguish the faulty phase, experimental tests were carried out with artificial electrical faults, including creating short circuit between HV bushing and tank, between LV bushing and tank, between tap and tank, inter-tap short circuit, between turn and tank, and inter-disc short circuit. Two transformers were studied. One is a three-phase 300 kVA 3300/440 V delta-wye transformer and the other is a 2 MVA 3150/460 V wye-delta transformer. Among all the statistical indicators, ASLE has the best performance with the highest accuracy in the diagnosis of aforementioned fault types. The authors also concluded that SSE often underrates the important information scattered around the trough or lower values in a magnitude response. CC is not suitable for the comparisons of similar patterns with different magnitudes, because it neglects the deviation in magnitude and emphasises the

similarity in shape. SSRE and SMMRE are altered versions of SSE. SSRE does not have the disadvantage of CC, but it is dominated by those Y_i/X_i ratios whose values are large than 1, as shown in Table 2-1. The drawback of SSMMRE is that its objectivity is affected by the deviation of high-frequency components.

In 2006 E. Rahimpour from Iran Transformer Research Institute, Iran and D. Gorzin from University of Zanjan, Iran proposed a new method, Weighted Amplitude function W_a and Weighted Frequency function W_f based on the resonance and anti-resonance points, as listed in Table 2-2, to compare the frequency responses and detect the location and level of winding faults, i.e., axial displacement [34]. A disc HV winding and a layer LV winding were used in the research. The windings were constructed for the aim of experimental test, and approximately correspond to windings of a transformer with rated voltage of 10 kV and rated power of 1.3 MVA. The presented method is independent of the frequency response types, and the errors from measurement do not need to be taken into consideration. The authors believe that this weight function method is more reliable and effective than other methods. However, the research is based only on the assumption that the amount of resonance and anti-resonance points remains unchanged.

In 2007, R. Wimmer, S. Tenbohlen and K. Feser from University of Stuttgart, Germany used the statistic indicators Expectation E, Standard deviation σ_e, correlation factor ρ to evaluate the frequency response [47]. The correlation factor can be viewed as a normalised correlation coefficient. When the difference between two sets of data is zero or owns an average zero value, E is zero. It changes monotonically when the axial displacement happens, but not for the radial deformation. Thus this parameter can be used for the diagnosis of axial displacement. When two sets of data own constant damping deviation, σ_e is zero. σ_e shows monotonous behaviour both to axial displacement and radial deformation. The correlation factor ρ reflects the linear dependence between two sets of data, and it has an inverted characteristic to the Standard Deviation σ_e. Sensitivity studies were carried out on a set of winding coils with 400 kVA power rating, by shifting the low voltage winding 1 cm per step against the high voltage winding. It was concluded that the correlation factor is the most sensitive. The authors also suggested that this parameter becomes more sensitive when splitting the whole frequency range into several frequency regions.

In 2008, P.M. Nirgude, D. Ashokaraju from Central Power Research Institute, India, A.D. Rajkumar from Osmania University, India, and B.P. Singh from Bharat Heavy Electrical

Limited, India evaluated Correlation Coefficient (CC), Standard Deviation (SD) and Absolute Sum of Logarithmic Error (ASLE) techniques using both magnitude and phase responses [42]. Three case studies were carried out respectively on a 31.5 MVA, 132/33 kV three-phase transformer with Interleaved Disc HV winding and Continuous Disc LV winding, a model Continuous Disc winding, and two three-phase transformers by the same manufacturer rated 2.5 MVA, 33/11 kV and 16 MVA, 33/11 kV . The authors proposed that 1% variation in axial height and/or internal diameter from the reference should be used as the threshold to indicate faults in the windings. The corresponding criteria is set as that when CC is smaller than 0.9998 for magnitude and 0.95 for phase, SD is larger than 1.0 for magnitude and 10 for phase, ASLE is larger than 0.4 for magnitude and 0.8 for phase, an fault is indicated. In addition, the severity of displacement/deformation can also be concluded from the amount of variation of the parameters from the suggested critical values. These methods can also be used when time based comparison is not available, by construction or type based comparison. However, the analysis result is more reliable when time based comparison is accessible.

In 2010, E. Rahimpour from ABB, Germany, M. Jabbari from University of Zanjan, Iran and S. Tenbohlen from University of Stuttgart, Germany used multiple numerical indices extracted from FRA traces to detect the types of mechanical faults, including Standardised Difference Area (SDA), Correlation Factor ρ, Index of Frequency Deviation IFD, Index of Amplitude Deviation IAD, Weighted Amplitude Function W_a, Weighted Frequency function W_f [44]. In order to increase the sensitivity of proposed comparison method, the whole frequency range was divided into several frequency regions, and the aforementioned indices were extracted from each frequency region. One test object has an disks high voltage winding and a layer low voltage winding, and the FRA data was collected when the space between two adjacent disks were alternated in different dimension. Another similar test object was used to study different degrees of radial deformation. The Standardized Difference Area (SDA) is regarded as an unhelpful numerical index for FRA diagnosis, while all of the other numerical indices are useful. The Correlation Factor ρ can be used to detect the fault types but it is less reliable than other numerical indices. The IFD and IAF have similar properties. Nevertheless, they are not applicable for the diagnosis of disc-space variation. Frequency and amplitude weight functions, W_a and W_f, can both reflect the location and level of faults. It suggested that the combination of the numerical indices improves the accuracy and reliability of the diagnosis results, especially in terms of fault level.

In 2012, K. Pourhossein from Islamic Azad University, Iran, G. Gharehpetian from Amirkabir University of Technology, Iran, E. Rahimpour from ABB, Germany, and B. Araabi from University of Tehran, Iran investigated the properties of Integral of Absolute Difference (IA), Maximum of Difference (MAX), Correlation Coefficient (CC), Weighted Amplitude Function (W_a), Weighted Frequency Function (W_f), Sum Squared Error (SSE), Sum Squared Ratio Error (SSER) and Euclidean Distance (ED) [40]. Two single phase transformer manufactured for special experimental purposes were used for the detection of axial displacement and radial deformation separately. The one for axial displacement study had Plain Disc HV winding and Multiple Layer winding, and the other for radial deformation had HV and LV windings both of Plain Disc type. The investigated transformers owned oil-immersed cylindrical tank. To identify the fault types, regional Euclidean Distance was used, which is the Euclidean Distance extracted from different frequency regions. It was found if all the indicators are standardized between 0 and 1, Euclidean Distance has the best linearity against the extent of defect of axial displacement. Therefore, it is possible to determine the extent of defect by the value of the index.

In 2016, S. Tenbohlen from Stuttgart University, Germany and M.H. Samimi from University of Thran, Iran introduced the index Complex Distance (CD) using both magnitude and phase data to evaluate frequency responses [41]. Experiments were carried out on the high voltage and low voltage windings corresponding to a 600 kVA transformer, with different level of axial displacement and disc space variation. The importance of phase data was emphasised, which is claimed to increase the diagnosis sensitivity and stability. Simulation models were used to verify and prove the conclusion. Complex Distance performs well even when uncertainties are introduced in the experiments. In 2017, the same authors used Correlation Coefficient (CC) and Euclidean Distance (ED) to test the sensitivity of different connection schemes, terminating resistors and measurement impedance, with different mechanical transformer faults, i.e., axial displacement, disc space variation and radial deformation [66]. Similar conclusion to [45] is drawn regarding the shortcoming of CC, in that the overall magnitude difference could not be reflected. Thus the authors used two indices together, i.e., CC and ED, to evaluate the frequency responses. The authors concluded that the capacitive inter-winding connection is the most sensitive connection method to winding mechanical faults, among the recommended standard connection methods. Larger terminal resistors and measurement impedance performed better, considering the injection of noise.

In this part of literature review, different statistical indicators used in FRA interpretation and diagnosis are summarised. Those statistical indicators have been applied in the FRA diagnosis to detect the fault type, location and level. Sometimes, the combination of several statistical indicators leads to a better identification result. However, when calculating the numerical indices, issue is that the magnitude response may be used in dB or not by the researchers. This may lead to differences in the interpretation and it is hard to identify the criterion.

2.2.2 Transfer Function Method

Transfer function method represents the FRA traces by mathematical equations, with a number of parameters. A rational transfer function in continuous frequency domain, or s domain, usually can be described as:

$$T_s = k_s \frac{\prod_{k=1}^{N1}(s-z_{rk})}{\prod_{j=1}^{M1}(s-p_{rj})} \times \frac{\prod_{k=1}^{N1}(s-z_{ck})(s-z'_{ck})}{\prod_{j=1}^{M1}(s-p_{cj})(s-p'_{cj})} \quad \text{Equation 2-1}$$

where k_s is the constant coefficient, z_{rk} and p_{rj} are real zero and real pole separately, and z_{ck}, z'_{ck} are conjugate complex zeros, p_{cj}, p'_{cj} are conjugate complex poles respectively. The variable s is related to frequency f and defined as $s=2\pi fi$.

There are three main algorithms for deriving transfer function, of a frequency response, i.e. the non-linear least square (NLS), vector fitting (VF) and sub-space representation.

In 1959, E.C. Levy from Space Technology Labs, USA invented the NLS method [67]. The NLS method approximates the frequency response use the ratio of two polynomials:

$$T_s = \frac{a_m s^m + a_{m-1}s^{m-1} + \cdots + a_1 s^1 + a_0}{b_n s^n + b_{n-1}s^{n-1} + \cdots + b_1 s^1 + b_0} \quad \text{Equation 2-2}$$

where the parameters a and b represent the coefficient of the transfer function numerator and denominator respectively. The method chooses a set of initial parameters, and then refines the chosen parameters by successive iterations. The sum of difference square between the estimated data and measured data are minimised, by setting the partial derivative to zero, to find an optimal solution. In 1992, J. Bak-Jensen, B. Bak-Jenson, S.D. Mikkelsen and C.G. Jensen from University of Aalborg, Denmark, firstly introduced theNLS method to the FRA field to parameterise the frequency response of a 50 Hz $(20/\sqrt{3})$kV/$(110/\sqrt{3})$V transformer [68]. A second or third order RLC equivalent circuit can be constructed according to the estimated transfer function for the frequency range

from 50Hz to 300 kHz [68]. However, the model only gave a rough approximation for the FRA trace, which means that the overall trend is matched while most of details on the FRA trace are neglected. Two reasons contribute to this drawback. One is the order of transfer function is too low to give an accurate match, the other is that the adopted circuit is too simple to demonstrate the details. However, this study was pioneering this research direction to estimate parameters for winding equivalent circuit modelling using the frequency response. In 1997, S.M. Islam from Curtin University of Technology, Australia, K.M. Coates and G. Ledwich from the University of Newcastle, Australia continued the modelling work by proposing a parametric system identification technique to model the frequency response [69]. The large frequency range, 50 Hz to 1 MHz was divided into low, medium and high frequency ranges. For each frequency range, the transfer function was estimated by the NLS method. Different models were built for different frequency ranges, and the corresponding electrical parameters are approximated according to the coefficients of the transfer function. However, due to the adopted models are of low order, the matching result was not accurate enough to reproduce the original complex FRA trace. Nevertheless, the alteration in parameters of the built model can still be correlated with transformer faults, which enables the transformer diagnosis study. The NLS method was also employed to construct equivalent circuit in [70] and to identify transformer fault types in [71].

In 1996, T. McKelvey, H. Akcay and L. Ljung from Linkoping University, Sweden developed two non-iterative state-space identification algorithms to parameterise frequency responses [72, 73]. Different from the transfer function representation which is limited to single-input single-output systems, a state space representation in matrix format can be applied to multiple-input multiple-output systems. The first algorithm used equally distributed data in the frequency domain and it can be used when the noise has zero mean. The second algorithm was able to deal with arbitrarily located data in the frequency domain, but it needed priori knowledge of noise. The transfer function in the frequency domain can be computed from the state space representation.

In 1999, Vector Fitting (VF) method was firstly proposed by B. Gustavsen and A. Semlyen from Electrical Power Research Institute, Norway and the University of Toronto, Canada [74]. It soon became the most widely used method of transfer function estimation in the frequency domain. VF uses partial fraction decomposition of a rational function in the following format:

$$T_s = \sum_{n=1}^{N} \frac{c_n}{S-p_n} + d + sh \qquad \text{Equation 2-3}$$

where p_n are poles which are either real or complex, c_n are their corresponding residues, both d and h are real numbers. a_n and c_n are achieved by renewing the pre-defined starting complex poles through least squares approximation. Poles and residues are identified at two different stages. This method is accurate, stable and effective. In 2006, B. Gustavsen improved the method by optimising the pole relocation procedure [75]. A more relaxed condition was adopted to replace the original asymptotic condition. By this improvement, the selection of initial poles becomes less important. In 2010, L.D. Tommasi from Energy Research Center, Netherlands, B. Gustavsen from SINTEF Energy Research, Norway, and T. Dhaene from Ghent University, Belgium, proposed a robust transfer function identification method [76]. In this method, only the magnitude spectrum was used. The estimated FRA phase trace may differ from the measured FRA data, since the method produces the minimum-phase shift function.

Researchers have started to apply Vector Fitting method to the analysis of FRA data after its invention. Studies have shown that different choices of transfer function orders for the same FRA trace may lead to different amounts and/or value of parameters. Standards should be set regarding the choice of transfer function's order, which should guarantee the accuracy as well as the conciseness of estimation. Moreover, it is reported this algorithm is sensitive to the predefined starting poles. This means that different starting poles for the estimation may lead to dissimilar final poles for the same FRA trace, which results in error when calculating indices. Hence, a standard method should also solve this problem so that all the rational functions which fit to a single FRA trace can be identical.

Satisfactory high order transfer function models can be produced by the three introduced estimation methods, i.e., non-linear least square (NLS), vector fitting (VF) and sub-space (SS). Through simulation of artificial winding faults, work can be conducted to associate the estimated parameters with different transformer faults. After the expression of transfer function is obtained, two indices can be used for the frequency response comparison, i.e., Sum of Absolute Displacement of Poles (SDP) and Faulted-Intact relation (FI):

$$SDP = \sum_{i=1}^{K} \left| \frac{P_{Yi} - P_{Xi}}{P_{Xi}} \right| \qquad \text{Equation 2-4}$$

$$FI = \left(\frac{\sum_{i=1}^{m}\left|a_{Y_i}\right|}{\sum_{i=1}^{m}\left|a_{X_i}\right|}\right)\left(\frac{\sum_{i=1}^{m}\left|b_{Y_i}\right|}{\sum_{i=1}^{m}\left|b_{X_i}\right|}\right) \quad \text{Equation 2-5}$$

where X and Y are the magnitude response of the diagnostic and reference FRA data normally in dB scale, P in Equation 2-4 is the pole of partial transfer function in Equation 2-3, a and b in Equation 2-5 are the coefficients of the rational transfer function in Equation 2-2.

In 2008, P. Karimifard, G.B. Gharehpetian from Amirkabir University of Technology, Iran and S. Tenbohlen from University of Stuttgart, Germany used Vector Fitting method to estimate the transfer function of frequency response, for the purpose of determination of extent of axial displacement [77]. A detailed transformer model was used, which was based on the disc winding (for both HV and LV winding), core geometry and materials to produce FRA data. Three different levels of axial displacement were created. The Sum of Absolute Displacement of Poles (SDP) was introduced to reflect the extent of displacement, which increases when the level of fault increases. However, the SDP values for the 3 degrees of deformation, i.e. displacement of 4%, 8% and 12 % of the winding height, are 141, 145 and 151, which means the alteration of SDP is relatively small compared with the SDP value itself.

In 2011, M. Bigdeli from Islamic Azad University, Iran, M. Vakilian from Sharif University of Technology, Iran, and E. Rahimpour from ABB, Germany used Vector Fitting method to detect and evaluate winding faults types, including axial displacement, radial deformation, disc space variation and short circuit [78]. To study axial displacement, a 31-disc high voltage winding, with 6 turn per disc, and a four layer low voltage winding, with 99 turn per layer, were used, which correspond to a transformer with rated voltage of 10 kV and rated power of 1.3 MVA. To study radial deformation and disc space variation, a 30-disc high voltage winding, with 11 turn per disc, and a one layer low voltage winding, with 23 turns, were used, which correspond to a transformer with rated voltage of 10 kV and rated power of 1.2 MVA. To study short circuit, a 30-disc high voltage winding, with 9 turn per disc was used, with rated voltage of 10 kV and rated power of 1.2 MVA. A new statistic indicator, Faulted-Intact relation (FI) was proposed. FI is less than 1 for axial displacement and radial deformation, and it is larger than 1 for disc space variation and short circuit. Axial displacement and radial deformation can be distinguished by calculating FI from 10 kHz to 100 kHz, whilst disc space variation and short circuit can be distinguished by the different alteration in FI depending on different levels of fault. The

location of disc space variation can be determined by FI, since it has a unique and specific value for a different level and location of disc space variation.

2.3 Modelling

To model transformer windings, it fundamentally requires the compromise between accuracy and complexity, and an appropriate model is needed when analysing a transformer winding at different lifetime stages, such as design, operation or post-failure analysis. Transformer winding models for FRA can be categorised into two groups: white box and grey box models. Each category of model is suitable for certain situations and possesses its own advantages.

2.3.1 White Box Model

In order to understand the deviations induced in the frequency responses and to correlate the changes to the winding components, sensitivity studies are needed. A white box simulation model is preferred, because it is more economic and flexible than experiments on a real transformer for sensitivity study. White box models require geometric information and construction documents of the transformer. They can be classified into distributed parameter model, lumped parameter model and hybrid model. Distributed parameter model is developed based on multi-transmission line theory, and each turn of the winding is viewed as a transmission line, for which the propagation of the voltage signal is described. However, the calculation becomes very complex when the amount of winding turns is huge [79-82]. The lumped parameter model consists of a certain amount of units. Each unit contains the electrical circuit components, i.e. the inductor, capacitor, and resistor, and represents a section of the transformer winding [33, 83, 84]. The lumped parameter model can accurately simulate the frequency response up to 1 MHz, whilst the distributed parameter model is suitable for frequency range up to 10 MHz. In [85], hybrid model has been first time mentioned in the FRA field, extensive discussion has been conducted. However, the definition of hybrid model is not very clear. Normally, by definition, a hybrid model can be regarded as a simplified distributed parameter model; firstly a part of the winding is modelled using multi transmission line theory, i.e., each turn of conductor is treated as a single unit of transmission line using traveling wave equation, and the rest of the winding is treated as a lumped parameter model where a couple of discs would be regarded as a lumped unit. The two parts are connected to each other as a hybrid

model which simplifies the calculation process while the accuracy of frequency response is kept.

In 1992, K. Cornick from UNIST, UK, B. Filliat and C. Kieny from EDF, France, and W. Muller from Siemens, Germany proposed a simulation model for the very fast transient overvoltage analysis. The simulation model was based on turn by turn analysis, using the multiconductor transmission line theory. Experiments were conducted on a three-phase 40 MVA, 220/34.5/6.9 kV transformer, which showed that the simulation result had a good agreement with the experimental result, especially the location of major resonant frequency [79]. Though this model predicts the winding's behaviour accurately, its computation time is too long. In 2007, M. Popovand L. van der Sluis from Delft University of Technology, Netherlands, R.P.P Smeets from Eindhoven University of Technology, Netherlands, J. Lopez-Roldan from Pauwels Trafo Belgium N.V, Belgium and V.V. Terzija from The University of Manchester, UK, presented a refined single-phase 15 kVA, 6.6/0.07 kV transformer model based on transmission line theory, as shown in Figure 2-4, for the purpose to study fast transient oscillations [81]. The elements of the electrical parametric circuit of the winding were calculated per turn, containing the capacitance, the inductance, the resistance as well as the dielectric losses. In this model, the frequency-dependent iron and copper losses were considered. Experimental measurements were carried out, and the validity of the model was verified. The voltage distribution on the winding can be simulated by this model over a large frequency range, up to 10 MHz.

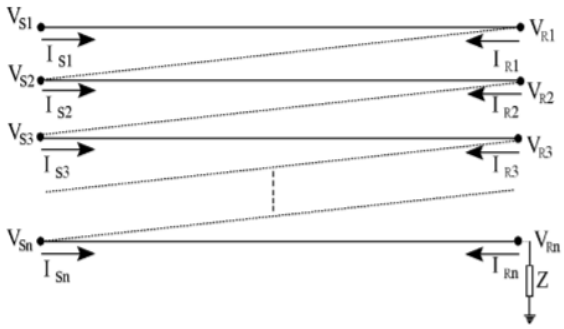


Figure 2-4 Multi-Conductor Transmission Line Transformer Winding Model [81]

The lumped parameter model has the advantages of simplicity and ease of application. The R-L-C-M lumped model is shown in Figure 2-5 [52]. It has been attracting increasingly attentions of researchers due to its ability to precisely represent the internal winding voltage distribution [52, 86]. To build this model, the first step is to calculate the electric components. Two groups of detailed calculation methods were used in [87, 88], including calculations of winding series capacitance, winding ground capacitance, interwinding

capacitance, core inductance, air-core self-inductance and mutual inductance. The skin effect should also be considered in resistance calculation. The second step was to find the system solution. In [69], the system solution was found by MATLAB. The frequency response can be produced automatically after the input of electric parameters calculated in the first step. In [88], the system solution is found by applying Kirchoff's Voltage Law (KVL) and Kirchoff's Current law (KCL) to the circuit using complex inductance and capacitance matrix for calculation. Through white box models, a one-to-one representative relationship can be built between transformer geometrical data and electric parameters.

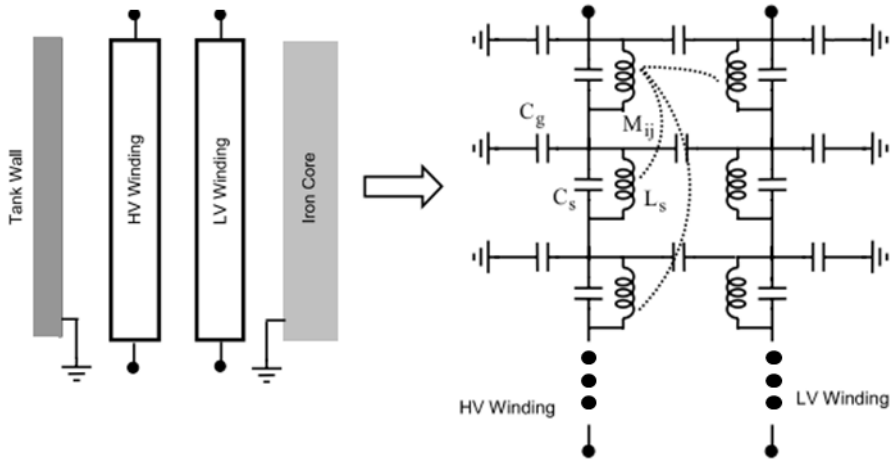


Figure 2-5 Equivalent Circuit of Two-Winding Transformer [52]

In 1996, S.M. Islam and G. Ledwich, from the University of Newcastle, Australia, constructed a single winding equivalent circuit based on winding geometry without core effect [89]. The circuit contained winding series capacitance, ground capacitance, air core self-inductance, mutual inductance between each two cells and winding losses. The model's winding geometry was artificially manipulated to simulate winding disc deformation, interturn and interdisc short circuit faults. The results demonstrate the ability of FRA technique in detecting different winding faults. However, the model can only represent a single winding without core, which is not practical.

In 2003, E. Rahimpour, H. Mohseni, from Tehran University, Iran, and J. Christian, K. Feser from University of Stuttgart, Germany, improved the accuracy of transformer model [33]. The core effect, frequency dependent losses have been considered in the improved model. The model was verified through the comparison between the simulation and measurement results of the model, which has a 30-disc high voltage winding and a single layer low voltage winding. It was concluded that the simulation results were in good agreement with the measurement results in the frequency range from a few kHz up to 1

MHz. The model developed in this work has the limitation that it can only represent single phase transformers. A more complicated modelling method is needed to represent 3-limb and 5-limb cores, as most power transformers operating in the power system are of three phase type.

In 2006, J.A.S.B. Jayasinghe, Z.D. Wang from The University of Manchester, UK, P.N. Jarman from National Grid, UK, and A.W. Darwin from AREVA T&D UK Ltd-Transformers, UK built a simulation model of a 132/11 kV, 30 MVA transformer [52]. In this equivalent circuit, a cell consists of the parallel inductor and capacitor. As shown in Figure 2-5, C_g represents the ground capacitance from winding to core or tank, C_s represents the series capacitance between winding turns and discs, L_s is the combination of leakage and core inductance and M_{ij} represents the mutual inductive couplings between every two cells. There are interwinding capacitances between the HV and LV winding. By applying Kirchoff's Voltage Law (KVL) and Kirchoff's Current law (KCL) to the circuit, the ratio of the voltages at terminal nodes can be computed and then a frequency response can be derived. In 2010, D. M. Sofian, Z.D. Wang, and J. Li from the University of Manchester, UK modelled a three phase 1000 MVA 400/275/13 kV auto-transformer [90]. The series capacitance is calculated by Stein's method [91]. The cross-capacitance (ground capacitance and interwinding capacitance) between units in adjacent windings is derived from the multi-conductor capacitance model [92]. The inductance calculation methods in [93] was used. The same system solution as in [52] was used. It was concluded that in the frequency range from 10 Hz to 1 MHz, the simulation results of series and common winding followed the trends of measurement results well, while the accuracy of tertiary winding modelling should be improved in the high frequency region.

In 2012, Naser Hashemnia, A. Abu-Siada and Mohammad A. S. Masoum, from Curtin University, Western Australia, summarised the influence of different faults on the different frequency regions of frequency responses [94]. Deformation caused by the buckling stress can be simulated by the decrease of the interwinding capacitance and mutual inductance. Besides, the decreased distance between windings leads to the smaller shunt capacitance. The inter disc fault due to axial disc space variation is simulated by increasing the series capacitance and mutual inductance of the two influenced discs. The change in the series capacitance and mutual inductance is used for the simulation of axial displacement. Leakage fault, which is fundamentally caused by the significant increase of dielectric losses, can be simulated by increasing the conductance to ground. One example of leakage

fault is disc to ground fault, which are caused by ground shield damage, insulation damage, high moisture content in the winding, abrasion, hotspot and aging insulation. When a short circuit fault occurs, the series resistance and inductance in the model should be altered. When a radial fault happens, the ground capacitance, capacitance and mutual inductance between high voltage and low voltage windings should be changed accordingly. The whole frequency range was divided into 3 frequency regions, and the influence from the aforementioned faults on different frequency regions are summarised, including the alteration in magnitude and the shifting of resonant points. In 2013, the same authors investigated the influence of different faults on the frequency responses [95]. It was found that the radial displacement has an influence on the whole frequency range, from 10 Hz to 1MHz, while the axial deformation only influences the frequency range higher than 200 kHz.

In short, the white box is suitable for FRA interpretation studies as it can lead to better understanding about how the change of physical dimension of transformer windings changes the features of corresponding frequency responses. However, the white box model requires transformer design data, which may not be accessible in most cases by utilities but original equipment manufacturers.

2.3.2 Grey Box Model

The structure of a grey box model's equivalent electrical circuit has the same circuit topology as the white box model, but its parameters of the circuit are estimated from the given frequency response.

In 2002, J. Pleite, E. Olias, A. Barrado, A. Lazaro and J. Vazquez from Univ. Carlos III, Madrid, Spain proposed a method to construct the grey box model [96, 97]. The magnitude response of the measured frequency response was used to demonstrate the proposed method. In this study, the amount of the basic cells/units of the model was decided by the amount of the resonance and anti-resonance. The whole frequency ranges were also divided into several frequency regions according to the amount of the resonant frequencies. Each cell was assumed to have very limited influence at other frequency regions, and thus only dominated in its own intended frequency range. However, this method does not have good repeatability. When it is applied to the artificial transformer equivalent winding circuits, the calculated values of the circuit elements may deviate from the original data, and sometimes the values may even become negative, which is not physically achievable.

The biggest problem with this method lies on the unrealistic assumption used by the authors.

In 2011, S. D. Mitchell and J. S. Welsh from University of Newcastle, Australia constructed a grey box model for the better interpretation of frequency responses [98]. Nine frequency responses were measured on a Dyn1 1.3 MVA 11 kV/433 V distribution transformer, which included high-voltage end-to-end open circuit test, low-voltage end-to-end open circuit test, and capacitive interwinding test on three phases. A common set of transformer parameters was estimated to simultaneously match all of the nine frequency responses. A nonlinear optimization algorithm was used to determine the parameters of the model by minimising a cost function representing the difference between the estimated and the measured frequency responses. External transformer measurements and knowledge of common manufacturing practices were used to constrain model parameters. The estimated core cross-sectional area, core yoke length, core limb length, and HV winding turns were in good agreement with the actual values, and the estimated frequency response matched the measured frequency response both in magnitude and phase up to 1 MHz. The developed model can be used as a test bed for the sensitivity study of model parameters to support the analysis of frequency response. In 2013, the same authors developed a method to make initial estimations and set constraints for major parameters of grey box model [99]. The proposed method can reduce the possibility that the model parameters converge on local minima without proper constraints, which may not be physically representative. The transformer design documentation is the safe guarded technical know-how of the transformer manufactures. Instead, the external transformer dimensions, routine test data, and nameplate details can be obtained and were used to develop the relationships for the estimation of initial parameter. With the initial parameter estimates and constraints of the winding conductor cross-sectional area, winding conductor length, mean high voltage winding diameter, number of turns in the high-voltage winding, the mean core yoke and limb lengths of a transformer's magnetic circuit, and core cross-sectional area, the representative parameters of the transformer physical dimensions can be achieved. A group of distribution transformers, with different ages, sizes, power ratings and manufacturers were used to demonstrate the applicability of proposed method.

In 2017, R. Aghmasheh and V. Rashtchi from University of Zanjan, Iran, and E. Rahimpour from ABB, Germany put forward some new thoughts for the grey box modelling [100]. The main contribution of this study is the application of appropriate functions to represent the electrical parameters of transformer winding. With the increase

of the distance between winding discs, the magnetic coupling between them decreases. Therefore, there exists an inequality constrains on the inductances of the model. Weibull Distribution function was used for the estimation of both self and mutual inductance. Taking the first row of n×n inductance matrix for example, the independent variable of the Weibull Distribution function is the number of an element, and the value of Weibull Distribution function is the value of the element, i.e. the inductance. Thus the parameter identification process was simplified, by reducing the number of unknown parameters, and eliminating the inequality constraints between inductances. An exponential function was used in the estimation of conductor resistance. The model became more simple and feasible because of the proposed method. The proposed method was tested using a 44-disc with 7 turns in each disc. The frequency response produced by the grey box model was compared with the measure frequency response up to 1000 kHz, and the matching result is acceptable, as the shape of measured FRA trace can be well followed by the simulated FRA trace.

2.3.3 Finite Element Model

The Finite Element Model (FEM) is essentially a white box model. The Finite Element Model (FEM) is based on 3-dimensional electromagnetic simulation to accurately calculate the electrical parameters of equivalent circuit, and an example of the single phase transformer FEM model is shown in Figure 2-6. The whole model was meshed into a large number of small sections, or the so-called finite elements. The response of the FEM model can be obtained by computing the response of each element. In the FEM model, the electrostatic energy, leakage fluxes, electromagnetic forces are considered. The valid frequency range of the Finite Element Model can be up to 1 MHz [101-104].

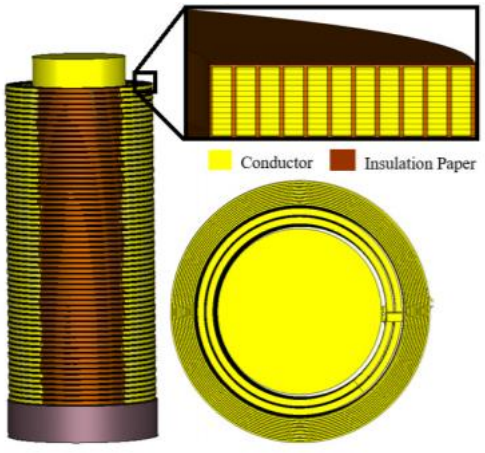


Figure 2-6 3-D Finite Element Model of Single-Phase Transformer [102]

In 2014, Z.W. Zhang from University of Liverpool, UK, W.H. Tang, T.Y. Ji and Q.H. Wu from South China University of Technology, China, investigated the radial deformations through FEM model using COMSOL [101]. A single-phase power transformer was modelled. It has a 30 double-disc high voltage winding, with 11 turns per disc, and a 23-turn, 6 conductors in parallel, helical low voltage winding. In their study, the FEM model was used as a computation tool to calculate the inductance and capacitance. The resistance and conductor were calculated by analytical formulae according to the transformer physical dimension. Those electrical parameters were applied to a hybrid model to produce a frequency response. Accuracy of the simulated frequency response was verified with the experimental result. The radial deformations lead to the alteration in ground capacitance and inductance and thus the change in the frequency response. The authors also suggested that other fault types such as axial movements can be studied using this model.

In 2015, N. Hashemnia, A. Abu-Siada and S. Islam from Curtin University, Australia, built a single-phase transformer FEM model using the software of ANSYS [103, 104]. The inductances and capacitances were extracted using electrostatic and magnetostatic solvers, while the frequency-dependent resistance was obtained by eddy current solver. The frequency response could be produced by applying the extracted electrical parameters to the lumped parameter model. The impacts from axial displacement and radial deformation on the frequency response were studied. In previous study such as [33, 44], the axial displacement was simulated by the alteration in the mutual inductance between high voltage and low voltage windings, while the alteration in capacitance is not considered. In previous study such as [98], the radial deformation was simulated by the alteration in the capacitance of high voltage and low voltage windings, while the alteration in inductance is not considered. Different from the past research, the study proved that a more realistic frequency response can be produced for axial displacement and radial deformation by considering the alteration in both capacitance and inductance. Besides, for different levels of axial displacement and radial deformation, the corresponding percentage changes of electrical parameters were suggested by the authors.

In 2018, S. Tenbohlen, M. Tahir, from University of Stuttgart, Germany, and E. Rahimpour from ABB, Germany, built an FEM model using CST Microwave studio [102]. A single phase transformer, with continuous disc high voltage winding and helical low voltage winding, was modelled, and the power rating of this transformer was about 1 MVA. The model developed in CST MICROWAVE STUDIO can produce the frequency response directly using high frequency solver, without the need to build circuit models.

The simulation was verified with experimental results for both healthy and faulty windings. Two types of Disc Space Variation (DSV) were investigated, i.e. DSV between discs and DSV in full winding. Numerical indices, including Standard Deviation (SD), Correlation Coefficient (1-CC), Euclidean Distance (ED), and Cross Correlation Factor (CCF) were evaluated for their ability to identify DSV. As the degree of mechanical failure increases, all numerical indicators are monotonous. The authors also suggested that the inductive interwinding connection scheme is the most sensitive measurement method to detect DSV.

In the white box model, the winding deformations are normally simulated by the percentage change in the electrical parameters, especially the capacitance. For FEM model, the deformation can be more accurately simulated. Some researchers applied the electrical parameters extracted by the FEM model to the lumped parameter model or the hybrid parameter, while some researcher use FEM model to produce the frequency response directly, for example using the high frequency solver in the design module of CST MW STUDIO. One disadvantage of Finite Element Model is that the computation requires long time and large memory.

2.4 Application of Artificial Intelligence in FRA

Different artificial intelligence methods have been applied for the transformer fault diagnosis, such as Support Vector Machines (SVM), Clustering Analysis (CA), Artificial Neural Networks (ANN), Genetic Algorithm (GA), for multiple purposes, such as parameter estimation for grey box model [105] and diagnosis of the type, level, and location of transformer faults[39, 64, 106-108]. Several artificial intelligence algorithms may be used together to solve some complex problems, or to find a solution with more effective performance. The most commonly used artificial intelligence methods and their applications will be briefly introduced in this section.

The artificial intelligence methods including machine learning can be divided into supervised and unsupervised types. The supervised machine learning methods require observations with known labels as input to build an identification model, and this model could be used to estimate the labels of new observations. The unsupervised machine learning methods classify the input observations according to the similarity or dissimilarity among them, without any guidance, i.e. the labels of observations.

2.4.1 Support Vector Machines (SVM) and Applications

SVM is a supervised machine learning method which was initially proposed for the two-type classification problem, and it is generally considered to be one of the best algorithms for classification problem [109]. Multiclass classification problems can also be solved by SVMs if using multiple binary classifiers. An SVM is a generalised linear classifier, which can be also applied to the nonlinear classification problem in combination with kernel method.

A binary SVM classifier finds an optimal hyperplane which isolates two groups of observations with known types. The distances from the hyperplane to the nearest observations from the two groups should be equal. For example, in the two-dimensional space in Figure 2-7, an optimal hyperplane, represented by the bold line, leans neither to the nearest observation from type C1 (triangles) nor the nearest observation from type C2 (circles). Each observation, either a triangle or circle, is called a feature in SVM algorithm. The boundary features, which are circled in dashed line in Figure 2-7, determine the hyperplane of binary classifiers, and they are the nearest observations to the classification hyperplane from each feature type. Such boundary features are called support vectors in SVM. There may be more than one support vector from each type of features.

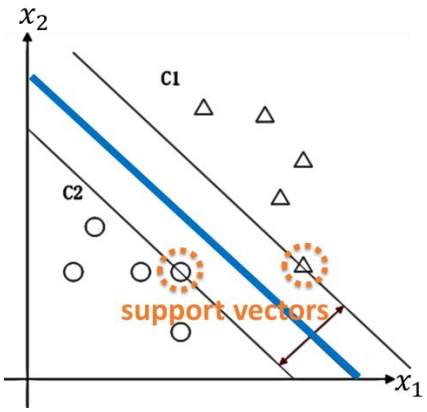


Figure 2-7 Binary SVM Classifier Training Data in 2-D space

In the two-dimensional space in Figure 2-7, the classification hyperplane g(x) =0 is a line and it is defined as:

$$g(\boldsymbol{x}) = \boldsymbol{w}^T \cdot \boldsymbol{x} + b = (w_1 \quad w_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + b = 0 \qquad \text{Equation 2-6}$$

where $\boldsymbol{x}$ is a point, a 2×1 vector, located on the classification line in the two-dimensional space, $\boldsymbol{w}$ is the vector of fitted linear coefficients and b is bias. The vector $\boldsymbol{w}$ and the bias b

determine the slope and the vertical intercept of the classification hyperplane in the two-dimensional space, respectively. The signs of two types of features on two sides of the classification hyperplane are different. In Figure 2-7, the sign of g($\boldsymbol{x}$) =$\boldsymbol{wx}$+b is positive for features from type C1 and negative for features from type C2.

Define y as the classification label:

$$y = \begin{cases} 1, & for\ type\ C1 \\ -1, & for\ type\ C2 \end{cases} \qquad \text{Equation 2-7}$$

For the convenience of computation, the whole plane can be scaled such that for the support vector $\boldsymbol{x}_i$, $g(\boldsymbol{x}) = |\boldsymbol{w}^T \cdot \boldsymbol{x}_i + b|$ equals to 1. To achieve an optimal classifier, the geometrical margin $1/\|\boldsymbol{w}\|$, between the support vector and the classification hyperplane, is to be maximized which is equal to minimizing $\frac{1}{2}\|\boldsymbol{w}\|^2$. Thus, the problem to find the optimal classification hyperplane in the two-dimensional space can be described mathematically as:

$$\begin{cases} min\frac{1}{2}\|\boldsymbol{w}\|^2 & (objective\ function) \\ y_i(\boldsymbol{w}^T \cdot \boldsymbol{x}_i + b) - 1 \geq 0 & (constrain\ function) \end{cases} \qquad \text{Equation 2-8}$$

This is a convex Quadratic Programming problem, which also applies to finding the optimal classification hyperplane in higher dimensional spaces. To solve it, a Lagrange equation can be defined, by combining the constrain function with the objective function with a non-negative Lagrange multiplier $\boldsymbol{\alpha}$:

$$L(\boldsymbol{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\|\boldsymbol{w}\|^2 - \sum_{i=1}^{n} \alpha_i (y_i(\boldsymbol{w}^T \cdot \boldsymbol{x}_i + b) - 1) \qquad \text{Equation 2-9}$$

Set $\theta(\boldsymbol{w})$ as:

$$\theta(\boldsymbol{w}) = \max_{\alpha_i \geq 0} L(\boldsymbol{w}, b, \boldsymbol{\alpha}) \qquad \text{Equation 2-10}$$

Since $\alpha_i \geq 0$, when the constrain function in Equation 2-8 is satisfied, there exists $\theta(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2$. Thus the objective function in Equation 2-8 can be expressed as:

$$\min_{\boldsymbol{w},b} \theta(\boldsymbol{w}) = \min_{\boldsymbol{w},b} \max_{\alpha_i \geq 0} L(\boldsymbol{w}, b, \boldsymbol{\alpha}) \qquad \text{Equation 2-11}$$

Under Karush–Kuhn–Tucker condition (constrain function is satisfied and L is differentiable regarding $\boldsymbol{w}$ and b), the problem in Equation 2-11 can be converted into its dual problem according to Lagrange Duality:

$$\max_{\alpha_i \geq 0} \min_{\boldsymbol{w},b} L(\boldsymbol{w}, b, \boldsymbol{\alpha}) \quad \text{Equation 2-12}$$

To solve Equation 2-12 L should be minimized regarding $\boldsymbol{w}$ and b, by setting their partial derivatives as 0, and hence:

$$\boldsymbol{w} = \sum_{i=1}^{n} \alpha_i y_i \boldsymbol{x}_i \quad \text{Equation 2-13}$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \quad \text{Equation 2-14}$$

Now the problem can be expressed by only the Lagrange multiplier $\boldsymbol{\alpha}$ as:

$$\begin{cases} \max\limits_{\alpha} \sum\limits_{i=1}^{n} \alpha_i - \frac{1}{2} \sum\limits_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \boldsymbol{x}_i{}^T \boldsymbol{x}_j \\ s.t. \alpha_i \geq 0, i = 1, \dots, n \\ \sum\limits_{i=1}^{n} \alpha_i y_i = 0 \end{cases} \quad \text{Equation 2-15}$$

The Lagrange multiplier $\boldsymbol{\alpha}$ can be solved using the sequential minimal optimization algorithm. After $\boldsymbol{\alpha}$ is computed, $\boldsymbol{w}$ and b are derived as:

$$\boldsymbol{w} = \sum_{i=1}^{m} \alpha_i y_i \boldsymbol{x_i} \quad \text{Equation 2-16}$$

$$b = -\frac{\max(\boldsymbol{w^T} \cdot \boldsymbol{x_i} | y_i = -1) + \min(\boldsymbol{w^T} \cdot \boldsymbol{x_i} | y_i = 1)}{2} \quad \text{Equation 2-17}$$

Once $\boldsymbol{w}$ and b are obtained, the classification hyperplane is found. Therefore, the type of new features can be identified according to its sign.

Since the binary SVM classifier can only distinguish two classes of features, multiple binary SVM classifiers are needed for the multiclass classification problem. 'One-versus-one', 'one-versus-all' and 'binary tree' are three commonly adopted multiclass classification strategy [110]. In this study, 'one-versus-one' method is used. This means that between every two classes, a decision is made on which class the new feature is more similar to. Finally, the new feature is identified as the class which wins the most votes.

For nonlinear classification problem, the SVM algorithm needs to convert the input features into a feature space, using kernel functions. The radial basis function, polynomial function and sigmoid function are the most commonly adopted kernel functions. The radial

basis function reflects the distance of training features from the origin, or any predefined point. The polynomial kernel function identifies the resemblance of input samples according to both their given features and the combination of them, i.e. the interaction features. It is suitable for the normalized training data. The sigmoid function is a mathematical function with an ‘s’ shape, and its value ranges between 0 and 1. The sigmoid function is widely used for the prediction of probability.

The SVM model also needs to be trained and tested before it is used for classification as a supervised machine learning method.

In 2012, M. Bigdeli, from Islamic Azad University, Iran, M. Vakilian from Sharif University of Technology, Iran, and E. Rahimpour from ABB, Germany used the SVM algorithm to recognise four types of transformer winding faults, including axial displacement, radial deformation, disc space variation and short circuit of windings [64]. Two groups of test objects were studied; one is a 6.5 MVA 20/6.3 kV transformer while the other group is four model transformers, with a 10 kV disc type high voltage winding, with or without a 0.4 kV low voltage winding, rated power ranging from 1.2 to 1.3 MVA. Vector fitting was used in this study to estimate the parameters of transfer function of measured FRA traces. The mathematical indices of frequency and amplitude variation at the resonant points were used as the input features of SVM classifier. The accuracy of the estimation result was very high, ranging from 80% to 100% for different fault types. The authors compared the results of proposed SVM method with those of the ANN algorithm, and concluded that SVM is more reliable than ANN.

In 2019, J. Liu, Z. Zhao, C. Tang, C. Yao, C. Li from Southwest University, China, and S. Islam from Federation University Australia, Australia used SVM algorithm to identify the winding fault types and levels, including disc space variation, inter-disc short circuit and radial deformation [107]. The frequency responses of faulty winding were produced using a 400 kVA, 10/0.4 kV, specially manufactured 3-phase model transformer. Eight statistic indicators are used as the input features of the SVM model, including Correlation Coefficient, Euclidean Distance, the maximum of difference, the Integral of Absolute difference (IA), the Sum Squared Error, the Sum Squared Ratio Error, the Sum Squared Max-Min Ratio Error and Root Mean Square Error. Particle swarm optimisation algorithm, grid search algorithm and genetic algorithm were used to optimize the parameters of the SVM identification model, and particle swarm optimisation algorithm was regarded as the best among the three optimisation methods in this study, in terms of convergence speed

and quality of the solution. The accuracy of the model reached 96.3% when identifying the fault types. When identifying the fault degrees, the accuracy of disc space variation was 95.24% while the accuracy of radial deformation is only 70%.

2.4.2 Clustering Analysis Methods and Applications

Clustering Analysis (CA) methods are unsupervised machine learning algorithms. The methods group similar observations into clusters. The observations within the same cluster are more similar to each other than to the observations from the other clusters. There are several types of CA methods, including connectivity-based clustering, centroid-based clustering, distribution-based clustering and density-based clustering. The connectivity-based clustering, which is also called hierarchical clustering, is the most commonly used clustering method. Hierarchical clustering includes agglomerative hierarchical clustering and divisive hierarchical clustering. For the agglomerative hierarchical clustering, initially, each observation is treated as a separate cluster. The algorithm repeatedly identifies the closest two clusters and merges them, until all the clusters are merged into one cluster. The divisive hierarchical clustering is opposite to the agglomerative hierarchical clustering. Initially, all observations are treated as a single cluster, and the whole big cluster are split into smaller clusters until the required cluster number is achieved.

In agglomerative hierarchical clustering, a distance matrix and a dendrogram can be used to aid the clustering process. The distance matrix lists all the distances between every two observations, and the closest two observations are merged together, according to the smallest distance in the matrix. Then the distance matrix is updated to find the new closest two clusters. There are many ways to define the similarity, or distance, between the observations, such as Euclidean Distance, Squared Euclidean Distance, Standardised Euclidean Distance, City Block Distance, Minkowski Distance, Chebychev Distance, Cosine Distance, Correlation Distance, Hamming Distance and Spearman Distance. Euclidean Distance, Standardized Euclidean Distance and Correlation Distance are some most commonly used distance types. Euclidean Distance is the root of accumulated squared data deviation, and Standardised Euclidean Distance is the Euclidean Distance of observations after standardisation. Correlation Distance uses a real number between 0 and 1 to reflect the resemblance between two observations.

After two objectives are merged, the distance of newly merged cluster to other clusters should be updated as well. The linkage method refers to the way how the distances

between the newly formed cluster and original cluster are calculated, after two clusters are merged. Different linkages criteria can be used to update the distances, including Single Linkage, Complete Linkage, Average Linkage, Weighted Linkage, Centroid Linkage, Ward Linkage and Median Linkage. The Single Linkage is also called the nearest-neighbour linkage. The smaller is the distance between the nearest two samples from two clusters, the greater is the similarity between the two clusters. The disadvantage of Single Linkage is that it may cause loose clustering. This effect is called Chaining, which means that two cluster can be merged even if they are separated relatively far, due to the existence of two close samples. The Complete Linkage is also called the farthest-neighbour linkage. Opposite to Single Linkage, the longest distance between two clusters' samples is used as the distance of the two clusters. Its disadvantage lies in that even if two clusters are very close to each other, due to the existence of two far away members, they cannot be merged. The Ward Linkage uses inner squared distance, and it is only applicable for the Euclidean Distances.

In 2018, Ali Reza Abbasi and Mohammad Reza Mahmoudi from Fasa University, Iran, and Zakieh Avazzadeh from Nanjing Normal University, China attempted to diagnose the transformer winding faults by the clustering method with Cross Correlation [108]. A detailed transformer model was built to simulate frequency responses under the healthy, radial deformation, axis displacement and short circuits conditions. The whole frequency ranges were divided into low, middle and high frequency regions, and the clustering were carried out on those different frequency regions. The Ward's Linkage and Squared Euclidean Distance were used during the agglomerative hierarchical clustering. A conclusion was made that satisfactory clustering results can be achieved, and the identification of different fault types is feasible.

2.4.3 Other AI Algorithms and Applications

Artificial Neural Networks (ANN) are computation models, and they can be either supervised or unsupervised. [111]. ANN methods are among the most popular artificial intelligence methods in the FRA field, and they have been successfully applied to identify the equivalent circuit of transformers [105], or to identify the type, location and severity of transformer winding faults[106, 112, 113].

ANN methods are nonlinear calculation models, which simulate the way that biological nervous systems analyse and process information. The network is able to evolve itself

based on the input and output, which means that the structures of the ANN models can be influenced by the information processed. An example is given in Figure 2-8 to illustrate the structure of ANN system, where the neural network consists of many basic units, nodes, and the nodes are distributed in three categories of layers, i.e., the input layer, the hidden layer and the output layer. The output layer may have multiple nodes or a single node [113]. The amounts of the important methodological factors including input/output layers, hidden layers, nodes, etc. varying problem from problem.

More hidden layers can be used for more complex problems. The output of each neuron is decided by its input and weight factors:

$$Output\ of\ hidden\ layers\ h_j = \boldsymbol{f}\left(\sum \boldsymbol{W_{ij}} \boldsymbol{x_i}\right) \quad \text{Equation 2-18}$$

$$Output\ layer = \boldsymbol{f}\left(\sum \boldsymbol{W_{jk}} \boldsymbol{h_j}\right) \quad \text{Equation 2-19}$$

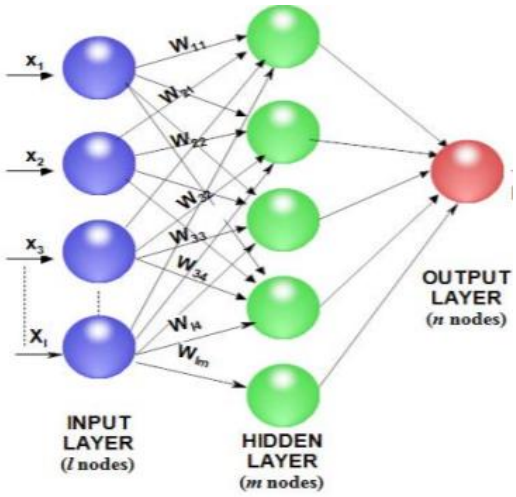


Figure 2-8 Example of Multiple Layer ANN Structure [113]

For the hidden layers, threshold function and sigmoid function are the commonly used functions. The threshold function decides whether the input value is larger than a predefined threshold. The sigmoid function is a mathematical function, which has a 's' shape, and it is often used to introduce nonlinearity in the model. The final classification result is related to the result of the output layer. ANN optimises its weight factors by reducing the difference between its predicted results and the known example results. However, the leaning process of ANN cannot be observed, and its output results are difficult to understand.

The Genetic Algorithms (GA) are artificial intelligence methods for parameters optimisation using search heuristic method. The algorithms were inspired by the natural

selection process, where the individuals with the fittest characteristics are selected to produce offspring. The iterative evolution is adopted in GA, and the candidate solutions in each iteration are called a generation. Initially, a group of random candidate solutions are formed, which are usually expressed in binary values i.e. string of 1s and 0s. The '1' or '0' of the candidate solution is regarded as a gene from a chromosome. The fitness function is a predefined objective function, which is used to describe how well the evaluated candidate solution meets the desired requirement. Then two candidate solutions are chosen as parents to create new candidate solutions, based on their performance according to the fitness function. Crossover and mutation are two importance phases of GA algorithm, where the genomes ('1's and '0's) are alternated. Then the least fitted candidate solutions are replaced with the newly produced candidate solutions, keeping the total amount of candidate solutions unchanged. Such iteration continues until the terminal condition is satisfied, which is normally set as a maximum iteration number or a threshold of the evaluated fitness [114].

Besides the aforementioned artificial intelligence algorithms which are widely applied, other algorithms such as the Particle swarm optimization (PSO) algorithm [115], imperialist competitive algorithm [116], ABC algorithm [117], chaos optimisation algorithm [118], etc. were used in the past research, but theories of those algorithms will not be introduced in detail here.

In 2000, S. Birlasekaran and Xingzhou Yu from Nanyang Technological University, Singapore used the ANN algorithm to identify the transformer fault types using frequency response data [71]. The transfer function was estimated by NLS method, which is a MATLAB inbuilt command, with a fixed order as 32. The frequency locations of the zeros and poles were used as the input of back-propagation ANN algorithm. Altogether 12 faults were investigated, such as unclamped or loosened screws, interturn faults impedance variation at different locations. The accuracy of prediction reached 100%. This research provided the study direction to use FRA trace's mathematical expression, zeros and poles, as the key parameters for transformer fault diagnosis. However, there are several shortcomings of the proposed method. Firstly due to the limitation of the MATLAB command, the transfer function estimation method may fail to match some complex measured frequency response which requires high order. Secondly, the location and severity of faults cannot be recognised. Lastly, the conclusion was based on the fact that all the faults are manually created.

In 2001, A. De and N. Chatterjee, from Jadavpur University, India, proposed a fault identification method based on ANN algorithm [119]. A group of frequency responses were simulated using an EMTP model of a three-phase, 31.5MVA, 132/33kV, y/d transformer with tap changer winding. The produced frequency response patterns were classified using an unsupervised method, self-organising mapping (SOM), and then fine-tuned by the supervised method, learning vector quantisation (LVQ). By this way, the training process did not require to input the fault categories of produced FRA traces. The fault types and locations could be identified. The authors claimed that the developed method is able to sense very delicate variation in the frequency response, which may be impossible for the experts' bare eyes to identify.

In 2006, G. M. V. Zambrano, A. C. Ferreira and L. P. Caloba, from COPPE/UFRJ, Rio de Janeiro, Brazil, estimated the parameters of transformer winding model by ANN method [105]. The frequency response was obtained from simulation studies based on a single-winding lumped parameter model. The mathematical expression, a rational transfer function, of the frequency response was estimated using ANN algorithm. GA algorithm was applied to choose the initial weight factor of each layer of ANN algorithm. Then the equivalent electrical circuit model was produced using conventional circuit theory. The estimated parameters, including values of the inductance, capacitance and resistance, of a four-cell electrical circuit, were very satisfactory, providing an accurate match with the simulated frequency response, in the frequency range from 1 Hz to 1 MHz. However the example used to test the proposed method was oversimplified. The simulated FRA trace was very simple with only a few resonance and antiresonance, while the actual measured data may have dozens of resonance and antiresonance. The proposed method still needs to be tested and verified using more realistic FRA measurement data. Later on, ABC algorithm [117], chaos optimisation algorithm [118], etc. were also applied to construct transformer winding models by other researchers. The estimation results were good. Still, the FRA data used for estimation are relatively simple, containing only a smaller amount of resonant points compared with the actual measured FRA data.

In 2006, R. Vahid from University of Zanjan, Iran and R. Ebrahim from Iran Transformer Research Institute, Iran, used GA algorithm to estimate the parameters of transformer R-L-C-M models [114]. The test object had a disc type high voltage winding, with 60 discs and 9 turns in per disc, and with a rated power of 1.2MVA. The unit number of the simulation model of the test object was also a parameter to be optimised. Not all the data points on the measured FRA trace were used, but the resonant points must be included. In this study,

randomly selected initial points were used for GA algorithm. In the worst estimation case, the number of iteration reached 5000 and it took about 4 hours to complete the iterations. The parameters estimated by the GA algorithm were compared with the calculated parameters by analytical formulas proposed in [33], including the self-inductance, mutual inductance, series capacitance, ground capacitance and resistance. The estimated parameters had a better performance than the calculated parameters in the high frequency range. In 2008, the same authors used the artificial intelligence algorithm, improved particle swarm optimization (PSO), for the parameter identification of transformer model [115]. The same simulation model as that in [114] was used, and every parameter to be estimated remained the same as that in [114]. PSO algorithm is an evolutionary computation technique and it is inspired by the cooperation and competition behaviour between the particles in a swarm. The PSO algorithm was improved by introducing the mutation process which is employed in GA algorithm. A fitness function was used to evaluate the estimation performance. However, the estimated result was not as good as [114], with larger difference between the measured and estimated resonant points.

In 2012, R. Vahid and H. Shahrouzi from University of Zanjan, Iran and R. Ebrahim from ABB, Germany, used an artificial intelligence method called the imperialist competitive algorithm for the model reduction of the detailed R-C-L-M model [116]. The test object was a 400 kV disc winding, with 86 discs and 9 turns in each disc. The imperialist competitive algorithm is the mathematical model of human social evolution. The algorithm was used to estimate the parameters of three reduced models, containing 21, 10 and 5 model units respectively. The reduced model can significantly save the calculation time and simplify the problem.

In 2012, M. Rahmatian, M. S. Naderi, G. B. Gharehpetian and A. J. Ghanizadeh from Amirkabir University of Technology, Iran, used the ANN algorithm to distinguish two typical fault types, i.e., axial displacement with different degrees of severity and radial deformation with different locations [112]. A detailed model was built on based on the geometrical parameter of a 1.2 MVA 10 kV/ 400 V transformer and its variation. The high voltage winding of the transformer has 30 double discs, while the low voltage winding is a double-layer winding. Different fault scenarios were simulated by changing the electrical parameters of the constructed model. The features, which were used as the input of the ANN algorithm, included Cross Correlation, Euclidean Distance, Sum Squared Error and Sum Squared Ratio Error of the diagnosis and reference data. The accuracy reached 75% for radial deformation locations and 81.25% for various degrees of severity of axial

displacement. Though in this study the quantity of fault types was limited, it still suggested a promising direction for the development of a more automated and multifunctional diagnosis method.

In 2014, A. J. Ghanizadeh and G. B. Gharehpetian from Amirkabir University of Technology, Iran proposed a more advanced classification method using ANN algorithm [39] compared with [112]. Their studies have much in common, such as the construction of a detailed transformer model, the employment of ANN algorithm and the features which were selected as input of the ANN algorithm. The transformer model has a 30 double-disc high voltage winding, and a two-layer low voltage winding. However, in addition to the mechanical faults of radial deformation and axial displacement, the study in [39] was able to detect the electrical faults disc-to-disc short circuit faults as well. Short circuit faults are easy to be detected. The accuracy for fault type identification was as high as 98.8%. When the aim changed to identify the fault types as well as their locations, the accuracy dropped to 95.4%. Their study suggests that the radial deformation is the most difficult fault type to be diagnosed.

In 2014, Ketan R. Gandhi from Lalbhai Dalpatbhai College of Engineering, India, and Ketan P. Badgujar from IITRAM, India used 9 numerical indices as the input parameters of ANN algorithm to identify the deviations in frequency responses [113]. Two 10 MVA, 66/11.55 kV delta/star-connected power transformers and one 12.5 MVA 132/11 star/star transformer were used for the case studies. The 9 indices included Correlation Coefficient, Mean Square Error, Sum Squared Ratio Error, Sum Squared Min-Max Ratio Error, Absolute Sum of Logarithmic Error, Absolute Difference, Min-Max Ratio, Comparative Standard Deviation, and Root Mean Square Error. Both the healthy and unhealthy transformer FRA data were used to train a multilayer feed-forward ANN with back propagation algorithm. Minor, moderate, and significant levels of difference can be recognized.

2.5 Summary

In this chapter, the researches regarding the evaluation methods of frequency responses, the modelling of transformer windings and FRA simulation, and the application of artificial intelligence in the FRA filed, are reviewed.

Subjective and objective methods have been applied in the interpretation and diagnosis of frequency response. The subjective methods rely on the expert's experience, while the objective methods use statistic indicators, including numerical indices and transfer function expression. The numerical indices are produced either using all the measured data points or only the resonant points. The transfer function uses several zeros, poles and a constant to mathematically express the frequency response. The fault types, location and level can be identified by the combination of statistic indicators, claimed by some of the publications.

There are three types of transformer models. The white box model is built using the geometric and construction information of the transformers. It contains the distributed parameter model, lumped parameter model and hybrid model. The white box model can be used to produce frequency responses. Meanwhile, the grey box model uses the frequency responses to estimate the parameters of the equivalent electrical circuit of transformer winding. The FEM model uses a more realistic 3D electrostatic and electromagnetic modelling approach to calculate equivalent parameters and produce frequency response, and it is essentially a white box model.

Artificial Intelligence algorithms have been widely applied in the FRA field, including the ANN algorithms, SVM algorithms, Generic Algorithms and so on. Those algorithms have been applied for the estimation of the parameters of winding models and for the identification of fault types, locations and levels of transformer windings.

In this PhD study, to achieve better evaluation of frequency responses, two transfer function estimation methods are developed. One method can precisely reproduce the original FRA trace, while the other method can give a satisfactory matching result with a physical achievable transfer function. By the transfer function method, the amount of the measurement data points can be effectively reduced without losing information.

Though the artificial intelligence methods have been applied in the FRA field for a long time, there is no research conducted on the winding type classification using FRA data. Considering the fact that the utilities have accumulated a large amount of frequency responses with or without transformer design information, such as the winding construction type, there is a need to develop methods to identify the winding types for more accurate interpretation of frequency response and better asset management. Two methods are developed in this PhD study for this purpose.

Chapter 3 FRA Database of UK National Grid Company

3.1 Introduction

The frequency responses investigated in this PhD study are introduced in Chapter 3, which are provided by the UK National Grid Company.

This chapter consists of four sections. The first section introduces how FRA data are stored in UK National Grid's FRA database, how FRA data are extracted from this database, and how to replot the FRA traces. The second section is about the design information of measured transformers, including their quantity distribution in terms of power rating, voltage ratio, and winding type. The information regarding failures of transformers recorded in the database is provided as well. The third section is about the fundamental understanding of the relationship between transformer design and the features of FRA traces. The fourth section introduces the frequency responses used in the following chapters.

3.2 Data Storage and Extraction

3.2.1 Data Storage

The FRA database provided by UK National Grid contains the measured FRA data for the time period from 02/07/1990 to 18/11/2005.

Due to the particular FRA measurement device used, i.e., HP network analyser 4195A, 5 measurements were taken for the frequency regions 5 Hz to 2 kHz, 50 Hz to 20 kHz, 500 Hz to 200 kHz, 5 kHz to 2 MHz, and 25 kHz to 10 MHz, respectively. For each region, 400 evenly distributed frequency points are recorded with sampling interval 0.005 kHz, 0.05 kHz, 0.5 kHz, 5 kHz and 25 kHz respectively. Both the magnitude and phase responses are recorded and they are stored separately in .txt format files. One .txt file only contains either magnitude spectrum or phase spectrum from a frequency region of one measurement.

Therefore, to obtain the full range FRA data for one transformer winding, information from all the related txt should be extracted and then combined.

3.2.2 Data Extraction and FRA Traces Plotting

An excel file is provided by UK National Grid to extract and plot the magnitude spectrum of a specific FRA measurement using Macros.

This excel file owns 7595 rows and 20 columns, and covers around 600 transformers. One row in this excel file corresponds to the FRA measurement record of a transformer winding. The different columns are the basic information of the transformer and the measurement, as well as the location and name of txt files storing the measurement data. An example of one row from the excel file is given below. Figure 3-1 shows the screenshot of the 1494th row. Column A to C is the basic information of the transformer, including its site i.e. substation, its unit number, and the transformer number (TNUMBER). Column D is the manufacturer of the transformer. The same manufacturer tends to use the same design, such as winding construction types. Column E is the measured phase. A phase and C phase have similar magnetic reluctance path whilst B phase has a different magnetic reluctance path due to the symmetric design of the core, thus its FRA trace at the low frequency region is different from the other two phases. Column F describes the measurement connection. Column G is about the tap winding information, i.e. how it is connected, though the 400/275/13 kV auto transformers do not have tap windings, and this voltage ratio group is the focus of this PhD study. When the tap winding's connection is changed, the winding under test will change accordingly, thus the frequency response will change. Column H describes the terminal connection method of other windings, which can also influence the frequency response, for example, delta connected tertiary windings can bring untested windings through the coupling effect. Column I is the test date.

A	B	C	D	E	F	G	H	I
SITE	**UNIT**	**TNUMBER**	**FAMILY**	**PHASE**	**WINDING**	**a to**	**to**	**TESTDATE**
Trawsfynydd	SGT4	T4668	AEI	A phase	N to LV		Other neutrals earthed	15/7/96

(a)

J	K	L	M	N	O	P	Q	R	S
DISK	**2K**	**20K**	**200K**	**2M**	**10M**	**DIRECTORY**	Serial No.	**Row No.**	Design group
B44	16	17	18	19	20	C:\Work\Databases\Databases\txc\fra\FRA14		1478	A01

(b)

Figure 3-1 Excel File Record for T4668 A Phase, N to LV

Columns J to O indicate the name of .txt file in which FRA data are stored. A text file's name consists of three parts. The first part is the element stored in column J 'DISK'. The second part is the element contained in one column from the columns K to O, which specify the frequency regions 5 Hz to 2 kHz, 50 Hz to 20 kHz, 500 Hz to 200 kHz, 5 kHz

to 2 MHz, and 25 kHz to 10 MHz. The last part is ‘C1’ or ‘C2’. ‘C1’ leads to the magnitude spectrum while ‘C2’ leads to the phase spectrum. By combining the three parts, the corresponding filename can be obtained. Column P suggests the location of folder storing the txt files of each measurement. For example, it is revealed by 1494th row that in the folder ‘C:\Work\Databases\Databases\txc\fra\FRA14’, a txt file called ‘B4416C1’ stores the FRA magnitude spectrum from 5 Hz to 2 kHz for the measurement on A phase, N to LV winding of transformer ‘T4668’. The frequency points, on which FRA data are sampled, are not recorded since they are in the same format for every FRA measurement. And those frequency points can be produced easily according to the frequency range and sampling intervals mentioned above.

However, this excel file can neither extract the phase data, nor combine the magnitude spectra from different frequency regions.

A MATLAB programme was written to extract and replot the frequency responses, requiring the information of transformer name, tested phase, tested winding, FRA measurement connection method etc. as input.

Different from the excel programme, the MATLAB programme used the combined data to plot a single FRA trace for the desired whole frequency range. For example, when extracting the frequency responses from 5 Hz to 200 kHz, to eliminate these duplicate measurement points at lower frequencies, 40 redundancy data were eliminated from the 2nd measurement, and 40 redundancy data from the 3rd measurement, resulting in 1120 data points per FRA trace. For the 1494th row of FRA measurement, the combined total of 1120 data points produced the magnitude and phase spectra against frequency as shown in Figure 3-2.

Due to the redundancy between the magnitude response and the phase response, researches have been focused on the magnitude response. Therefore, the phase plots will not be shown in the following part of this chapter.

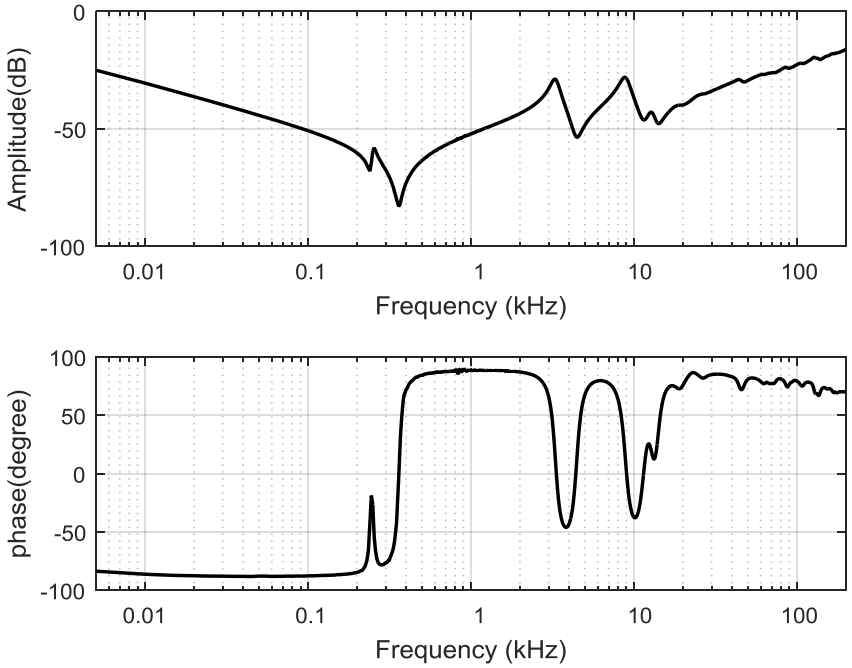


Figure 3-2 5 Hz to 200 kHz FRA Plot for T4668 A phase, N to LV

3.3 Analysis of Transformers Basic Information

UK National Grid FRA database contains 5963 frequency responses with transformer identity, from 474 known transformers, and 1631 frequency responses without transformer identity. By counting sites and unit numbers, it can be found that the 1631 frequency responses are from 134 other transformers.

Among the 474 transformers with TNUMBER recorded in the FRA database, the voltage ratios and power ratings of 269 transformers out of 474 can be acquired, and 149 out of 269 transformers have known winding types, with the available information sources.

3.3.1 Voltage Ratio and Power Rating

For the 269 transformers with known voltage ratios and power ratings, their voltage distribution is shown in the following pie chart in Figure 3-3. The 275/132 kV transformers have the largest quantity, 90. The numbers of 400/132 kV transformers and 400/275/13 kV transformers are 68 and 44 respectively. All of the above mentioned transformers are autotransformers, which add up to the amount 202.

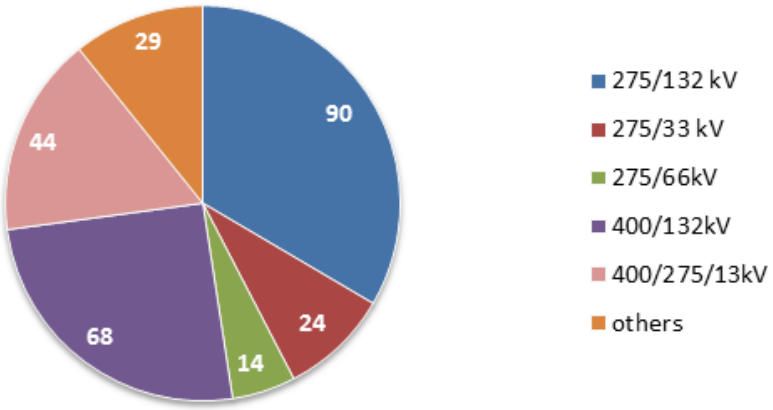


Figure 3-3 Pie Chart of Voltage Distribution of Transformers with known Tnumber

For an autotransformer, when measuring the frequency response, their neutrals are separate. Normally the FRA test is a single phase end-to-end open circuit test, with the other neutrals earthed, but in some cases they may be floating. The tertiary winding terminals are delta-connected, and they are normally connected internally with the tank and brought out by small bushings.

Out of the 269 transformers, there are 128 transformers of which the high voltage side is 275 kV. The quantity distribution of the 128 transformers in terms of different levels of voltage ratio and power rating is shown in Table 3-1. The number in bracelet indicates the corresponding quantity of each power rating group. There are mainly 3 categories, 275/132 kV, 275/33 kV and 275/66kV, with or without tertiary winding. It can be seen that 275/132 kV transformers take up the majority, and more than half of them own a 13 kV tertiary winding. Their power ratings range from 75 MVA to 240 MVA. The highest power rating reaches 240 MVA, which is also the most populated power rating of this voltage level.

Vector group is a method used by IEC to classify the winding configurations of three phase transformers. To describe HV winding connection, capital letters are used. For description of LV and tertiary winding connection, small letters are used. ‘D/d’ and ‘Y/y’ mean that the winding is delta and star connected respectively. ‘Z/z’ means interconnected star. ‘N/n’ means that the neutral is brought out. The number followed by the letters representing LV and tertiary winding shows difference in the phase angle between them with the HV winding. The hour indicator is used, where ‘1’ represents 30°, ‘2’ represents 60°, and ‘12’ represents 360° or 0°. The digital ‘11’ means that the LV winding lags the HV winding by 330°, or that the LV winding leads the HV winding by 30°. The winding configuration impacts the interaction between windings, and thus has an influence on the frequency response.

For the 275/132 kV autotransformers, the vector group YNa0d11 is used. This means that the HV winding is star connected with natural brought out. ‘a0’ means that it is an auto transformer and LV winding, i.e. common winding, is in phase with the series winding, i.e. part of the HV winding, with 0° displacement. TV winding is delta connected and leads the HV winding and LV winding by 30°. For 275/33 kV transformers, the winding connection scheme is YNd1 or YNd11. For 275/66 kV transformers, the winding connection scheme is YNd1, YNd11, or for 275/66/13 kV YNynd11. YNynd11 refers to the winding configuration that both the HV winding and LV winding are star connected, HV and LV windings are two physically separate windings, and the tertiary is delta connected, leading the HV winding 30°.

Table 3-1 Transformer with 275 kV High Voltage Side

Voltage Ratio	Number		Number	Power Rating		
275/132 kV	90	275/132/11 kV	31	120MVA (27)	180MVA (3)	240MVA (1)
		275/132/13 kV	56		180MVA (14)	240MVA (42)
		275/132kV (no tertiary)	3			240MVA (3)
275/33 kV	24			75MVA (2)	100MVA (15)	120MVA (7)
275/66kV	14	275/66/13kV (star/star)	5	120MVA (1)	180MVA (4)	
		275/66/33kV (star/star)	2	120MVA (2)		
		275/66kV (Star/delta)	7	120MVA (3)	150MVA (1)	180MVA (3)

There are 112 out of 269 transformers whose high voltage side is 400 kV, as shown in Table 3-2. 68 of them are 400/132kV transformers and 44 of them are 400/275/13kV transformers. For both 400/275/13 kV and 400/132/13 kV autotransformers, the vector group is YNa0d11.

Table 3-2 Transformers with 400 kV High Voltage Side

Voltage Ratio	Number		Number	Power Rating			
400/132 kV	68	400/132/13kV	55	220MVA (1)	240MVA (52)	276MVA (1)	288MVA (1)
		400/132kV no tertiary	13			240MVA (13)	
400/275/ 13 kV	44			500MVA (5)	750MVA (22)	900MVA (1)	950MVA (1)
				1000MVA (14)	1100MVA (1)		

3.3.2 Winding Construction Types

The winding construction types of all known transformers are summarised in this section. This design information can be acquired in different ways. The winding construction type information is most reliable if it is from the transformer manufacturer. The winding type can also be acquired by disassembling retired or failed transformers. Besides, the transformers from the same design group would have the same design. Therefore, it is also feasible to obtain the winding construction information of the transformer if the winding type of another transformer from the same design group is known.

Table 3-3 shows the winding type distribution of 275/132kV transformers. In each row, the transformers have exactly the same winding types in terms of common, series, tap and tertiary windings. It can be seen that there are only a few common windings and series winding are layer type for 275/132kV transformers. The Plain Disc winding takes up the majority of disc windings, followed by Interleaved Disc winding and Intershielded Disc winding. For the tap winding and tertiary winding, Single Helical and Layer windings are used. Because of the existence of high-frequency components in the spectrum of a voltage surge and lightning impulse, the voltage distribution along the winding is non-linear, which leads to larger local stress on the beginning of the winding. Higher Basic Insulation Level (BIL) is needed for transformers with higher voltage rating. Plain Disc winding are often used for lower voltage rating up to 132kV windings. To overcome the non-linear distribution of impulse voltage, Intershielded Disc winding and interleaved winding, which own higher series capacitance, are often used.

Table 3-3 275/132kV Transformer Winding Types

	Group	Tx Number	Series winding	Common winding	Tap winding	Tertiary winding
275/132kV	**A**	3	NA	NA	NA	-
275/132/11 kV (31)	**A**	7	Plain Disc	Plain Disc	Layer	Layer
	B	1	Layer	Layer	NA	NA
	C	6	Disc	Disc	Layer	NA
	D	8	Interleaved Disc	Interleaved Disc	Helical	NA
	E	9	NA	NA	-	NA
275/132/13 kV (56)	**A**	15	Plain Disc	Plain Disc	NA	NA
	B	8	Plain Disc	Plain Disc	NA	Helical
	C	1	Plain Disc	Plain Disc	Helical	Helical
	D	1	Interleaved Disc	Intershielded Disc	NA	NA
	E	1	Disc	Disc	NA	NA
	F	1	Five-layer	Disc	Layer	NA
	G	1	Layer	Layer	Layer	
	H	28	NA	NA	-	NA

'NA' is abbreviation for 'Not Available.'

It can be seen in Table 3-4 that 275/33 kV transformers are divided into four types according to the construction of windings. Due to the low voltage windings is only 33 kV, helical winding can be applied for this low voltage level in group B. For 275/66 kV transformers with or without tertiary windings, the layer winding type and disc winding type take up 50% of the total amount separately.

Table 3-4 275/33kV & 275/66kV Transformer Winding Types

	Group	Tx Number	HV winding	LV winding	Tap winding	Tertiary winding
275/33kV (24)	**A**	8	Interleaved Disc	Plain Disc	Interleaved Disc	-
	B	1	Plain Disc	Helical	NA	NA

	C	2	Double Concentric coils (no more information)	NA	NA	NA
	D	13	NA	NA	NA	NA
275/66 kV (7)	**A**	1	Plain Disc	Plain Disc	Double Concentric	NA
	B	1	Plain Disc	Plain Disc	NA	NA
	C	5	NA	NA	NA	NA
275/66/13 kV (5)	**A**	3	Layer	Layer	Double Concentric	NA
	B	1	Plain Disc	Plain Disc	NA	Helical
	C	1	NA	NA	NA	NA
275/66/33 kV	**A**	2	NA	NA	NA	NA

'NA' is abbreviation for 'Not Available.'

For 400/132 kV transformers in Table 3-5, 55 out of 68 400/132kV transformers have a 13 kV tertiary winding. For the 13 transformers without tertiary winding, all of the common and series windings are of disc type. For the common and series windings from the 55 transformers with a 13 kV tertiary winding, only a small portion of them are layer type.

Table 3-5 400/132kV Transformer Winding Types

	Group	**Tx Number**	**Series winding**	**Common winding**	**Tap winding**	**Tertiary winding**
400/132kV (13)	**A**	8	Plain Disc	Plain Disc	Two-layer	-
	B	2	Interleaved Disc	Plain Disc	NA	-
	C	3	NA	NA	NA	-
400/132/13kV (55)	**A**	5	Plain Disc	Plain Disc	Two-layer	NA
	B	11	Plain Disc	Plain Disc	NA	NA
	C	5	Interleaved Disc	Plain Disc	Helical	NA
	D	4	Interleaved Disc	Plain Disc	NA	Helical
	E	6	Layer	Plain Disc	Layer	NA
	F	2	Layer	Layer	NA	Layer
	G	22	NA	NA	NA	NA

'NA' is abbreviation for 'Not Available.'

In Table 3-6 44 400/275/13kV transformers are summarised. Same as 400/132 kV transformers, the disc type winding has a larger amount than the layer winding type. Plain Disc winding is the most widely used winding construction type, and then follows the Interleaved Disc winding. The Multiple Layer winding type has a larger proportion than the Intershielded Disc winding.

Table 3-6 400/275/13kV Transformer Winding Types

	Group	Tx Number	Series winding	Common winding	Tertiary winding
400/275/13 kV (44)	**A**	4	Plain Disc	Plain Disc	Single Helical
	B	9	Plain Disc	Plain Disc	NA
	C	6	Interleaved Disc	Intershielded Disc	Single Helical
	D	1	Interleaved Disc	Interleaved Disc	Single Helical
	E	4	Interleaved Disc	Interleaved Disc	NA
	F	2	Multiple Layer	Multiple Layer	Single Layer
	G	3	Multiple Layer	Multiple Layer	NA
	H	1	Disc	Disc	NA
	I	14	NA	NA	NA

'NA' is abbreviation for 'Not Available.'

The focus of this PhD study is on the 400/275/13kV transformers. This is because that the 400/275/13kV transformers are auto-transformer and they have no tap windings. This is regarded as comparatively simpler winding structure. Besides, the winding type information of this voltage ratio's transformers is the most complete as compared with transformers from other voltage ratios.

3.3.3 Record of Transformer Failures

According to the record of UK Nation Grid Company, the statistics of transformer failures from 1962 to 2002 are listed in the first column of Table 3-7. For the 20 failed 400/275/13 kV transformers, there are 7 transformers whose frequency responses are recorded in the FRA database, as shown in second column of Table 3-7. Considering the fact that the time-based comparison may not be available for all transformers, the construction-based comparison is also used when analysing the frequency responses of faulty transformers. Nevertheless, data are often not available due to incomplete records. Hence the numbers of

failed transformers with obvious alteration in frequency responses are listed in third column of Table 3-7 for information.

Table 3-7 Statistic of Faulty Transformers from Year 1962 to 2002

	No. failed transformer	No. failed transformer with FRA record	No. failed transformer with obviously altered frequency response
400/275/13 kV	20	7	1
400/132 kV	13	5	1
275/132 kV	19	6	1
275/66 kV	3	0	0
275/33 kV	7	2	2
400/66 kV	2	0	0
400/7.9 kV	1	1	0

Table 3-8 gives the information of the seven 400/275/13 kV transformers with FRA records, including rated power, manufacturer, winding construction type, manufacture year and failure year, and a brief note of fault diagnosis. T3983, T4296, T4297 and T4281 are rated at 500 MVA, T6606 and T5611 are 750MVA, and T6463 is 1000 MVA. T3983, T4296, T4297 and T4281 are manufactured by Bruce Peebles with the same design in the years 1965-1967. It is believed that they are the manufacturer's first design for the voltage level of 400 kV. All of those transformers were rewound, repaired or modified to a slightly different design in the years 1967-1970. The transformers have Plain Disc series and common windings and a helical tertiary winding. For this design, it is believed by the experts that short-circuit withstand capability might not be adequate, which partially contributes into loose clamping and leads to arcing/sparking at the loose clamping bolts. T4281, T4296 and T4297 all had dielectric failures.

As stated, the incomplete frequency response measurements and different measurement connection schemes lead to the limited amount of frequency responses available for comparison between reference and diagnosis measurements. As shown in Table 3-7, under the same measurement connection scheme, using either time-based or construction-based comparison, only one 400/275/13 kV transformer, i.e. T6463, has the reference and diagnosis frequency responses with obvious alteration. It failed in 1999 because of a short circuit current passing through the transformer, and a post-failure examination was also

carried out. This transformer was scrapped, and visual inspection indicated significant common winding movement, as shown in Figure 3-4.

Table 3-8 Information of 400/275/13 kV Failed Transformers

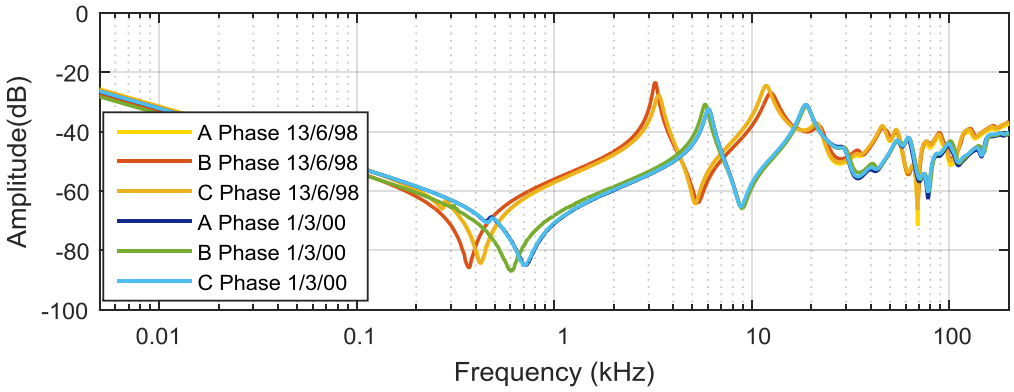
Tnumber	Rated Power (MVA)	Manufacturer	Winding construction type	Made year	Failed year	Fault
T3983	500	Bruce Peebles	Plain Disc series and common windings, helical TV winding	1965	1966	possibly winding fault
T4296				1966	1973&1996	inter-phase insulation; lightning strike, major insulation
T4297				1967	2000	arcing/sparking fault
T4281				1996	1976&1993	possibly gassing fault; arcing/sparking fault
T6606	750	Hawker Siddeley Power Transformers	Not Available	1978	1991	Not Available
T5611	750	Hackbridge and Hewittic	Plain Disc series and common windings	1969	1993	inter-phase insulation
T6463	1000	Parsons Peebles	Not Available	1976	1999	short circuit

(a) Photo 1

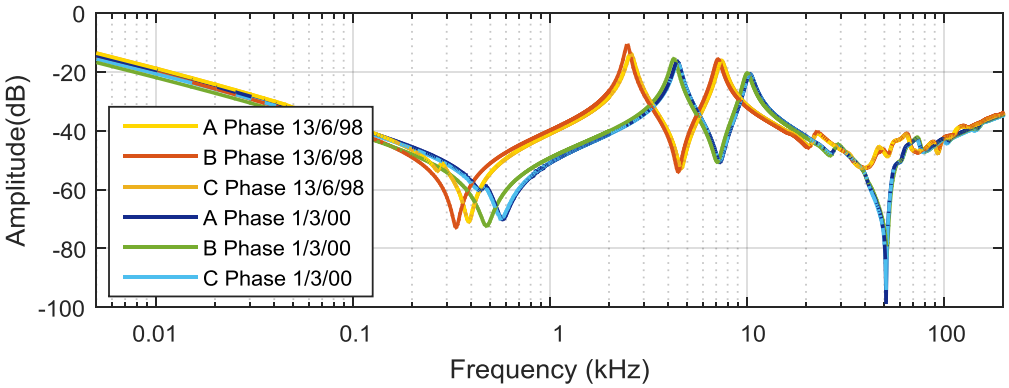
(b) Photo 2

Figure 3-4 Observed Damage to the Shield Ring on the Common Winding Top End

The reference frequency response measurement was made on 13/06/1998, and the diagnosis one was on 1/3/2000. The connection method used is end-to-end open-circuit test, with the other neutrals earthed and the delta-connected tertiary windings' terminals floating. The measured frequency responses of A, B, and C phase on N to LV and HV to LV are plotted in Figure 3-5. It can be see that nearly all the features, resonances and anti-resonances, of the diagnosis frequency responses have shifted to the right when comparing with the references. For the HV to LV measurements, there is a significant change around 50 kHz where the diagnosis FRA magnitude of an antiresonance dropped to around -100 dB.



(a) N to LV



(b) HV to LV

Figure 3-5 Frequency Responses of Transformer T6463

3.4 Fundamental Understanding on Influence of Transformer Design

The transformer's winding design affects its frequency response. In this section, examples will be given to investigate the influence from voltage ratio and power rating, using the examples from 400/275/13 kV autotransformers group.

Figure 3-6 shows the FRA traces from A, B, and C phases of common winding of 400/275 kV autotransformers T5291. Empirically, for 400/275/13 kV autotransformers, the frequency regions 5Hz -2 kHz, 2 kHz -20 kHz, 20 kHz -1000 kHz are dominated by transformer core, interaction between windings, winding structure respectively, and the region higher than 1 MHz is governed by the measurement setup. The frequency boundaries may change due to the variation in the voltage, power rating, winding types etc. On the frequency region 5Hz -2 kHz, the FRA traces from A phase and C phase own two troughs and one peak, while the FRA trace from B phase own only one trough. This can be explained by the difference in their magnetic reluctance path, which is caused by the transformer core structure. B phase has two symmetric reluctance paths, while A phase and C phases both have two different reluctance paths. On the frequency region 2 kHz to 20

kHz, the double peak characteristic can be seen. The first peak is dominated by the tertiary connection, and the second peak is controlled by the interaction of common and series windings from the same phase [120].

According to [120], the frequency region 20 kHz to 500 kHz is controlled by the winding types, mainly the winding series capacitance in proportion to the shunt capacitance. Around the frequency 1000 kHz, the interaction between inductance of the earthing lead and the capacitance of the transformer causes an evident peak. In this PhD study, for 400/275/13 kV transformers, the frequency region 20 kHz to 1000 kHz is regarded to be dominated by the properties of winding-under-test, after observing a large amount of frequency responses in the National Grid FRA database.

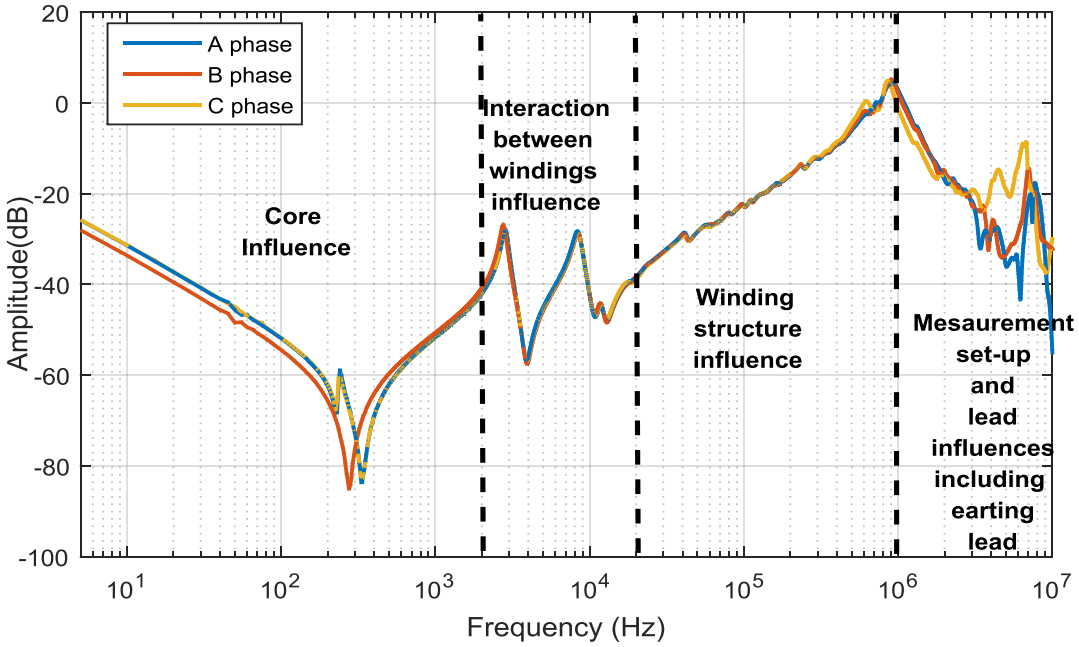


Figure 3-6 Influence Factors of FRA traces from Auto-Transformers without Tap Winding

3.4.1 Voltage Influence

Comparison of the measured FRA traces is done on the common, series, and tertiary windings from the same transformer, T5291, which own different voltages, as shown in Figure 3-7.

In low frequency region, i.e. 5 Hz to 2 kHz, the transformer core is the dominant equivalent component of the equivalent circuit of the windings under-test. The starting magnitude of the common and series winding FRA trace is dominated by the core inductance L_c, which can be calculated as:

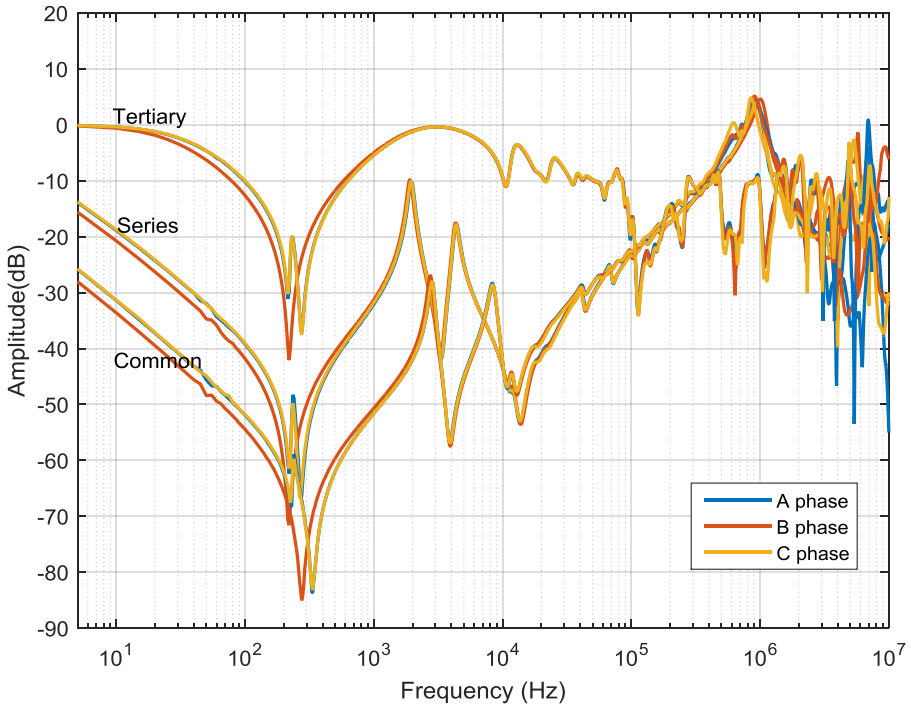


Figure 3-7 Comparison of Frequency Responses from Different Windings of the Same Transformers

$$L_c = \frac{N^2 u_0 u_r A}{l_c} \qquad \text{Equation 3-1}$$

where N is the number of winding turns, u_0 and u_r are the permeability of vacuum and core separately, A is the cross-sectional area of the core, and l_c is the mean length of the leg. The basic SI units for the recorded values are used in the calculation.

The common windings, 275 kV, have higher voltage than the series windings, 125 kV, thus the common windings have more turns than the series windings. As a result, when other influence factors are the same, more turns leads to a higher core magnetising inductance, a higher reactance, and thus a lower starting magnitude. This leads to the smaller starting magnitude of the common windings in the core controlled frequency region. The tertiary have the smallest voltage 13 kV, thus it has the highest starting magnitude. Note that the equivalent core impedance of the TV winding is too small to the measuring impedance of 50 Ω, the starting magnitude at 5 Hz is close to 0 dB.

The starting magnitude of B phase is smaller than that of A phase and C phase. This could be explained by the core structure. B phase has lower magnetic reluctance paths compared with A phase and C phase, which results in higher core magnetizing inductance and thus lower starting magnitude.

The common winding and series winding are of Multiple Layer winding type. For the frequency region controlled by winding structure, i.e., winding series capacitance in proportion to the shunt capacitance, from 20 kHz o 1000 kHz, the common winding and series winding own similar FRA magnitude traces, which rise roughly at 20 dB per decade with oscillations. Though their voltages are different, their magnitudes on the winding controlled frequency range are similar. The FRA traces of tertiary winding own the highest magnitude from 0 to 100 kHz. In the frequency region larger than 1000 kHz, although for the common and series windings, it is controlled by the measurement set up, for the FRA traces of tertiary winding, its own winding characteristics is shown, i.e. the most distinct resonances caused by tertiary winding construction, i.e. the smallest series capacitance.

3.4.2 Power Rating Influence

The influence of power rating is investigated in this section, through three 400/275/13 kV transformers from the same manufacturer. The vector groups of the three transformers are the same. The frequency responses of both series windings and common windings, on B phase, from transformer T6987 of 1000 MVA power rating, 16% impedance, transformer T7008 of 750 MVA power rating, 20% impedance, and transformer T6933 of 500 MVA, 12% impedance, are plotted in Figure 3-8. As stated, when measuring the frequency responses, the other neutrals are earthed and the tertiary windings are delta-connected and floating. Both the series winding and common windings of the transformers are of Plain Disc winding type. The frequency region 5 Hz to 200 kHz are shown in Figure 3-8 in a linear frequency plot to better illustrate the characteristics of the frequency response.

The core sizes of the transformers with difference power rating are expected to be different. A shorter and fatter core design leads to a higher core inductance, which influence the starting magnitude of frequency response. The permeability of the core and residue of magnetization also influence the starting magnitude. Thus, more information is needed for a more meaningful and convincing discussion for the low frequency region which would be just ignored here.

Equation 3-2 and Equation 3-3 show how to calculate the leakage inductance of winding at 50 Hz frequency, neglecting the influence of the resistance, using voltage bases of 275 kV and 125 kV, and power rating bases 500 MVA, 750 MVA, and 1000 MVA:

$$Z = Z_{pu} \times Z_{base} = Z_{pu} \times \frac{V_{base}2}{S_{base}}$$ Equation 3-2

$$L = \frac{|Z|}{2\pi f}$$ Equation 3-3

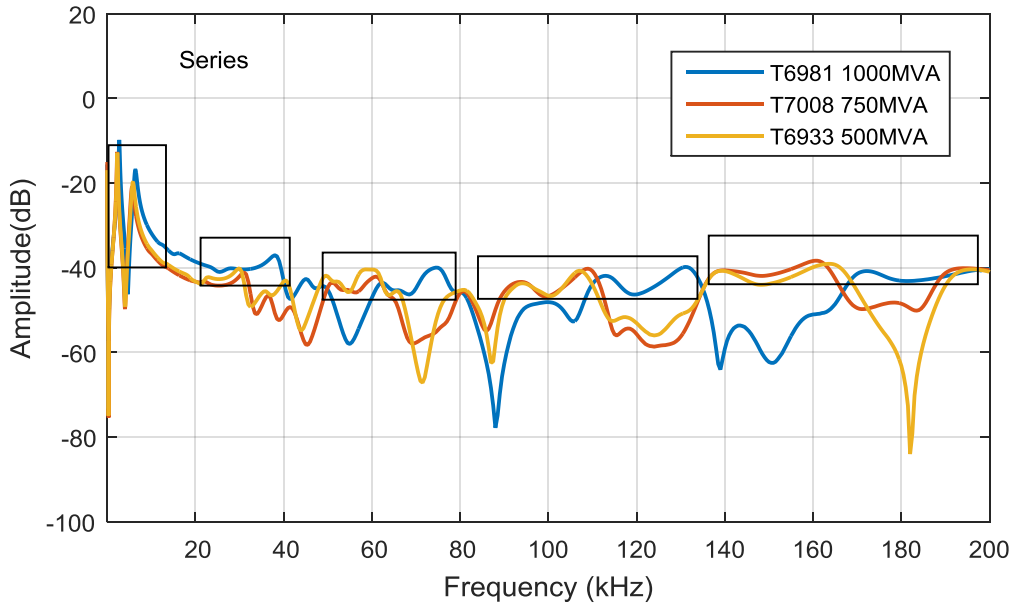
Table 3-9 Information of Transformers with Different Power Rating

TNUMBER	T6981	T7008	T6933
Power Rating(MVA)	1000	750	500
Design Group	S01	S02	S03
Impedance (%)	16	20	12
Leakage inductance of Series winding(mH)	7.958	13.262	11.937
Leakage inductance of Common winding(mH)	38.515	64.192	57.553

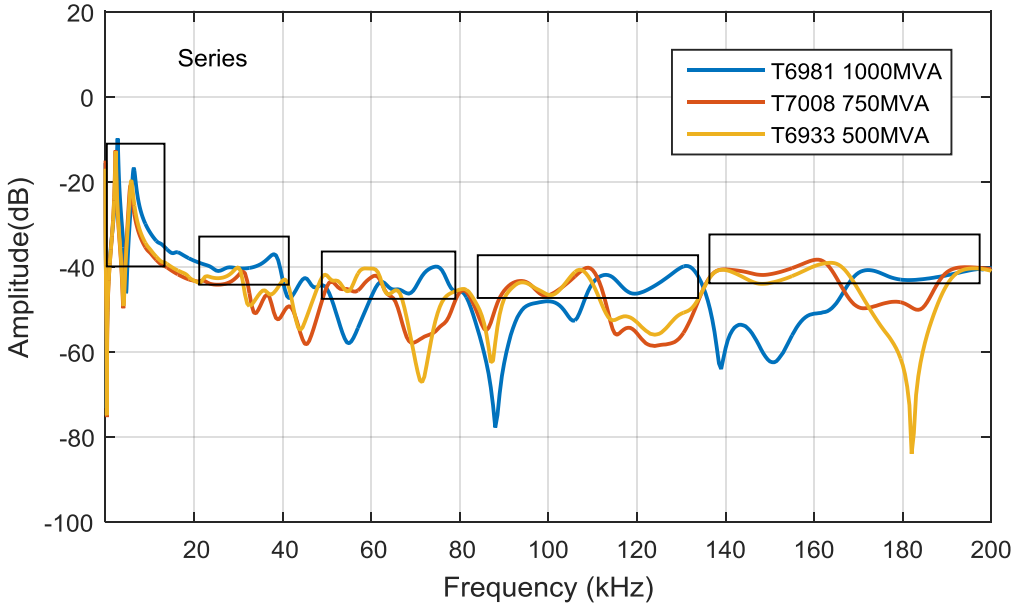
According to the equations, a smaller power rating indicates a higher winding leakage inductance. However, there are also other impact factors including the voltage and the impedance, thus in Table 3-2, the inductance of the 500 MVA transformer is smaller than that of the 750 MVA transformer. The leakage inductance is a lumped global element, hence it would be reflected more in the low and middle frequency region <20 kHz, as seen in Figure 3-7. In general, a higher inductance makes the resonance shifting to the left (i.e. shift to the lower frequency end), and this is particularly true for the 2rd peak of the 'double-peaks' in the middle frequency region.

The capacitance of the winding is influenced by the winding construction types, the physical dimensions of windings, the turn numbers, the coil cross section etc. Though for the three transformers, their winding types are the same, there is limited information on other influence factors. In general, smaller capacitance makes the resonance shifting to the right (i.e. shift to the higher frequency end).

In Figure 3-8, it can be seen that for both series and common windings, along with the increase of the power rating, the resonance frequency locations shift to higher frequency locations, which are individually marked in the rectangles. The right shifting patterns in the high frequency region may be explained by the combined influence of the winding capacitance and inductance. In a way, although the values of inductance and capacitance may be different, the inductance and capacitance matrixes may have a similar pattern for the three transformers.



(a) Series Winding



(b) Common Winding

Figure 3-8 Comparison of Different Power Ratings with Same Voltage

3.4.3 Winding Construction Types Influence

The 'typical' FRA traces of four winding types from 400/275/13 kV autotransformers are plotted Figure 3-9, i.e., Multiple Layer, Plain Disc, Interleaved Disc and Single Helical windings. These are the measured FRA data on a winding with its winding type known, selected based on empirical experiences after observing thirty three-phase 400/275/13 kV transmission transformers with known winding types, including series, common and tertiary windings. In the database, the same windings may have several measurements which were taken at different time, or with different connection arrangements of winding terminals.

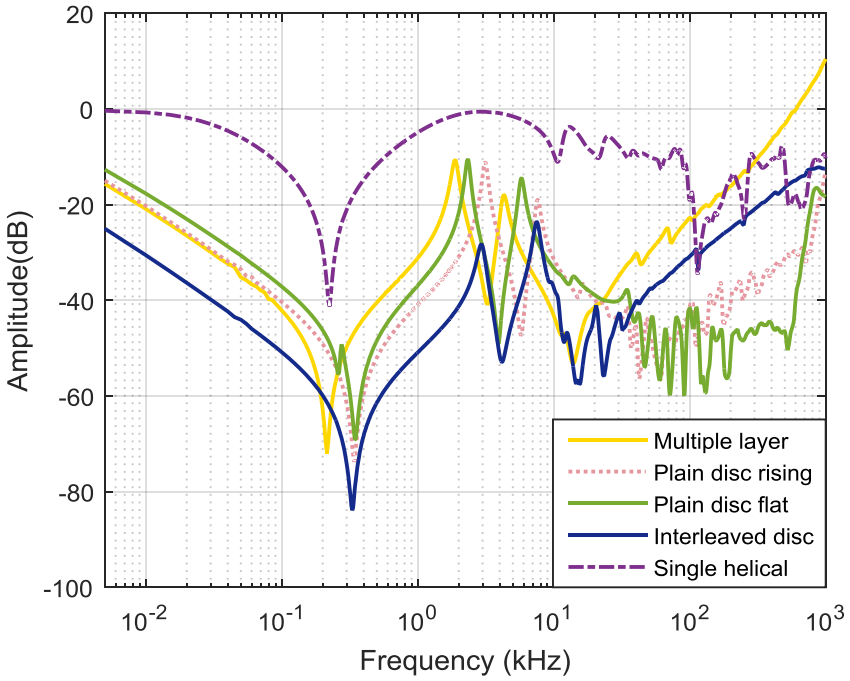


Figure 3-9 Typical FRA Traces of Different Winding Types

The Single Helical winding, used for the tertiary windings, has the highest magnitude roughly from 5 Hz to 100 kHz. This is because the tertiary winding has the lowest voltage and power rating (13 kV, 60 MVA), hence the overall winding inductance and winding series capacitance are the smallest. The Plain Disc winding has a lower winding series capacitance, and its FRA trace has the oscillatory camel humps characteristics which can be observed between about 20 kHz and 500 kHz. Either an overall rising or flat trend may appear in this frequency region for a Plain Disc winding. Multiple Layer and Interleaved Disc windings have high winding series capacitance, and their FRA traces have a rising trend from 20 kHz to 1 MHz. The FRA trace of Interleaved Disc winding has a smoother rising trend, while that of Multiple Layer winding has some obvious fluctuations. The reason behind this observation could be explained, as the Interleaved Disc winding has a higher winding series capacitance in proportion to the winding shunt capacitance, as compared to the Multiple Layer winding, and this higher ratio eliminates the appearance of resonant frequencies in this region. Based on the experience of the limited amount of Intershielded Disc windings which were scrapped, it was concluded that the FRA characteristics of Intershielded Disc winding are not unique; it looks like Plain Disc winding type.

As stated, Plain Disc windings and Intershielded Disc windings have been widely used in various voltage levels, normally less than or equal to 132 kV but in some cases up to 275 kV. Interleaved Disc windings are appropriate for higher Basic Insulation Level, normally

for a voltage level larger than or equal to 275 kV. Multiple Layer windings are used for all voltage levels by some particular manufacturers due to historic reasons. It is also applicable for voltages from 33 kV right to 400 kV, or even up to 800 kV. In general, windings with high series capacitance, such as Interleaved Disc winding type and Multiple Layer winding type, are more capable to withstand the lighting impulse. All of the four winding types mentioned above can be used for both series and common windings of the 400/275/13 kV transformers to be studied. As for the tertiary winding, Single Helical winding type is customarily used.

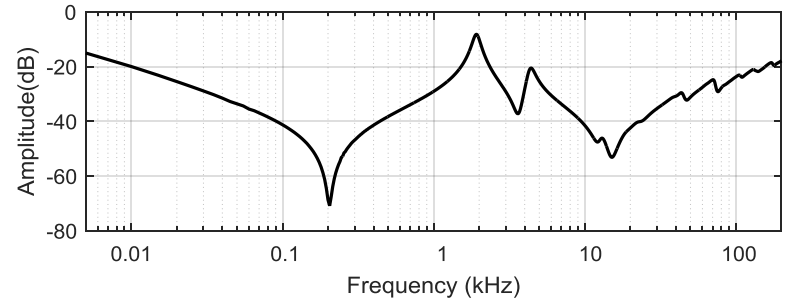
3.5 Investigated Transformers

The 400/275/13 kV autotransformer group will be mainly studied in the following chapters, and are introduced in this section. One faulty transformer of 275/132/11kV used for demonstration is also introduced in detail.

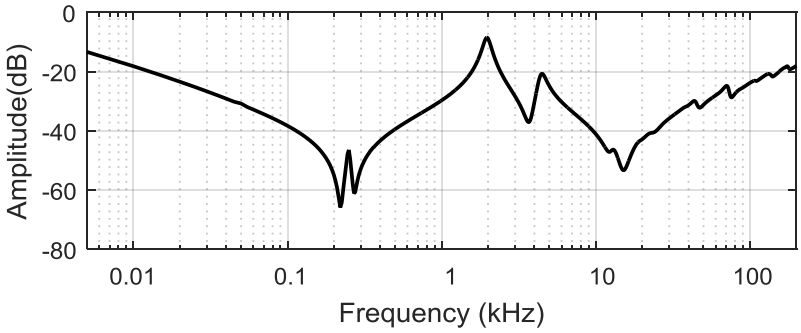
3.5.1 Transfer Function Estimation

Two transfer function estimation methods are introduced in Chapter 4. Two frequency responses from a 500 MVA, 400/275/13/kV autotransformer, as shown in Figure 3-10, are used for demonstration purpose. This 500MVA five-limb autotransformer is manufactured in 1966 by AEI Wythenshawe, and it has Multiple Layer type series and common windings. The measured frequency response on A phases HV to LV terminals is used to demonstrate Feature Extraction Method, while the measured frequency response on B phases HV to LV terminals is used to demonstrate the Extreme Points Identification Method.

The frequency responses from 5 Hz to 200 kHz are investigated, including 1120 frequency points per trace. For this transformer, the frequency region governed by the core is up to 2 kHz; the frequency region controlled by interaction between windings is from 2 kHz to 20 kHz; and frequency region 20 kHz to 200 kHz is dominated by winding properties, i.e., design parameter, and winding type.



(a) B Phase



(a) A Phase

Figure 3-10 FRA Plots from Multiple Layer Winding, measured on Series Winding of T4668

To verify the Feature Extraction Method, the frequency responses of series and common windings from A, B, and C phases of eight transformers are studied. The information of the eight transformers is listed in Table 3-10. Particularly, three frequency responses of different winding types are used to demonstrate the effectiveness of the developed method, which are plotted in Figure 3-11, Figure 3-12, Figure 3-13, including the frequency response from the Interleaved Disc winding (B phase series winding of transformer T6976), and the frequency responses from Plain Disc windings (B phase common and series windings of transformer T3983).

Table 3-10 Transformers to be studied in Chapter 4

TNUMBER	Common winding	Series winding	Power Rating	Manufacturer
T4668	Multiple Layer	Multiple Layer	750MVA	AEI
T5562	Multiple Layer	Multiple Layer	750MVA	AEI
T5291	Multiple Layer	Multiple Layer	750MVA	AEI
T4439	Multiple Layer	Multiple Layer	750MVA	FER
T5492	Plain Disc	Plain Disc	750MVA	FER
T3983	Plain Disc	Plain Disc	750MVA	PPT
T4673	Multiple Layer	Multiple Layer	750MVA	CAP
T6976	Intershielded Disc	Interleaved Disc	1000MVA	GEC

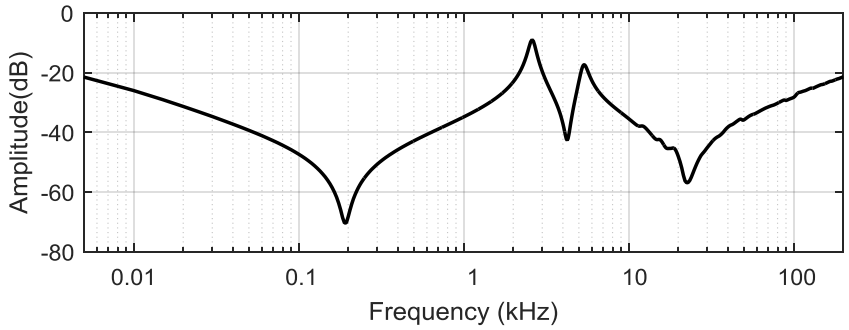


Figure 3-11 FRA Plot of Interleaved Disc Winding, measured on B Phase Series Winding of T6976

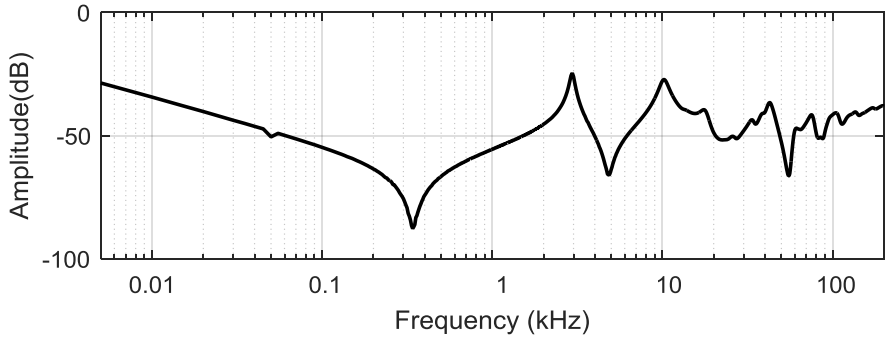


Figure 3-12 FRA Plot of Plain Disc Winding, measured on B Phase Common Winding of T3983

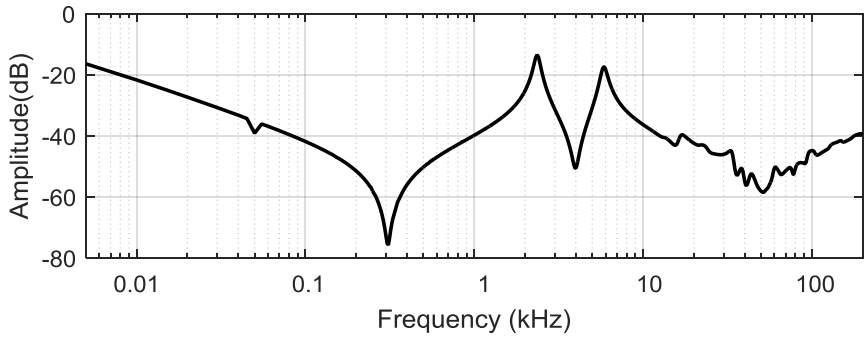


Figure 3-13 FRA Plot of Plain Disc Winding, measured on B Phase Series Winding of T3983

To verify the Extreme Points Identification method, the frequency response of Plain Disc winding type is investigated as shown in Figure 3-14. The example used in Chapter 4 is the measurement made on C Phase series winding of transformer T5492. This transformer is one of the eight transformers in Table 3-10.

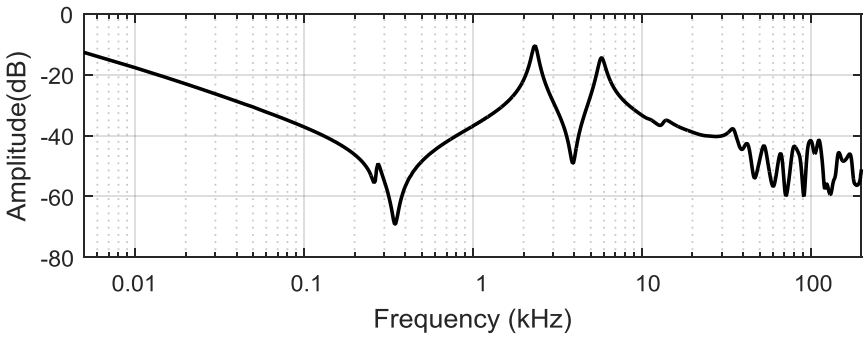
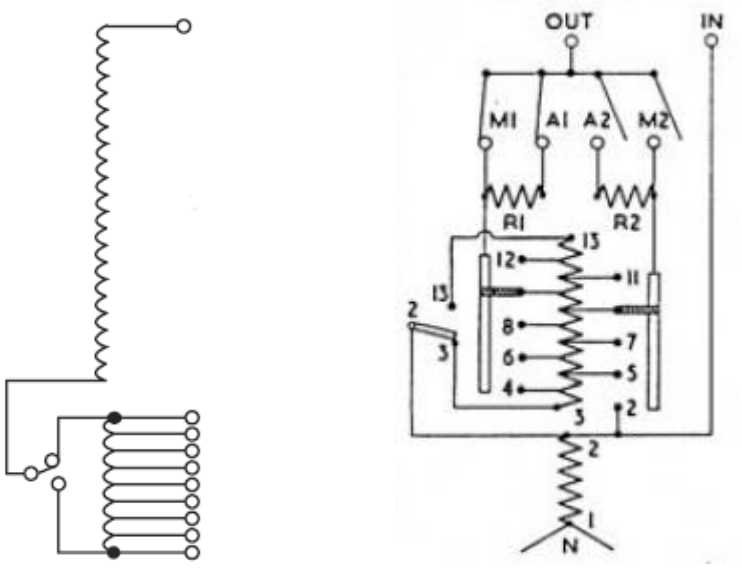


Figure 3-14 FRA Plot of Plain Disc Winding, measured on C Phase Series Winding of T5492

Both of the proposed methods for transfer function estimation are applied for diagnosis of faulty transformers. As an example, a 275/132/11kV, 120MVA, Yan0d_ autotransformer manufactured by Metropolitan Vickers in 1954 and installed at the site of Carrington substation in 1956, numbered as T2305 in the UK National Grid's database, is used in Chapter 4. It has a three-limb core. The transformer has a helical tertiary winding, Plain Disc common winding, five-layer series winding, and a two-layer multi-start tap winding. A line-end revising tap changer made by AEI was used at the terminal of 132 kV with a tapping range of ±15%, in ±9 steps of 1.67% per step. The schematic and physical arrangement of the reversing tapping configuration is shown in Figure 3-17. The winding direction is clockwise for series winding, and anticlockwise for common winding. The winding direction of reversible regulating winding can be set either as clockwise or anticlockwise. Therefore, the winding output voltage can be reduced in a subtractive sense (i.e. the switch is connected from 2 to 3 in Figure 3-17(b)), that is 'buck,' and increased by an additive sense (i.e. the switch is connected from 2 to 13 in Figure 3-17(b)), that is 'boost.' Noticeably, the numbers in Figure 3-17(b) are not the tap positions, but the connection points.

T2305 failed in September 2004, due to an internal short-circuit caused by damage to tap-changer from incorrect oil handling. The B phase tap winding axially collapsed. Figure 3-16 shows the significant tap winding deformation.



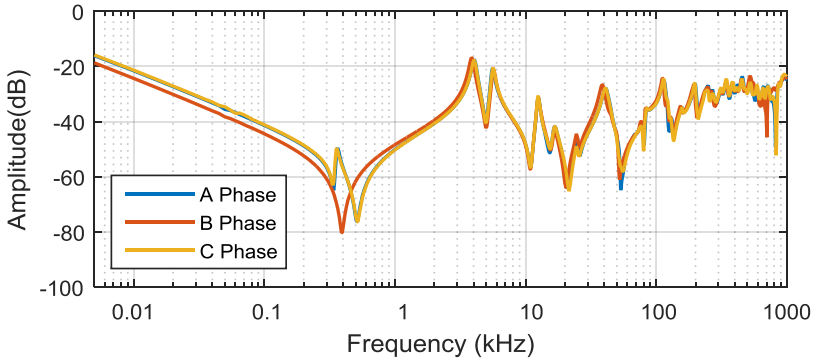
(a) Schematic Arrangement (b) Physical Arrangement

Figure 3-15 Buck/Boost Tapping Arrangement [1]

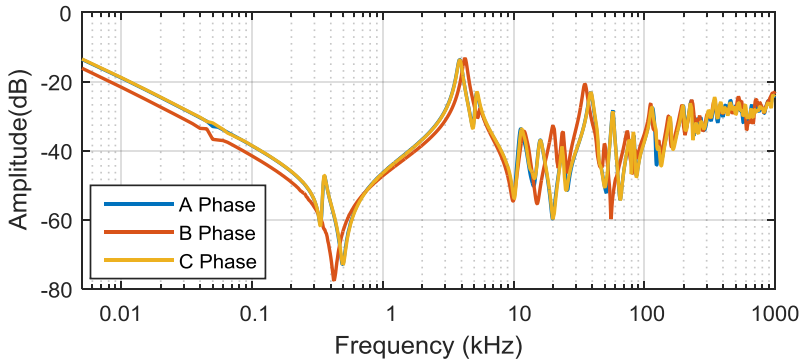
Figure 3-16 Winding Movement of Tap Winding from B Phase T2305

For this transformer, there are no previous frequency response measurement results available before the fault. Therefore, the construction-based comparison should be used. Three FRA measurements were conducted using end-to-end open-circuit test, with the other neutrals earthed. Because the short circuit fault occurred on the LV winding side, the first FRA measurement was conducted on N to LV, with tap position from N to 2. According to Figure 3-15 (b), under this connection scheme, only the LV winding is tested, at nominal voltage. It can be seen in Figure 3-17 (a) that the three frequency responses from A, B, and C phases do not show obvious alteration from 5 Hz to 1 MHz. The minor alterations at about 3 kHz, which is in the frequency region controlled by winding interaction, may be caused by the direct electric connection between the LV and tap winding, similar to the auto-transformer connection which is very familiar to the readers, and the shift of the resonance and antiresonance around 10 kHz to 100 kHz and at about 800 kHz, is caused by the electrostatic and electromagnetic coupling and interaction between LV and tap windings, note that tap winding end is floating. Thus the second FRA measurement was conducted on N to LV with tap position at Tap 19, which is the maximum tap of +15%. This means that the diagnosis measurement include LV + tap winding. The frequency responses from 5 Hz to 1 MHz of A, B, and C phases are plotted in Figure 3-17 (b). The frequency response of B phase is significantly different from the frequency responses of A and C phases, and this helped site test engineers to suspect that the tap winding of B phase is faulty. The third FRA measurement was conducted on the tap winding only, with tap lead position from 3 to 13. According to Figure 3-15 (b), with this connection scheme, the LV and HV windings are not under measurement but only the tap winding. The frequency responses from 5 Hz to 1 MHz are plotted in Figure 3-17(c). The

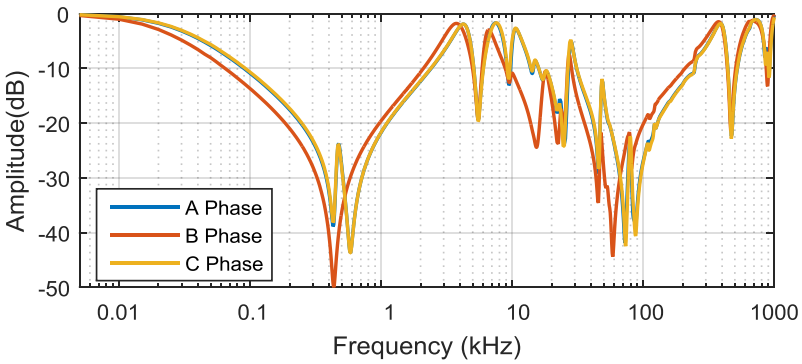
shift of the resonances and antiresonances in the frequency response of B phase can be easily observed.



(a) Series Winding in Nominal Condition



(b) Series Wining with +15% Tap Winding



(c) Tap Winding

Figure 3-17 Frequency Responses of A, B, and C Phase of Transformer T2305

In Chapter 4, the frequency responses from 5 Hz to 200 kHz of the second FRA measurement results shown in Figure 3-17(b) are used to test the proposed methods for transfer function estimation. Within this frequency region, the alteration in the frequency response of B phase can be clearly reflected. The diagnosis frequency response of B phase is compared with the reference frequency response of A phase.

3.5.2 Winding Type Identification

In the FRA database of UK National Grid, both the magnitude and phase spectra were recorded, and only the magnitude frequency spectrum is used for winding type recognition.

As mentioned before, the boundary frequencies of 2 kHz and 20 kHz are empirical; hence the whole range from 5 Hz to 1 MHz will be used in the following study of the SVM model in Chapter 5. There exists a great similarity for FRA traces in the frequencies lower than 20 kHz for all the Multiple Layer, Plain Disc and Interleaved Disc windings, and this similarity is regarded reasonable as this frequency region is dominated by the core and the interaction between windings. Inclusion of the low frequency regions would increase the complexity for classification but enhance the confidence level of the SVM model developed.

For the Multiple Layer winding type, 30 FRA traces are used from two 500 MVA 400/275/13 kV transformers and three 750 MVA 400/275/13 kV transformers, including the common and series windings from A, B, and C phases, and they are plotted in Figure 3-18. For the Plain Disc winding type, 36 FRA traces are investigated from four 750 MVA and two 1000MVA 400/275/13 kV transformers, including common and series windings from A, B, and C phases, and they are shown in Figure 3-19. The choices of FRA traces for the Interleaved Disc winding type are rather limited because only a small quantity of frequency responses from this type is available in the database. For this winding type, 18 FRA traces are used, and they are from three 750MVA 400/275/13 kV transformers of the same manufacturer, from A, B, and C phase of common and series windings, as plotted in Figure 3-20.For the Single Helical winding type, 6 FRA traces from the tertiary windings of two 750 MVA 400/275/13 kV transformers and 9 FRA traces from two 1000 MVA and one 750MVA 400/275/13 kV transformers are used, as plotted in Figure 3-21.

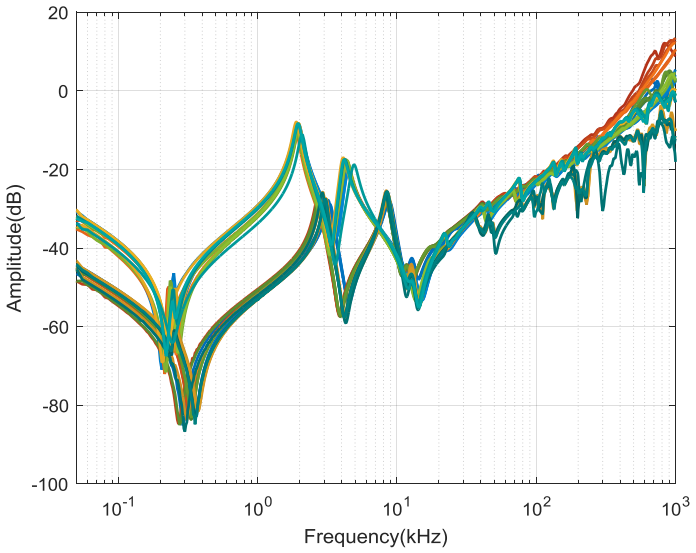


Figure 3-18 FRA Plots of Multiple Layer Windings

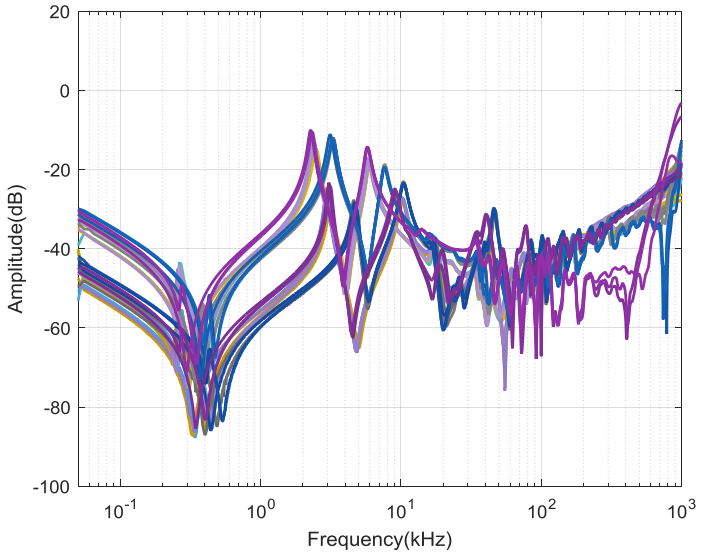


Figure 3-19 FRA Plots of Plain Disc Windings

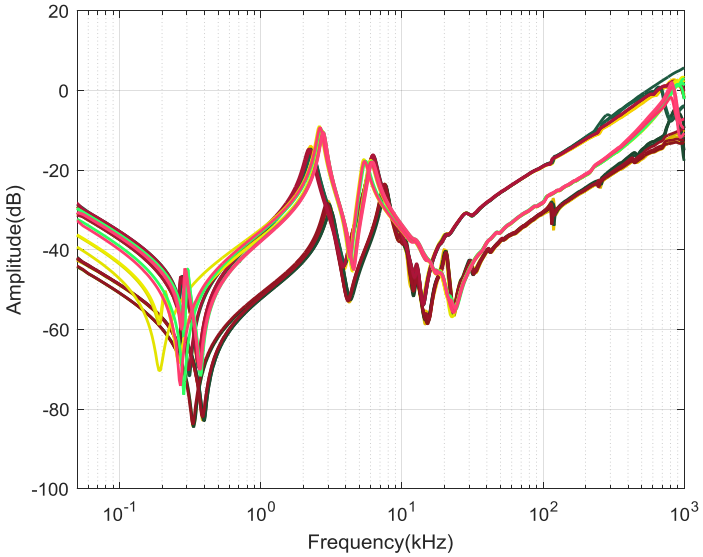


Figure 3-20 FRA Plots of Interleaved Windings

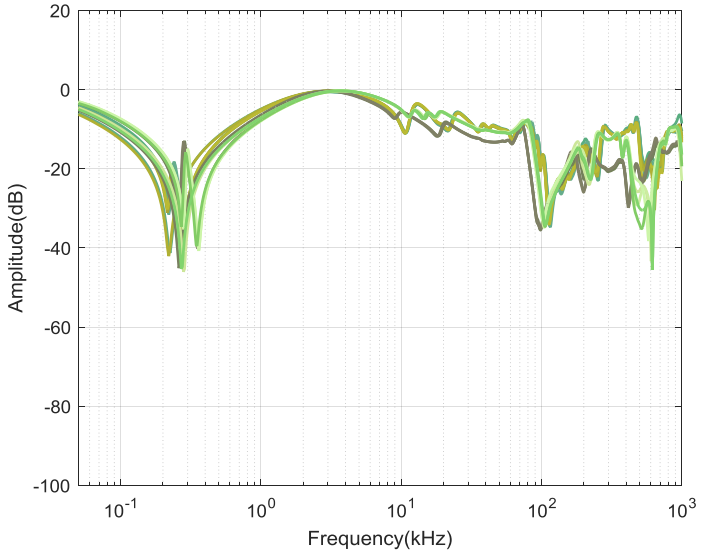


Figure 3-21 FRA Plots of Single Helical Windings

For the study of unsupervised machine learning methods in Chapter, to avoid the influence of measurement setup, the measured FRA data up to 200 kHz are used. The 28 frequency responses under investigation are from five different winding types, i.e., Intershielded Disc winding type, Plain Disc winding type, Multiple Layer winding type and Single Helical winding type. The frequency responses are plotted in Figure 3-22 and their information are listed in Table 3-11.All of the 28 frequency responses are from 11 400/275/13 kV autotransformers, as shown in Table 3-11. Among the 11 transformers, 6 of them are used in Chapter 4, i.e. T6976, T3983, T4668, T4673, T5562 and T5492, and their information can be found in Table 3-10. The information of other 5 transformers is listed in Table 3-11(c).

The FRA traces [34] of Intershielded Disc windings are from the common and series windings of T7014 and T6976. The FRA traces {5-10} of Plain Disc windings are from the common and series windings of T3983, T4297, and T6124. The FRA traces {11-14} of Multiple Layer windings are from the common and series windings of T4668 and T4673. For the 6 FRA traces from Interleaved Disc windings, {15, 16} are from the series windings of T7014 and T6976, while {17-20} are from the common and series windings of T4083 and T4082. All of the {1-20} traces are from A phase, expect that {2, 11, 13} are from B phase. FRA traces from different phase have different patterns in the low frequency region <2 kHz, thus the clustering difficulty is increased by using FRA traces from different phases. For the Single Helical winding, {21-24} are from T5562, {25-27} are from T5492 and {28} is from T3983.

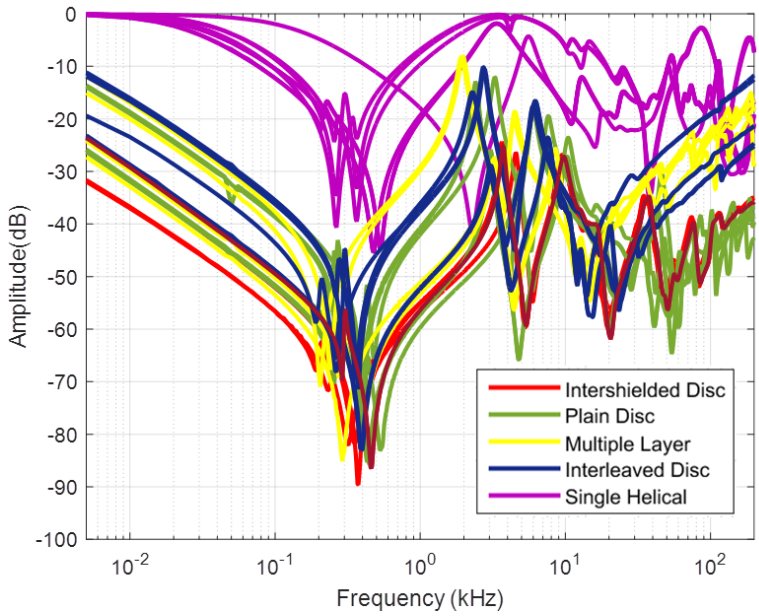


Figure 3-22 FRA Plots of Common, Series and Tertiary Windings under Investigation in Chapter 6

Table 3-11 Basic Information of FRA Traces under Investigation in Chapter 6

(a) Winding Construction Types

Winding Type	Amount of Trace	Trace Label	Transformer
Intershielded Disc	4	{1, 2, 3, 4}	T7014 T6976
Plain Disc	6	{5, 6, 7, 8, 9,10}	T3983 T4297 T6124
Multiple Layer	4	{11,12,13,14}	T4668 T4673
Interleaved Disc	6	{15,16,17,18,19,20}	T7014 T6976 T4083 T4082
Single Helical	8	{21,22,23,24,25,26,27,28}	T5562 T5492 T3983

(b) HV/LV/TV Winding

Winding Type	Trace Number	Trace Label
Series Winding	9	{5,6,7,12, 15,16,17,18}
Common Winding	11	{1,2,3,4,8,9,10,13,14,19,20}
Tertiary	8	{21,22,23,24,25,26,27,28}

(c) Transformer Information

Tnumber	Common winding	Series winding	Power Rating	Manufacturer
T7014	Intershielded Disc	Interleaved Disc	1000MVA	GEC
T4297	Plain Disc	Plain Disc	500MVA	Bruce Peebles
T6124	Plain Disc	Plain Disc	1000MVA	Hackbridge and Hewittic
T4083	Interleaved Disc	Interleaved Disc	750MVA	English Electric
T4082	Interleaved Disc	Interleaved Disc	750MVA	English Electric

3.6 Summary

In this chapter, the FRA Database of UK National Grid is introduced and analysed. The FRA data were acquired by the same measurement device and they are stored in the database in the same format, hence it is easy to extract the FRA data.

Then the basic information of the transformers are analysed, including the voltage ratio, power rating and their winding types. The statistics of failed transformers are given. The quantity distribution of transformers for different voltage ratios summarized, and the most common used voltage ratios includes 275/132 kV, 400/132 kV and 400/275/13 kV. For each voltage ratio, the quantity distributions for different power ratings are summarised in Table 3-1 and Table 3-2. The correlation between the transformer designs and the FRA traces are investigated. All the transformers with known winding types are summarised in Table 3-3 to Table 3-6. It is found that the Plain Disc winding is the most commonly used winding structure. The influence from the voltage and power rating are analysed by comparing the FRA traces.

The influence of transformer winding designs on FRA traces is introduced using the 400/275/13kV auto transformers as example. The typical FRA traces of different winding types are illustrated, and their characteristics are introduced. Lastly, the frequency responses to be used in the following chapters are introduced along with the corresponding transformer information.

Chapter 4 Transfer Function Estimation

4.1 Introduction

Frequency Response Analysis (FRA) is a sensitive technique to identify the mechanical health condition of transformer windings. If variation occurs to the reference and diagnostic frequency responses measured on the same windings with the same connection scheme, a winding displacement or deformation may be indicated.

Although the FRA technique was introduced to diagnose transformer mechanical faults several decades ago, no IEC/IEEE FRA interpretation standard has been established yet. There have been three routes to investigate FRA interpretation; one is to develop the fundamental understanding of FRA traces through simulation and modelling, among which there are white box and grey box models, and/or using artificial intelligence algorithms. White box models are based on transformers' design data, and grey box models are based on various measurement results at transformer terminals including frequency responses. The second route is more direct, attempting to establish "yes/no" rules by quantifying the similarity or difference between the reference and diagnostic frequency responses, among which both numerical indices and transfer function expression were used. The third investigation route is to utilize artificial intelligence algorithms including pattern recognition methods to classify fault type/location/severity and so on.

As stated in Chapter 2, the numerical index is a single value reflecting the relationship of two sets of data to be compared, and windings can be mechanically classified into different levels of health condition according to the indices extracted from the low, middle and high frequency ranges [40-42, 44, 45, 47, 56, 65]. However, this method is highly experience based and the associated criteria developed based on limited experience are often challenged by new cases in field. On the other hand, a transfer function uses a set of parameters to mathematically represent a FRA trace, as long as the representation is accurate, the set of parameters contains more information than the numerical index, which creates possibilities for detailed interpretation and a possibly more realistic development route for interpretation standard.

The rational transfer function, T_s, can be expressed by zeros and poles as follows:

$$T_s = k_s \frac{\prod_{k=1}^{N1}(s-z_{rk})}{\prod_{j=1}^{M1}(s-p_{rj})} \times \frac{\prod_{k=1}^{N1}(s-z_{ck})(s-z'_{ck})}{\prod_{j=1}^{M1}(s-p_{cj})(s-p'_{cj})} \quad \text{Equation 4-1}$$

where k_s is the constant coefficient, z_{rk} and p_{rj} are real zero and real pole respectively, and z_{ck}, z'_{ck} are conjugate complex zeros, p_{cj}, p'_{cj} are conjugate complex poles respectively. The complex zeros and poles are in the form $s=\sigma+j\omega$, where $\omega=2\pi f$ and f is the frequency location of resonance or anti-resonance, σ is the frequency dependent damping rate. For real zeros and poles, $\sigma=0$. The transfer function is physically achievable when its real zeros and poles are negative, and its complex zeros and poles own negative imaginary parts.

Two methods to estimate the transfer function representation of FRA measurement results are proposed in this chapter, both of which use the rational transfer function expression.

For the first method, a program is developed to auto-process the measured data and to obtain the transfer function expression. It is implemented in MATLAB using the inbuilt command 'invfreqs', which is based on Non Linear Least Square algorithm. This method requires both the FRA magnitude and phase spectra as the input. Measurement results across the large frequency range are divided into several frequency regions, and key information, i.e., complex poles and zeros, is abstracted from each frequency region to form the Feature Transfer Function. Afterwards, correction is needed to eliminate the difference between the Feature Transfer Function and measured FRA data. This method is verified as valid by estimating the transfer functions for 48 groups of measured frequency responses of 400/275/13 kV auto-transformers.

The second method is called the Extreme Points Identification Algorithm, which aims to find a unique solution for a physically achievable transfer function, and this is crucial for the interpretation of FRA results. The proposed method uses the FRA magnitude spectrum only. The extreme points, i.e., resonance and antiresonance, on a FRA magnitude trace are identified to initialise parameters for its mathematical expression. Both the magnitude and frequency location of an extreme point are not only determined by itself but also are mutually influenced by other extreme points. Hence iterations are carried out to optimise the parameters. All the finalised parameters predicted by this method are physically achievable, and the quantity of parameters is controlled to be the same as twice of the total number of peaks and troughs plus 1, which ensures the simplification of the mathematical transfer function expression. This algorithm has been successfully applied on the selected FRA traces from the UK National Grid's database.

As mentioned, the FRA measurement data in the UK National Grid's database contains 400 evenly distributed data points each for the following five frequency regions: 5Hz to 2 kHz, 50 Hz to 20 kHz, 500 Hz to 200 kHz, 5 kHz to 2 MHz and 25 kHz to 10 MHz. The frequency responses studied for the transfer function estimation in this chapter are from 5 Hz to 200 kHz, and each contains 1120 frequency points.

4.2 Feature Extraction Method

The FRA trace in Figure 4-1, which was introduced in Chapter 3, is used to demonstrate the proposed Feature Extraction Method. The oscillations from 5 Hz to 100 Hz in the phase response data are resulted from noise.

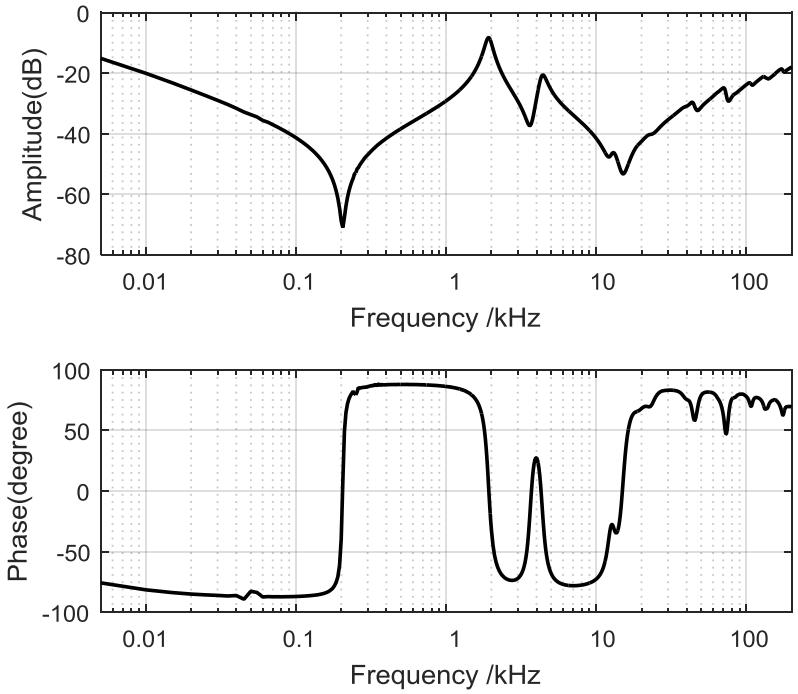


Figure 4-1 FRA Plot from Multiple Layer Winding, measured on B Phase Series Winding of T4668

4.2.1 Methodology

A flow chart, as given in Figure 4-2, is used to show the estimation process of the proposed Feature Extraction Method. The frequency response data is firstly pre-processed as input, and the whole frequency range from 5 Hz to 200 kHz is divided into several frequency regions, according to the number of resonance points and the amount of data points in each frequency region. Complex zeros and poles are extracted from each frequency region and combined to construct a Feature Transfer Function (FTF), which has a similar shape to the measured frequency response. The difference between FTF and the measured frequency response is corrected to produce the Finalised Transfer Function (FDTF).

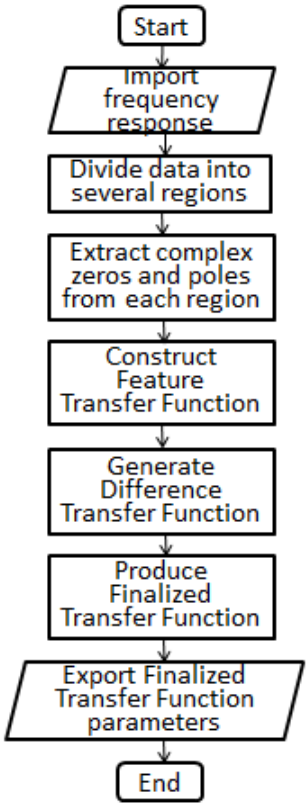


Figure 4-2 Flow Chart of Feature Extraction Method

MATLAB command 'invfreqs' can be employed to estimate transfer function expression of the measured frequency response in the continuous frequency domain. Both zeros and poles, complex or real, can be produced by 'invfreqs'. The complex zeros and poles, $s=\sigma\pm j\omega$, are corresponding to the resonances or antiresonances, and the location f of which can be roughly calculated as:

$$f \approx \frac{|\omega|}{2\pi} \qquad \text{Equation 4-2}$$

After the resonant point which is controlled by a pair of complex poles, the magnitude drops 40 dB per decade, whilst the magnitude rises 40 dB per decade after an antiresonant point controlled by a pair of complex zeros. A pair of complex zeros cause the phase response to rise 180° after the antiresonant points, whilst a pair of complex poles cause the phase response to drop 180° after the resonant points. The real zeros and poles do not cause any resonance or antiresonance and they both contribute to the flat magnitude at the beginning part of a frequency response, and afterwards the real zero makes the magnitude rise 20 dB per decade, and the real pole makes the magnitude drop 20 dB per decade. A real zero causes the phase response to rise 90° while a real pole causes the phase response to drop 90°.

Location of the inflection point f' is related to the absolute value $|\omega|$ of real zero and pole:

$$f' = \frac{|\omega|}{2\pi}$$ Equation 4-3

The ‘invfreqs’ command cannot be used directly for the whole frequency range, because that the number of resonant points for the whole frequency range is huge, as a result a transfer function with very high order is needed. However, the ‘invfreqs’ command is only able to process the estimation of transfer function with low order. Experience shows the ‘invfreqs’ command is able to estimate transfer functions with an order up to 30, and this number may vary slightly with different situations. Therefore, the division of frequency regions is needed to guarantee that the amount of the resonant points in each frequency region is within the estimation ability of this inbuilt command. Noticeably, in the low frequency region from 5 Hz to 2 kHz, the measured frequency points have a larger quantity of 400, and there may be noises in this region. Though there may be only several resonant points with in this frequency region, it is found that better estimation result can be obtained by dividing this region into two smaller regions, i.e. 5 Hz to 1 kHz and 1 kHz to 2 kHz.

Although the frequency range from 5 Hz to 200 kHz is investigated in this study, the proposed method is applicable to any frequency range. In the case of a wide frequency range, the order of the Difference Transfer Function, which represents the difference between the Feature Transfer Function and the measured data, may exceed the estimation ability of the inbuilt command, and the whole frequency range should then be divided into several regions to estimate this Difference Transfer Function.

The frequency boundaries of division may vary accordingly. Although they are fixed in this study and are based on experience, it is expected that as long as the appropriate amount of resonant points and data points are include in each region, the estimation accuracy would not be affected. This method intrinsically ensures the key information to be extracted from each frequency region without any effect from the other frequency regions.

4.2.1.1 First Step: Feature Transfer Function

There are three stages to construct a Feature Transfer Function (FTF).

A. Frequency Range Division

In the first stage to construct the Feature Transfer Function, it is compulsory to divide the whole frequency range into several frequency regions for the aforementioned reasons. Each frequency region needs to have sufficient number of data points to avoid some of the zeros

or poles which should be extracted to be ignored. The frequency response in Figure 4-1 is used to show how to estimate its transfer function for the frequency range from 5 Hz to 200 kHz. All the 1120 data points are divided into 7 frequency regions, i.e., 5 Hz to 1 kHz, 1 kHz to 2 kHz, 2 kHz to 20 kHz, 20 kHz to 65 kHz, 65 kHz to 110 kHz, 110 kHz to 155 kHz and 155 kHz to 200 kHz.

B. Selection of Order of Transfer Function

The command 'invfreqs' requires to input the orders of numerator and denominator of the transfer function to be estimated, which are equal to the amounts of zeros and poles respectively. When estimating a transfer function for each frequency region, the choice of the order has an impact on the precision of estimation. If the order is smaller than what is demanded, some features on the FRA trace may be ignored. Nevertheless, redundant zeros and poles may appear if the selected order is more than adequate. There is no such a relationship as that the estimation accuracy can be improved with the increase of the order. For simplification, the order of the numerator is set equal to that of the operator. An optimal order should be selected for each frequency region.

There exists a wide range selection for the transfer function's order for each frequency region. The possible range of transfer function order is predefined for each frequency region, based on the approximate amounts of zeros and poles. For each candidate order, estimation can be conducted using 'invfreqs'. If there exist redundant zeros and poles, the redundant parameter are supposed to be cancelled before comparing one candidate order with other orders.

The complex zeros and poles satisfy the following criteria are viewed as redundant:

- For the imaginary part, the variation between the frequency location of zeros and poles should be smaller than 2/5 of the sampling interval;
- For the real part, the absolute difference value should be smaller than 60, or when the absolute difference value is larger than 60, the variation in percentage should be less than 0.6%.

The criteria of '2/5 of the sampling interval', '60', and '0.6%' are empirical values and experience based.

The first criterion ensures that the redundant complex zeros and poles which are located close enough to be cancelled, and the second criterion ensures that the damping rate of the redundant complex zero and pole are viewed to be nearly the same. In the low frequency band, the damping rate is small. It may occur that although the absolute difference in the real part is minor, the relative percentage difference is huge. Therefore, the second criterion is made such that the redundant complex zeros and complex poles in low frequency region, which have small real part, can be cancelled. Of course for the redundant real zeros and real poles, only the second criterion should be satisfied.

To compare different orders, the maximum magnitude and phase errors between the measured and estimated data is recorded for each order, and the Correlation Coefficient is also calculated for the estimated and measured frequency responses for each order. The optimal order should be selected according to the following rules:

- The maximum magnitude error of the optimal order should be smaller than 3 times the minimum value of those from all the rest orders. The same rule applies to phase. This rule guarantees that the local deviations for both magnitude and phase are within acceptable range.
- The magnitude Correlation Coefficient of the optimal order should be larger than 0.9999. This rule guarantees that the global similarity in the shape of the measured and estimated magnitude data.
- The phase Correlation Coefficient should be as close to 1 as possible. This rule guarantees that the global similarity in the shape of the measured and estimated phase data. Because the phase is more sensitive than the magnitude, when the local deviation is controlled and the match of magnitude data is ensured, the optimal order is finally chosen as the one with the maximum phase Correlation Coefficient.

The ‘3 times’ in first criteria of order selection is empirically based. Noticeably, though Correlation Coefficient has the drawback that it only reflects the similarity in the shape of the traces, this problem is solved by using maximum errors of the magnitude and phase in combination.

For instance, according to the criteria set above, for the frequency region from 2 kHz to 20 kHz, the optimal order is selected as 22. The estimated and measured frequency responses are plotted for comparison in Figure 4-3.

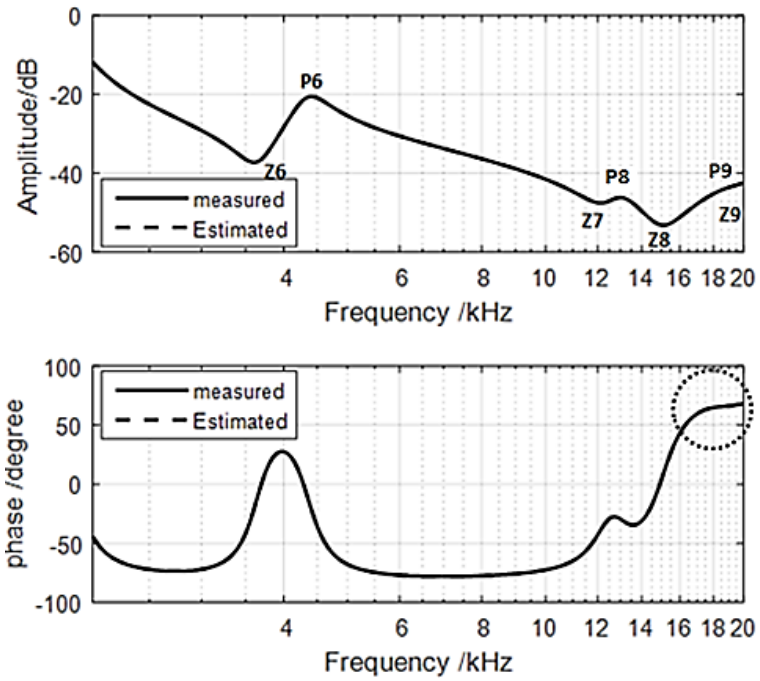


Figure 4-3 Measured and Estimated FRA Plot on 2-20 kHz Frequency Region

Table 4-1 lists the parameters of the estimated transfer function with order 22, as shown in Figure 4-3, over the frequency region from 2 kHz to 20 kHz, after redundant zeros and poles are cancelled. For simplification, for a pair of conjugate complex zeros/poles, only the zero/pole with positive imaginary part is listed, in the order of frequency from low to high. It can be seen that the pair of complex zeros and pair of complex poles listed in row 5 are located at 0.07 kHz and 1.91 kHz separately, outside the frequency region 2 kHz to 20 kHz. Though for the concerned frequency region such complex parameters have an influence on magnitude of the estimated transfer function, no resonance or antiresonance are caused. The complex zeros Z9 and poles P8 locate at 18.71 kHz and 18.55 kHz, though the resonance and antiresonance cannot be obviously seen on the magnitude response, their effects can be clearly observed on the phase response, as shown in the circle in Figure 4-3.

Table 4-1 Estimated 2-20 kHz Transfer Function Parameters

	Zero	Pole	fz(kHz)	fp(kHz)
1	-244413	-108552	-	-
2	-90929	-2314	-	-
3	115561	126547	-	-
4	188838	152950	-	-
5	-1080+413i	-717+12025i	0.07	1.91
6	-1381+22869i	-1763+27247i	3.64	4.34
7	-5488+77217i	-339680+76899i	12.3	12.25
8	-6941+94283i	-5601+80918i	15.01	12.89
9	-13412+117507i	-12298+116477i	18.71	18.55

C. Extraction of Zeros and Poles

The last stage of step one is to build the Feature Transfer Function using the extracted conjugate complex zeros and poles from each frequency region. Each antiresonance should be corresponding to a pair of conjugate complex zeros, whilst each resonance should be corresponding to a pair of conjugate complex poles.

When extracting complex zeros and poles from each frequency region, the following two criteria should be met:

- Only the complex zeros and poles located within the estimated frequency region should be selected.
- Complex zeros or poles whose ratio between real and imaginary parts larger than 2 should not be selected.

It is found that the relative height as compared to those surrounding frequency points caused by complex zeros or poles is related with the ratio of their real and imaginary parts. If the ratio of real and imaginary parts is larger than 2, this pair of complex zeros/poles can be regarded as a pair of real zeros/poles, as the effects of this pair of complex zeros/poles are identical to the real zeros/poles. Thus the second selection criterion should be used. This conclusion is based on the sensitivity study on the effects of the ratio between the real and imaginary parts on the frequency responses, and a ratio equal to 2 is used as the boundary based on experience.

Among all the zeros and poles in Table 4-1, Z_6 to Z_9 and P_6, P_8, P_9 are extracted. The complex pole P_7 is not corresponding to any resonance. As mentioned, it can be seen that Z_9 and P_9 are located closely in frequency, though no resonance nor antiresonance can be seen in the magnitude spectrum around 18.5 kHz. However, in Figure 4-3, the corresponding frequency band in the dashed circle on the phase spectrum shows the slight downward and upward trends, and this proves the presence of the complex zeros and poles. It means that the criteria set are appropriate.

After the application of the criteria above, all the chosen parameters from seven frequency regions are listed in Table 4-2, which contains19 complex zeros and 18 complex poles in total. A constant coefficient k with value 1 is used for the Feature Transfer Function. The FRA trace constructed using the chosen complex zeros and poles of Feature Transfer Function, T_1, are plotted in Figure 4-4.

Table 4-2 FTF Parameters

k=1				
1	**Zero**	**Pole**	**fz(kHz)**	**fp(kHz)**
2	-33+1282i	-722+12029i	0.204	1.915
3	-1381+22869i	-1763+27247i	3.640	4.337
4	-5488+77217i	-5601+80918i	12.290	12.879
5	-6941+94283i	-12298+116477i	15.006	18.538
6	-13412+117507i	-13110+146501i	18.702	23.317
7	-11946+148301i	-19154+239420i	23.604	38.106
8	-20283+239837i	-14415+265817i	38.172	42.307
9	-13739+264539i	-17389+280834i	42.104	44.697
10	-15783+288870i	-22272+396277i	45.976	63.071
11	-23034+396277i	-23585+427230i	63.071	67.998
12	-23285+428596i	-19294+454204i	68.215	72.291
13	-19310+466778i	-22937+519877i	74.292	82.743
14	-23621+521010i	-22932+675461i	82.924	107.506
15	-21766+679055i	-20724+836681i	108.078	133.166
16	701958+775922i	-35857+861613i	123.495	137.134
17	-19639+837936i	899351+885796i	133.366	140.983
18	-34724+866957i	-30636+1102403i	137.985	175.458
19	-28682+1106887i	-83579+1216002i	176.172	193.538

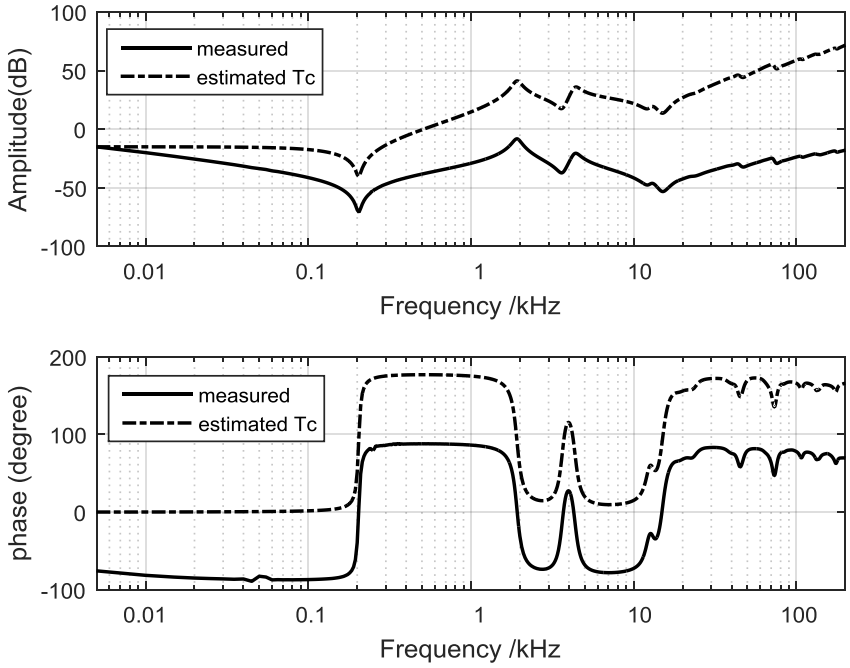


Figure 4-4 Comparison of FTF Plot and Measured Data

It can be seen that a pair of complex zeros and a pair of complex poles are located both at 63.071 kHz in Table 4-2, and their real parts are very similar. However, those parameters cannot be cancelled; otherwise small difference will be caused both on the magnitude and phase traces. This is because that the resolution on the magnitude spectra is not fine enough

to show the resonance or antiresonance, whilst the phase spectra are very sensitive to reflect the existence of such closely located complex poles/zeros.

At this stage, the Feature Transfer Function T_1 is constructed in Equation 4-4:

$$T_1 = \frac{\prod_{k=1}^{n2}(s - z_{ck})(s - z'_{ck})}{\prod_{j=1}^{m2}(s - p_{cj})(s - p'_{cj})} \quad \text{Equation 4-4}$$

4.2.1.2 Second Step: Difference Transfer Function

It can be seen in Figure 4-4 that the magnitude and the overall trace shape of FTF are not consistent with the measured FRA data. Therefore, in the second step, FTF is compared with measured data, and the Difference Transfer Function (DTF) should be used to correct and represent the deviation.

The magnitude data m_2 and phase data h_2 of the DTF can be computed as:

$$m_2 = \frac{m_{measured}}{m_1} \quad \text{Equation 4-5}$$

$$h_2 = h_{measured} - h_1 \quad \text{Equation 4-6}$$

where $m_{measured}$ and $h_{measured}$ are the measured magnitude and phase data, m_1 and h_1 are the magnitude and phase of the FTF respectively. The difference magnitude and phase spectra are shown in solid line in Figure 4-5.

When choosing the optimal order of DTF using 'invfreqs', the aforementioned criteria apply, except that the data error for each possible order is ignored in low frequency region below 10 Hz. This is because that the first couple of data points tend to have large variation whilst those at the high frequency region match with each other well, especially for phase estimation. This may be caused by the fact that the frequency range is too large. The estimation of DTF is done for the whole frequency range. Such relaxation in the criteria is within the acceptable range since only the first one or two data are significantly different.

After the redundant zeros and poles are cancelled from the DTF, Figure 4-5 compares the estimated frequency response representing the difference, and Table 4-3 lists the parameters of this transfer function in the following format:

$$T_2 = k_s \frac{\prod_{k=1}^{N1}(s - z_{rk})}{\prod_{j=1}^{M1}(s - p_{rj})} \times \frac{\prod_{k=1}^{n2'}(s - z_{ck})(s - z'_{ck})}{\prod_{j=1}^{m2'}(s - p_{cj})(s - p'_{cj})} \quad \text{Equation 4-7}$$

The difference between the calculated and estimated difference frequency responses in the first couple of data points is enlarged to view due to the logarithmic frequency scale used in Figure 4-5.The DTF is only about the offset, because all of the complex zeros and poles causing resonance and antiresonance, which can be easily observed on the frequency response, are excluded.

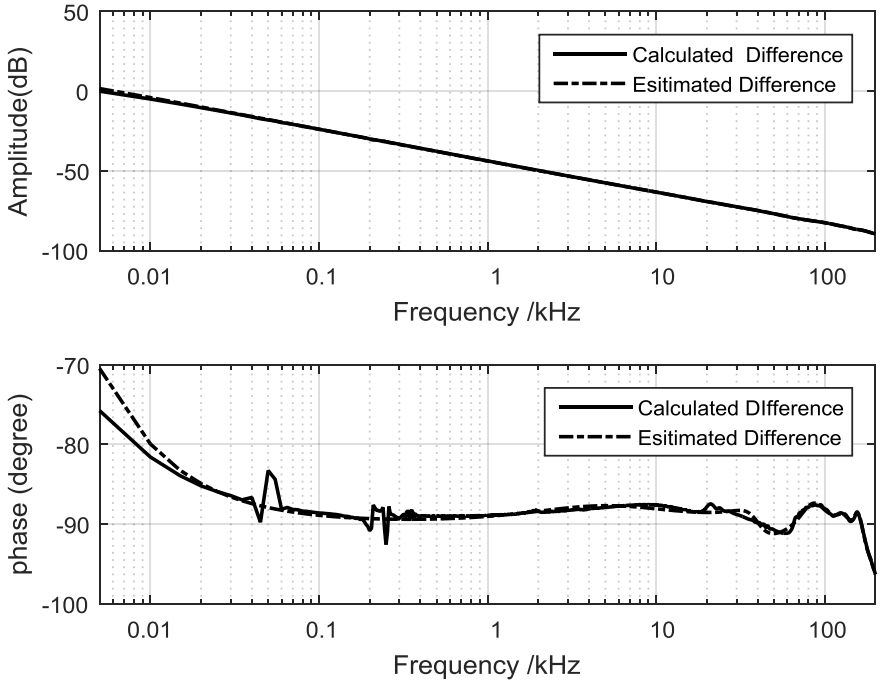


Figure 4-5 Comparison of DTF Plot and Calculated Data

Table 4-3 DTF Parameters

k=-3.4369e-14				
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	3068285	-11	-	-
2	-30400	-32878	-	-
3	-61165+240663i	-58293+241886i	38.304	38.498
4	-137983+522372i	-138914+533428i	83.141	84.900
5	-69725+872229i	684191+852497i	138.824	135.683
6	-112405+994077i	-67101+871369i	158.217	138.687
7	807949+1028917i	-109242+1002993i	163.762	159.636

4.2.1.3 Third Step: Finalized Transfer Function

The last step is to correct the FTF with the DTF, and thus the Finalised Transfer Function (FDTF), is obtained. FDTF can be achieved by multiplying FTF and DTF:

$$T_s = T_1 \times T_2$$ Equation 4-8

When combining the FTF with the DTF, the constant coefficient of FDTF is the same as the constant coefficient of DTF. All the zeros and poles listed above for FTF and DTF are contained in FDTF.

Table 4-4 lists all the parameters of FDTF. For no-source network, zeros and poles of transfer function are expected to appear alternatively. However, for some zeros and poles in Table 4-4, the sequence is irregular. And such areas can be called fuzzy areas, noted in grey colour in Table 4-4, while the remaining areas are called explicit areas. Actually, there are roughly 11 troughs and 10 peaks which are dominant and obvious on the measured frequency response. However, there exist 24 pairs of complex zeros and 23 pairs of complex poles in the estimated transfer function FDTF, as shown in Table 4-4, the amount of which is much larger than the actual number of peaks and troughs. This is because that the proposed method is able to describe the delicate features of the frequency response. Figure 4-6 compares the final estimated transfer function traces with measured data, and it can be seen that the result is satisfactory.

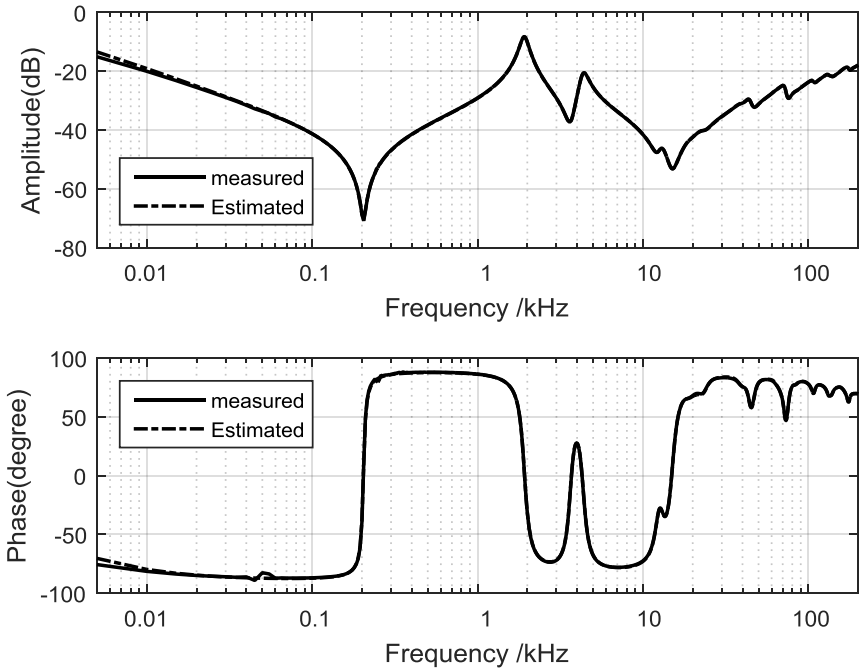


Figure 4-6 Comparison of FDTF Plot and Measured Data of Multiple Layer Winding

Table 4-4 FDTF Parameters

k=-3.4369e-14				
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	3068285	-32878	-	-
2	-30400	-11	-	-
3	-33+1282i	-722+12029i	0.204	1.915
4	-1381+22869i	-1763+27247i	3.640	4.337
5	-5488+77217i	-5601+80918i	12.290	12.879

6	-6941+94283i	-12298+116477i	15.006	18.538
7	-13412+117507i	-13110+146501i	18.702	23.317
8	-11946+148301i	-19154+239420i	23.604	38.106
9	-20283+239837i	-58293+241886i	38.172	38.498
10	-61165+240663i	-14415+265817i	38.304	42.307
11	-13739+264539i	-17389+280834i	42.104	44.697
12	-15783+288870i	-22272+396277i	45.976	63.071
13	-23034+396277i	-23585+427230i	63.071	67.998
14	-23285+428596i	-19294+454204i	68.215	72.291
15	-19310+466778i	-22937+519877i	74.292	82.743
16	-23621+521010i	-138914+533428i	82.924	84.900
17	-137983+522372i	-22932+675461i	83.141	107.506
18	-21766+679055i	-20724+836681i	108.078	133.166
19	701958+775922i	684191+852497i	123.495	135.683
20	-19639+837936i	-35857+861613i	133.366	137.134
21	-34724+866957i	-67101+871369i	137.985	138.687
22	-69725+872229i	899351+885796i	138.824	140.983
23	-112405+994077i	-109242+1002993i	158.217	159.636
24	807949+1028917i	-30636+1102403i	163.762	175.458
25	-28682+1106887i	-83579+1216002i	176.172	193.538
26	-80724+1215272i	-	193.422	-

4.2.2 Application on Other Winding Construction Types

The proposed method has been tested on 48 groups of FRA data of 400/275/13 kV auto-transformers and 3 typical examples are given in the following sections.

4.2.2.1 Interleaved Disc Winding

An estimation example of the frequency response from an Interleaved Disc winding is plotted in Figure 4-7, and the estimated parameters are listed in Table 4-5. For the Interleaved Disc winding, many closely located resonant and antiresonant points exist in the 20 kHz to 200 kHz frequency region. All the fuzzy zooms are noted in grey, and it can be seen that the more the small oscillations, the more frequent the fuzzy areas appear.

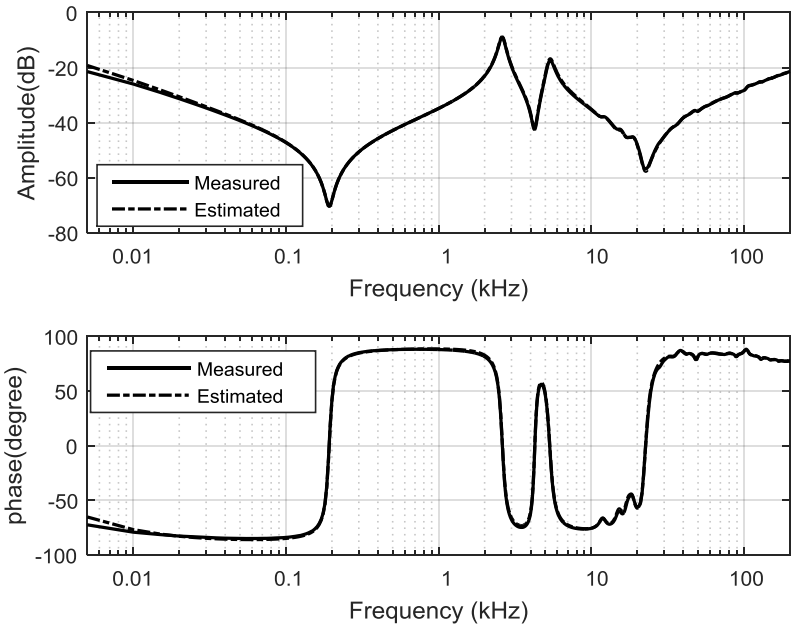


Figure 4-7 Comparison of FDTF Plot and Measured Data of Interleaved Disc Winding

Table 4-5 Estimated Parameters of Interleaved Disc Winding by Feature Extraction Method

	k=1.1023e-1			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	1538798	1290631	-	-
2	-321119	-43770	-	-
3	-39735	-16	-	-
4	-57+1205i	-793+16374i	0.192	2.607
5	-912+26724i	-1893+33467i	4.255	5.329
6	-5017+74358i	-4905+75235i	11.840	11.980
7	-5486+94979i	-5148+96166i	15.124	15.313
8	-5706+109015i	-5099+109969i	17.359	17.511
9	-7620+113403i	-7886+117228i	18.058	18.667
10	-7907+142554i	10213+144180i	22.700	22.959
11	9575+143832i	-58730+152390i	22.903	24.266
12	-60708+163167i	-4978+206797i	25.982	32.929
13	-5347+206680i	-13611+247769i	32.911	39.454
14	-14621+246943i	-10551+303233i	39.322	48.286
15	-10755+304547i	-22448+352930i	48.495	56.199
16	-22742+353935i	-425976+374986i	56.359	59.711
17	-222224+405564i	-377142+411845i	64.580	65.580
18	-16061+442673i	-16210+442809i	70.489	70.511
19	-45413+464536i	-46926+461022i	73.971	73.411
20	-197046+536565i	-178586+505898i	85.440	80.557
21	-16625+564337i	-17117+563668i	89.863	89.756
22	-33871+648590i	-33879+653296i	103.279	104.028
23	-93487+665262i	-95205+657829i	105.933	104.750
24	235926+784996i	220180+770378i	124.999	122.672
25	256876+805660i	-19924+808350i	128.290	128.718
26	-20119+809179i	268839+833256i	128.850	132.684
27	-52146+890834i	-51097+889583i	141.853	141.653
28	-24958+910135i	-24524+910423i	144.926	144.972

29	-26842+1016365i	-26612+1016453i	161.842	161.856
30	-16350+1051480i	-16146+1051345i	167.433	167.412
31	-41087+1152820i	-40726+1152502i	183.570	183.519
32	-813275+1215919i	-680935+1796577i	193.618	286.079

4.2.2.2 Intershielded Disc Winding

An estimation example of the frequency response from an Intershielded Disc winding is plotted in Figure 4-8, and the estimated parameters are listed in Table 4-6. All the fuzzy areas are noted in grey. Although the Intershielded Disc winding has more obvious resonances compared with the Interleaved Disc winding, there still exist many small oscillations on its typical camel hump feature.

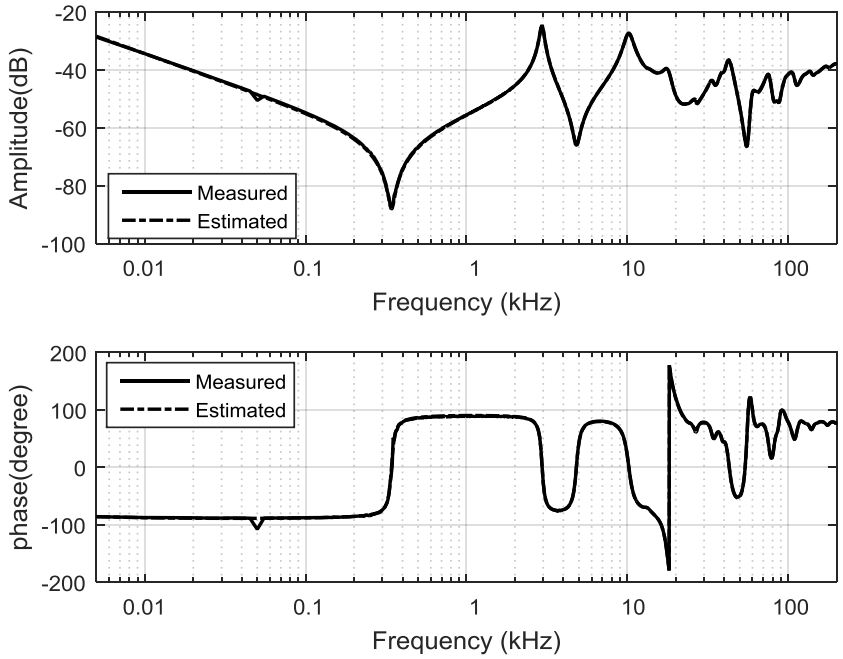


Figure 4-8 Comparison of FDTF Plot and Measured Data of Intershielded Disc Winding

Table 4-6 Estimated Parameters of Intershielded Disc Winding by Feature Extraction Method

	k=-8.772e10			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1		783683	-	-
2	-9363	-10154	-	-
3	-66+1238i	-10	0.200	-
4	-126+1996i	-63+1288i	0.320	0.205
5	-1308+33012i	-782+22996i	5.257	3.662
6	-4847+76961i	-4140+59665i	12.255	9.501
7	-6040+89640i	-4484+77967i	14.274	12.415
8	-18219+109304i	-5786+90688i	17.405	14.441
9	9440+118634i	-6257+105714i	18.891	16.833
10	-11069+132651i	-15857+127617i	21.123	20.321

11	-69820+160567i	-51111+148293i	25.568	23.614
12	1218061+171103i	-1203+215182i	27.246	34.265
13	-1480+215228i	-18775+224906i	34.272	35.813
14	-1801+244980i	-1672+244679i	39.010	38.962
15	-10866+293784i	-8364+301376i	46.781	47.990
16	-12519+308950i	-16304+355396i	49.196	56.592
17	-507891+362647i	-60248+377497i	57.746	60.111
18	-16187+365469i	-5936+440267i	58.196	70.106
19	-53686+380271i	-27274+474007i	60.553	75.479
20	-5998+440160i	-216040+485032i	70.089	77.234
21	-11941+501807i	-13679+502460i	79.906	80.010
22	-20261+516618i	-30854+587577i	82.264	93.563
23	-31345+582265i	12349+636948i	92.717	101.425
24	-31105+621561i	-55637+652470i	98.975	103.896
25	12667+636600i	-23333+658130i	101.369	104.798
26	-26509+658958i	-87302+681162i	104.930	108.465
27	-78634+694593i	-28535+742209i	110.604	118.186
28	-27214+754253i	-14435+804854i	120.104	128.161
29	-14292+805109i	-24418+875226i	128.202	139.367
30	-24070+875159i	-17586+932695i	139.357	148.518
31	-17738+932613i	-73446+951946i	148.505	151.584
32	-73959+949758i	-32258+996558i	151.235	158.688
33	-32347+1000076i	-46416+1221930i	159.248	194.575
34	-47065+1224111i	1217945+1234389i	194.922	196.559
35	-30968+995493i	-640418+2363117i	158.518	376.293
36	-43511+1225830i	-43235+1224033i	195.196	194.910
37	-52914+1198785i	958485+2013659i	190.889	320.646
38	2930+1301082i	5499+1300699i	207.179	207.118

4.2.2.3 Plain Disc Winding

An estimation example of the frequency response from a Plain Disc winding is plotted in Figure 4-9 and the estimated parameters are listed in Table 4-7. All the fuzzy areas are noted in grey. The noise, which appears as a small peak around 50 Hz is eliminated by using an appropriate order for the frequency band 0 to 1 kHz. Though the local deviation is noticeably, it is still within the acceptable range according to the criteria set.

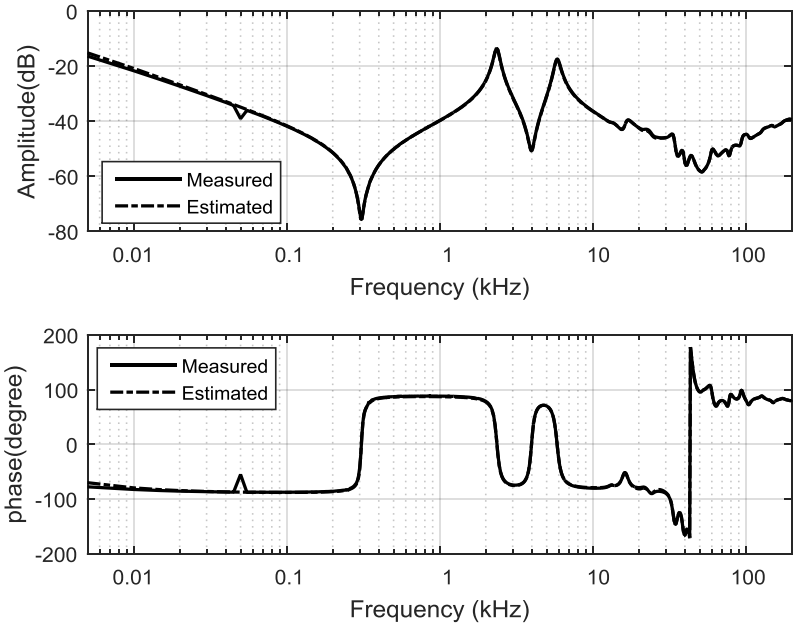


Figure 4-9 Comparison of FDTF plot and Measured Data of Plain Disc Winding

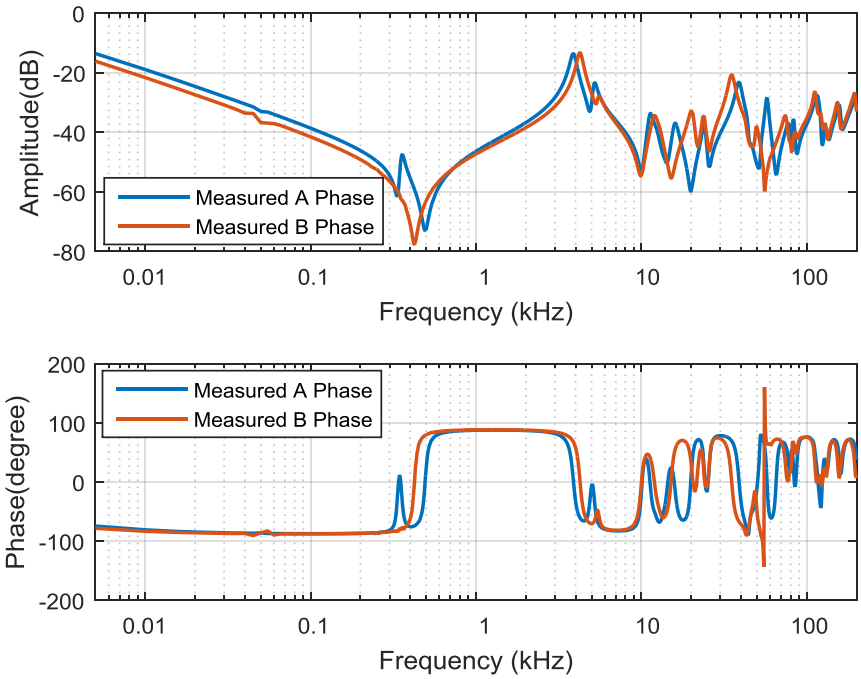
Table 4-7 Estimated Parameters of the Plain Disc Winding by Feature Extraction Method

	k=2.8002E-28			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	-723174	-24016	-	-
2	-22135	-9	-	-
3	-52+1936i	-604+14831i	0.310	2.360
4	-954+25001i	-1747+36703i	3.981	5.844
5	-4943+83212i	-4927+83623i	13.250	13.316
6	-5896+100351i	-6103+103604i	15.979	16.497
7	-10102+148314i	-8539+146831i	23.617	23.381
8	-59437+160490i	-66056+167415i	25.556	26.658
9	-8104+221491i	-8375+213405i	35.269	33.982
10	-7727+254217i	-9659+246309i	40.480	39.221
11	41897+279350i	-16096+275543i	44.482	43.876
12	-55113+315027i	-5824+380747i	50.164	60.629
13	-5475+381153i	-11610+381551i	60.693	60.757
14	-23270+382554i	-43629+396271i	60.916	63.100
15	-40951+398951i	-7995+400182i	63.527	63.723
16	-8670+400523i	-91840+454669i	63.778	72.400
17	-124585+453500i	-9927+478779i	72.213	76.239
18	-9266+480399i	-13318+495680i	76.497	78.930
19	-10509+494072i	-415617+531664i	78.674	84.660
20	-933428+502719i	-20603+589485i	80.051	93.867
21	-17685+582954i	-25542+656886i	92.827	104.600
22	-23884+661554i	-77060+730207i	105.343	116.275
23	-71678+726903i	-29542+798108i	115.749	127.087
24	-28612+795538i	-16021+889831i	126.678	141.693
25	-15876+889787i	-26628+947585i	141.686	150.889
26	-26620+949796i	-50982+1015226i	151.241	161.660
27	-49931+1013977i	-56560+1075183i	161.461	171.207
28	-55444+1074726i	-38466+1096961i	171.135	174.675
29	-38659+1095426i	-26975+1253996i	174.431	199.681

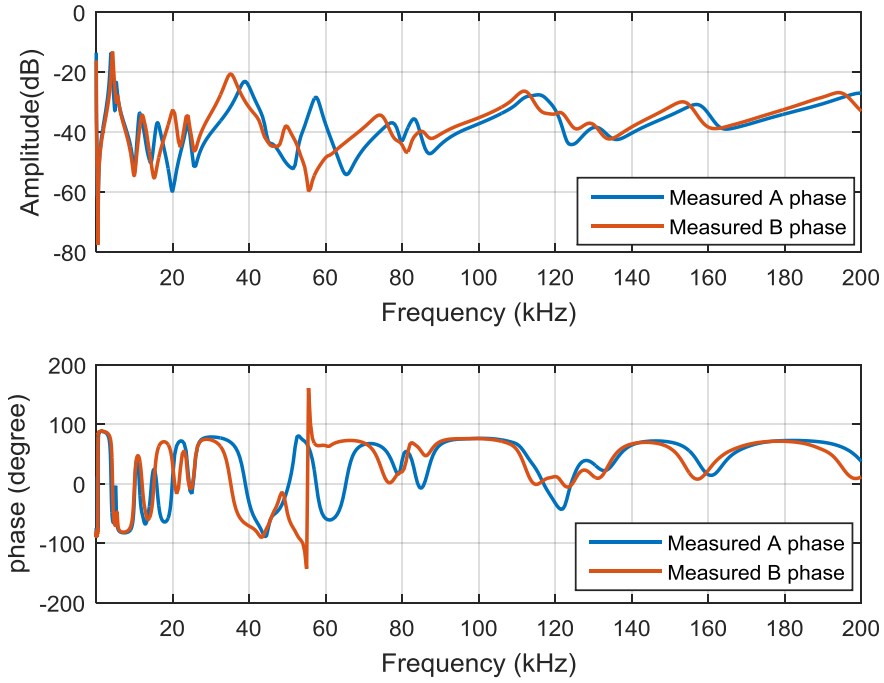
30	-25354+1254568i	-	199.772	-
31	3137759+2127287i	-	338.740	-

4.2.3 Application on Faulty Winding

The frequency responses measured on A phase and B phase, common winding, N to LV (Tap 19) of the 275/132/11 kV autotransformer T2305 are used to demonstrate the effectiveness of the transfer function estimation method. Axial collapse occurred on the tap winding of B phase. In Figure 4-10, the frequency responses of A phase and B phase from 5 Hz to 200 kHz given and they are different at the low frequency region, i.e., A phase has two troughs and B phase only has one, due to the asymmetry of the three-limb core. Attention should be paid to the mid- and high frequency regions where significant differences exist including the shifting of resonances and antiresonances; and new resonances also appear around 80 kHz and 120 kHz.



(a) Logarithmic



(b) Linear

Figure 4-10 Comparison of Frequency Responses from A Phase and B Phase N to LV T2305

The proposed Feature Extract Method is applied to the frequency responses in Figure 4-10. The estimated and the measured frequency responses of A phase are compared in Figure 4-11. The estimated and the measured frequency responses of B phase are compared in Figure 4-12. The parameters of their estimated transfer function are listed in Table 4-8 and Table 4-9 respectively. In Figure 4-14, the measured frequency responses from A phase and B phase are plotted together, while the estimated frequency responses for A phase and B phase are plotted together to compare if the difference between the two frequency responses can be reflected by the proposed estimation method .

It can be seen that both of the estimated frequency response traces are in good agreement with the measured results. The shifting of the resonant points can be reflected by the change in the estimated parameters. For example, the original resonance located at 39.0 kHz, as the same as that in frequency response of A phase, is shifted to 35.0 kHz in frequency response of B phase after the fault, which are circled in dashed line in Figure 4-13. Accordingly, the corresponding pair of complex poles changes from -9696±244344i to -8797±220568i. The difference in the frequency locations of resonances is reflected by the difference of the imaginary parts of the two pairs of complex poles. The higher the frequency, the larger is the imaginary part of the complex poles. The difference in the real parts is not important hence would not be discussed here. In general, the smaller the absolute value of the real part, the sharper the shape of the resonance would be.

Although the shifting of the resonant points can be reflected, the change in the number of resonant points is concealed by the proposed method, as mentioned before the number of estimated resonant points are not always the same as what is seen by naked eyes in the FRA plot, and sometimes more parameters than needed are used to describe the delicate features of the frequency response. The frequency response of B phase has more resonant points than that of A phase. For the frequency response from A phase, there are 36 pairs of complex zeros and 34 pairs of complex poles used for its expression. For the frequency response from B phase, there are 34 pairs of complex zeros and 33 pairs of complexes poles in its transfer function.

This method focuses more on a well matched result and the appropriate amount of parameters matched with the number of peaks and troughs sometimes becomes secondary. Therefore due to the uncertainty in the amount of key parameters, the difference in the amount of resonant points of A phase and B phase is not considered.

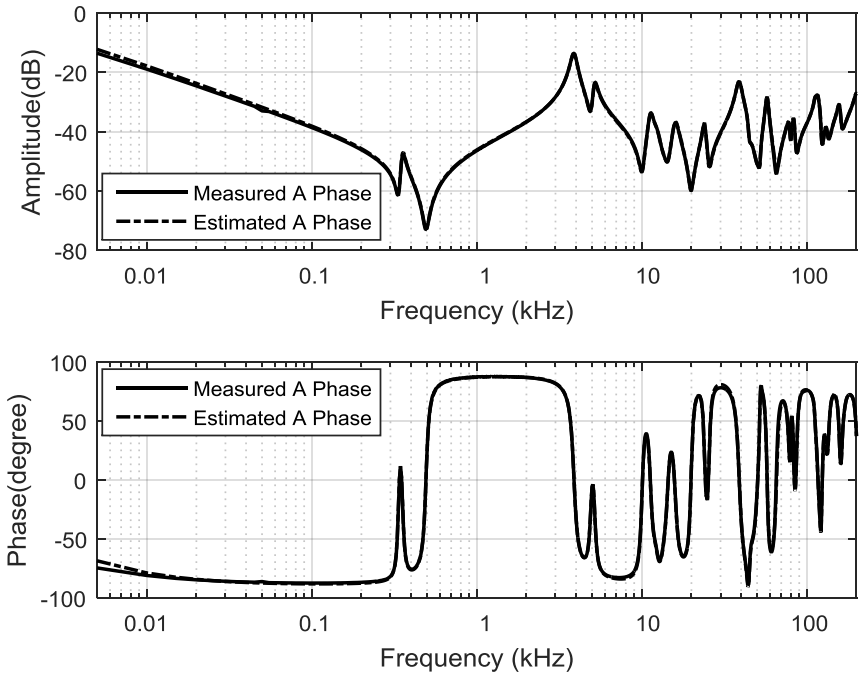


Figure 4-11 Comparison of Measured and Estimated Frequency Responses of T2305 N to LV A Phase

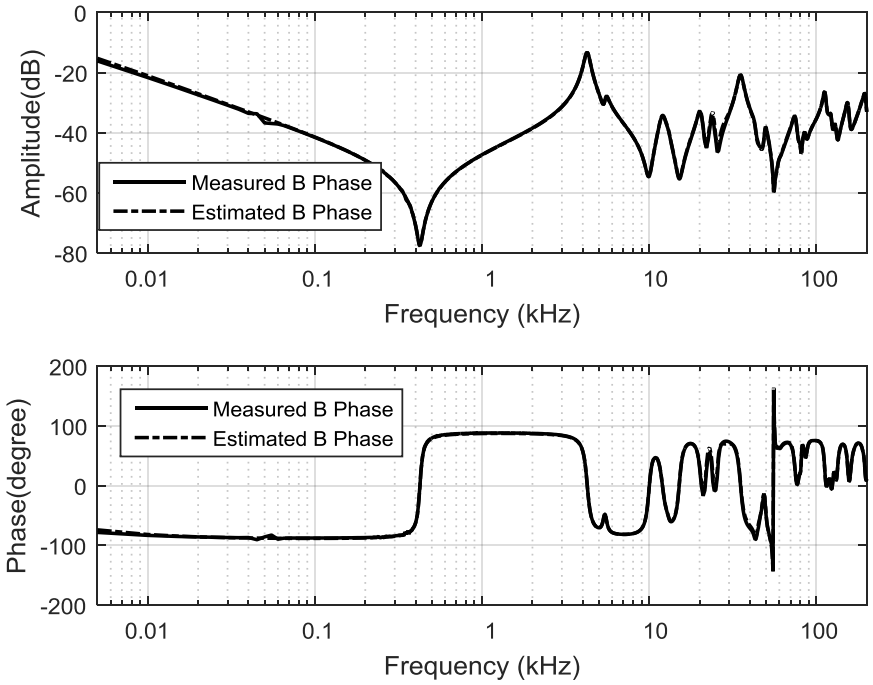


Figure 4-12 Comparison of Measured and Estimated Frequency Responses of T2305 N to LV B Phase

Table 4-8 Estimated Parameters of T2305 N to LV A Phase

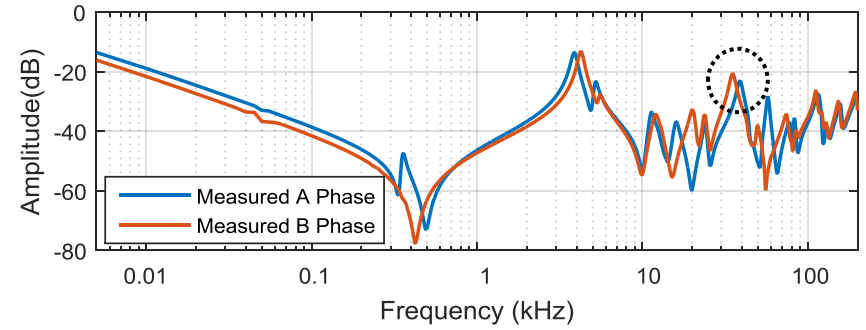
	k= 2.4835e-14			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	368222	-1580896	-	-
2	49155	1289698	-	-
3	-54+2107i	38881	0.335	-
4	-105+3094i	-12	0.492	-
5	-974+30923i	-60+2230i	4.922	0.355
6	104165+51929i	-1050+24346i	8.265	3.875
7	-1924+63076i	-1003+32461i	10.039	5.166
8	-4690+77703i	-2620+70842i	12.367	11.275
9	-6133+84915i	82767+76644i	13.515	12.198
10	-3762+90216i	-3563+78036i	14.358	12.420
11	-3373+125058i	-7024+86529i	19.904	13.772
12	-3915+160513i	-4343+100052i	25.546	15.924
13	-67465+189225i	-3609+151177i	30.116	24.060
14	-185240+194231i	-69081+189291i	30.913	30.127
15	-5286+281482i	-203838+202292i	44.799	32.196
16	-12138+316078i	-9696+244344i	50.305	38.889
17	-3814+325953i	-8092+278238i	51.877	44.283
18	-13234+357898i	-6989+327299i	56.961	52.091
19	-17787+396505i	-12424+356406i	63.106	56.724
20	-11634+412958i	-7161+361414i	65.724	57.521
21	-637745+476956i	-19114+397046i	75.910	63.192
22	-6369+501016i	366633+435426i	79.739	69.300
23	-72741+521043i	-11553+493015i	82.927	78.466
24	-10556+542863i	-62727+510849i	86.399	81.304

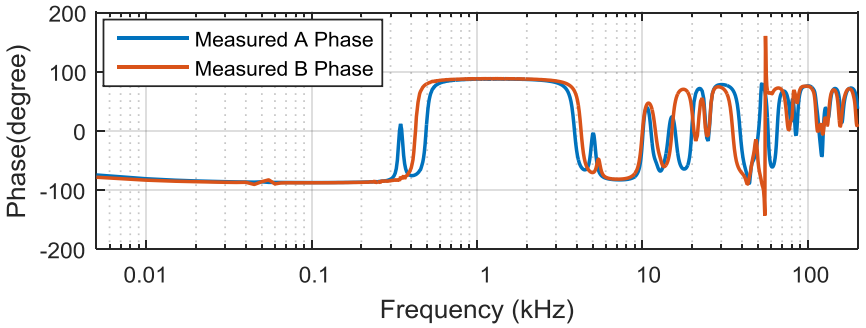
25	421265+550818i	-10335+523691i	87.665	83.348
26	-15138+558274i	-16362+557221i	88.852	88.684
27	-15439+706272i	-12938+705218i	112.407	112.239
28	7739+727525i	-802029+709859i	115.789	112.978
29	-11836+765550i	7434+727818i	121.841	115.836
30	-16133+778452i	-22931+733012i	123.895	116.662
31	-22943+848113i	-9714+761542i	134.981	121.203
32	-38616+1015362i	-24561+824891i	161.600	131.286
33	-23455+1024978i	-23533+993160i	163.130	158.066
34	-37304+1044750i	-38977+1016856i	166.277	161.838
35	-117586+1122257i	-31977+1040379i	178.613	165.581
36	-17518+1264473i	-103500+1129952i	201.247	179.837
37	-107671+1265546i	-52627+1254433i	201.418	199.649
38	-81581+1983203i	-15760+1261551i	315.637	200.782

Table 4-9 Estimated Parameters of T2305 N to LV B Phase

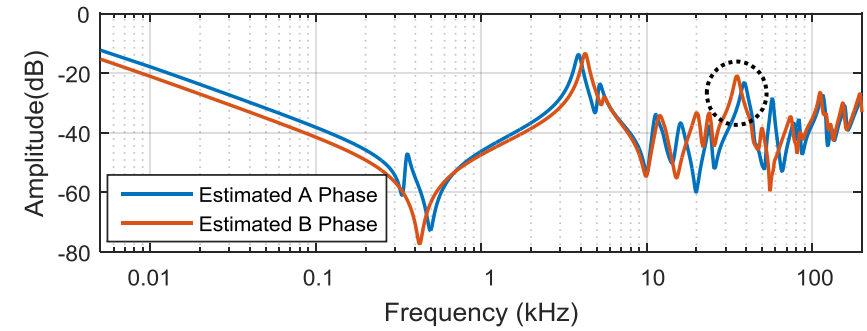
	k= -1.8279e-14			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	-379871	-653552	-	-
2	199184	-9	-	-
3	-85+2667i	-1085+26695i	0.424	4.249
4	-1029+33785i	-1053+34372i	5.377	5.471
5	-2447+62489i	-3925+75166i	9.945	11.963
6	-3346+82140i	-2884+82090i	13.073	13.065
7	-3759+94973i	33720+120964i	15.115	19.252
8	33234+120200i	-5393+127836i	19.130	20.346
9	-3296+137177i	183187+143976i	21.832	22.914
10	-4205+159923i	-4349+150507i	25.453	23.954
11	-20101+175335i	-16042+171434i	27.905	27.285
12	-6523+275575i	-8797+220568i	43.859	35.105
13	167309+287335i	-7898+272589i	45.731	43.384
14	-10071+300542i	-7306+310980i	47.833	49.494
15	3292+349347i	238844+344282i	55.6	54.794
16	-24924+360952i	-11385+348628i	57.447	55.486
17	703877+367410i	-4597+369412i	58.475	58.794
18	-4723+368353i	31340+437935i	58.625	69.699
19	32771+440499i	-14395+470014i	70.108	74.805
20	-15189+489961i	-10193+506709i	77.98	80.645
21	-5533+509613i	-14200+537036i	81.107	85.472
22	-13854+545535i	-36130+633821i	86.825	100.876
23	-35224+632484i	-119622+656457i	100.663	104.478
24	-96780+679918i	-16142+705851i	108.212	112.34
25	-21317+736512i	-11290+743663i	117.22	118.358
26	-13618+746137i	-15144+765616i	118.751	121.852
27	-17580+784654i	-19868+814712i	124.882	129.665
28	-20030+837099i	381119+821385i	133.229	130.727

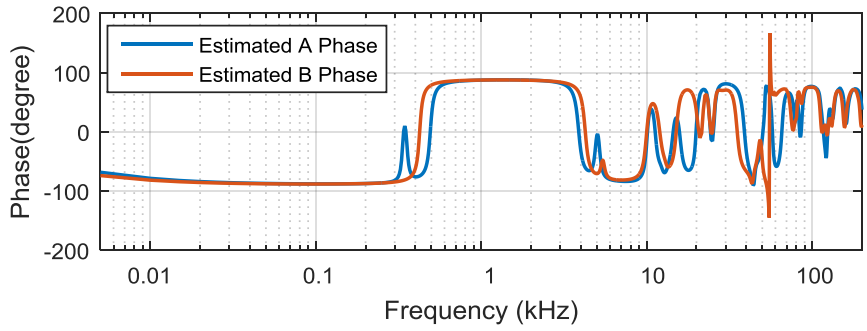
29	420300+839718i	10531+857658i	133.645	136.501
30	10459+857340i	-22175+970609i	136.45	154.477
31	-79769+981924i	-72817+980045i	156.278	155.979
32	-26827+1000474i	-77512+1064418i	159.23	169.407
33	-66163+1061786i	-26242+1189723i	168.988	189.35
34	-26237+1192105i	-25091+1226938i	189.729	195.273
35	-38313+1261568i	13624+1368608i	200.785	217.821
36	1288+1358302i	-	216.181	-





(a) Measured





(b) Estimated

Figure 4-13 Comparison of Measured and Estimated Frequency Responses of T2305 N to LV B Phase

4.2.4 Summary

The proposed method estimates the transfer function of a frequency response over a large frequency range by extracting key information, the complex zeros and poles, from several frequency regions to construct a Feature Transfer Function (FTF), and then correct FTF with a Difference Transfer Function (DTF).

The method is accurate and efficient, and the frequency responses can be processed automatically by the developed MATLAB program. Compared with other function fitting methods it has a better performance to reflect subtle features.

This method has been tested on eight 400/275/13 kV auto-transformers. For each transformer, six groups of frequency responses from three phases, two windings (series and common windings) are used, which are 48 groups of data all together. All of the estimation results match well with the measured frequency responses.

However when the method is tested on the frequency responses on the tertiary windings, not all of them can find good match. This is because the exact values set in cancellation criteria are related to transformer characteristics, which means the set criteria suitable for the series and common windings are not suitable for tertiary windings.

This method can be used for transformer winding diagnosis by comparing the parameters of transfer functions representing the reference measurement data and the diagnostic measurement data. If the parameters are changed, a mechanical fault may be indicated on the winding. In addition, the transfer functions obtained can be used to help design black box models of transformers, if the excitation and terminal connections are properly taken in account. The advantage of such models is that they are capable to represent the transformers with accuracy for a wider frequency range.

4.3 Extreme Points Identification Algorithm

The Extreme Points Identification Algorithm is proposed in this section for the transfer function estimation of frequency response. In this method, the extreme points on FRA magnitude trace are used for the computation of transfer function parameters, using their magnitudes and corresponding frequency locations. The proposed algorithm is applied both on artificially constructed and measured FRA traces.

4.3.1 Methodology

The understanding of the effects of real/complex zeros/poles on the frequency response is the foundation to develop the proposed algorithm. After the introduction of fundamental understanding, an artificially constructed FRA trace is used as an example to explain the procedure of the proposed estimation algorithm with the help of a flow chart. Only the magnitude response is needed for the proposed methodology.

4.3.1.1 Fundamental Understanding of Poles and Zeros

As mentioned before, for a physically achievable system, the real zeros and poles of the expected transfer function should be negative, whilst the real part of a pair of complex zeros/poles are supposed to be negative too.

The height and location of the resonance and antiresonance on the magnitude FRA trace are decided by the real and imaginary part of complex zeros and poles. Normally the imaginary part is dominant and hence it can be used to approximate the location of resonant point as show in Equation 4-2. However the real part definitely has some considerable effects and it would be taken into consideration to help improve the precision of the calculated parameters of the transfer function.

The transfer function of a single pair of complex poles, σ+ωi and σ-ωi, can be expressed as:

$$T_s = \frac{k}{(s-(\sigma+\omega i))(s-(\sigma-\omega i))} \qquad \text{Equation 4-9}$$

where k is the constant controlling the overall magnitude of the trace, σ mainly controls the sharpness of the resonance, and ω dominants the location of resonance. Normally ω is about dozens of times of σ.

Examples of complex poles' Bode plots are given in Figure 4-14 with different value of σ, ω and k. For each trace, its magnitude remains as a constant before its peak, and drops at 40dB per decade after its peak. When k changes, the trace shifts down or up without changing its shape, as shown by the two traces H1 and H2. For convenience, k is defined to fix the starting magnitude as 0 dB:

$$k = \sigma^2 + \omega^2 \qquad \text{Equation 4-10}$$

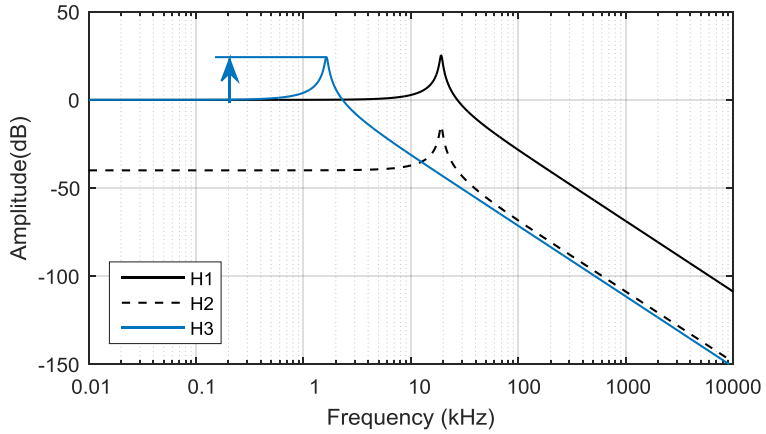


Figure 4-14 Examples of Different Pairs of Complex Poles

If ω varies while k is defined using Equation 4-10, the peak shifts to the left or right accordingly. When ω decreases, the peak shifts to the left, as shown by traces H1 and H3. When frequency f changes, the magnitude m of a pair of complex poles can be calculated according to Equation 4-11:

$$m = \frac{\sigma^2+\omega^2}{\sqrt{16\sigma^2\pi^2f^2+((\sigma^2+\omega^2)-4\pi^2f^2)^2}} \quad \text{Equation 4-11}$$

By setting the derivative of the magnitude m to zero, it can be found m reaches the maximum amplitude at the frequency point:

$$f = \frac{\sqrt{\omega^2-\sigma^2}}{2\pi}(\omega > \sigma) \quad \text{Equation 4-12}$$

This means the location of resonance is decided by both the real and imaginary parts. The real part normally has a smaller influence because it has much smaller value than the imaginary part.

For a specific FRA trace, of which the transfer function contains only one pair of complex poles, once the magnitude m_0 and its corresponding frequency f_0 of its resonance are known, σ and ω could be computed using the following equations:

$$\sigma = -\sqrt{\frac{2\pi^2{f_0}^2m_0}{\sqrt{{m_0}^2-1}} - 2\pi^2{f_0}^2} \quad \text{Equation 4-13}$$

$$\omega = \sqrt{\frac{4\pi^2{f_0}^2m_0}{\sqrt{{m_0}^2-1}} - \sigma^2} \quad \text{Equation 4-14}$$

The standardised transfer function of a real pole is defined as:

$$T_s = \frac{p}{s-p} \quad \text{Equation 4-15}$$

For a real pole p, there exists no resonance in its magnitude trace. Its magnitude trace starts from 0 dB and goes horizontally before frequency point p/2π, smoothly bends around this frequency point, and drops at 20 dB per decade after p/2π Hz.

The value of p can be calculated by the frequency *f'* and magnitude *m'* of any point on slope part of its magnitude trace:

$$p = -\sqrt{\frac{2\pi f' m'}{\sqrt{1-m'^2}}} \qquad \text{Equation 4-16}$$

When the resonance is located beyond the estimated frequency range, which means the location and magnitude of extreme point are unknown, an alternative solution can be used to roughly estimate this pair of complex poles. The parameters can be calculated using its induced magnitude m_1 and corresponding frequency f_1 at an arbitrary frequency point within the measured frequency range, which is defined as the higher boundary of estimated frequency range in this study, and a manually defined resonance frequency f_0 outside the estimated frequency range:

$$c = \frac{f_1}{f_0} \qquad \text{Equation 4-17}$$

$$\sigma = -\sqrt{\frac{2\pi^2 {f_1}^2 m_1 \sqrt{2c-1}}{\sqrt{{m_1}^2-1}} - 2\pi^2 {f_1}^2 c} \qquad \text{Equation 4-18}$$

$$\omega = \sqrt{\frac{4\pi^2 {f_1}^2 m_1 \sqrt{2c-1}}{\sqrt{{m_1}^2-1}} - \sigma^2} \qquad \text{Equation 4-19}$$

f_0 is normally set as several times of the higher boundary of the estimated frequency range. By this way of estimation, the measured and estimate magnitudes of higher boundary of the measured frequency range can be matched.

Both the magnitude plots of a pair of complex zeros and a real zero own horizontal 0 dB starting magnitude, in the predefined standardised format similar to the complex and real poles. A pair of complex zeros' magnitude owns an antiresonance, which behaves as a trough, and rises at 40 dB per decade after the antiresonance. The magnitude trace of real zero, *z*, rises at 20 dB per decade after $z/2\pi$ Hz. The values of a real zero and a pair of complex zeros can be calculated in similar ways as a real pole and a pair of complex poles stated above.

4.3.1.2 Estimation Procedure

The flowchart in Figure 4-15 illustrates the prediction process of the proposed method. The parameters are initialised firstly. Then iterations are carried out to optimize the parameters by eliminating the mutual influence between zeros and poles, and recalculate the locations of the resonant points. Iteration process would stop if the number of iteration is reached or the locations of extreme points are updated adequately to meet the correct precision. Lastly, the parameters located out of the concerned frequency range are estimated to give a

better match of the higher frequency boundary.

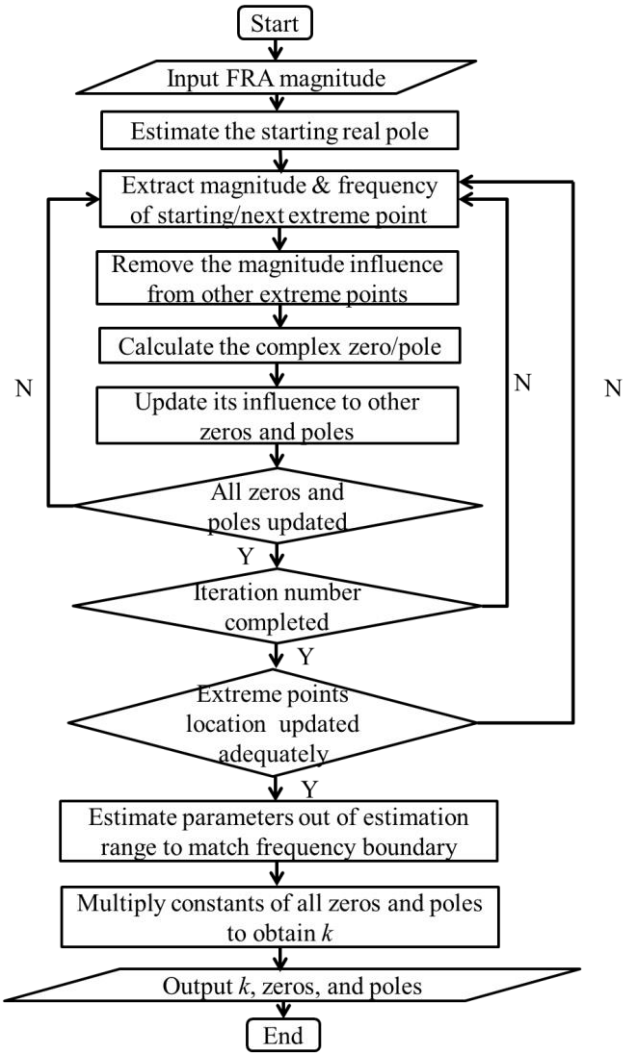


Figure 4-15 Flow Chart of Extreme Points Identification Algorithm

An artificially constructed frequency response given in Figure 4-16 is used as an example to explain the estimation procedure.

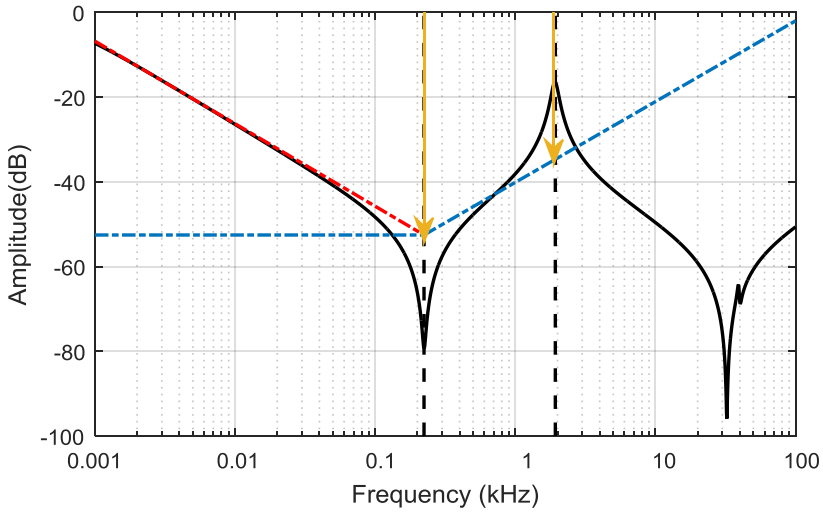


Figure 4-16 Example of Artificial FRA Trace

A. Parameters Initialization

The first step of parameter initialisation is to initialise a real pole. The starting part of the trace in Figure 4-16, before 0.1 kHz, is dominated by the slope part of a real pole. This pole could be identified by Equation 4-16.

Then all local minimum and maximum magnitude points, e.g. the two peaks and three troughs in Figure 4-16, which are corresponding to two pairs of complex poles and three pairs of complex zeros, are identified. In fact, the magnitude of resonance or antiresonance point consists of two parts, one of which is the pole's or zero's own magnitude, and the other is the influence from other zeros and poles. The trough/peak height of a pair of complex zeros/ poles can be obtained by removing the induced magnitude by other zeros and poles from the total magnitude. Once the magnitude and corresponding frequency are known, according to Equation 4-13 and Equation 4-14, the parameters can be solved. In the stage of parameter initialisation, the influence from the uninitialised parameters to the initialised parameters is viewed as 0.

In Figure 4-16, the red dashed slope is the latter part of the real pole's Bode plot, which drops 20 dB per decade. The blue dash line is the asymptotic Bode plot of the first pair of complex zeros, of which the antiresonance is located at about 0.2 kHz. Considering the fact that the magnitude trace of a pair of complex zeros or poles either rises or drops at 40 dB per decade after the resonance or antiresonance, the combined influence from the lower frequency range to higher frequency range causes the magnitude response either rises or drops at 20 dB per decade roughly. Therefore, the magnitude of the red dashed at about 0.2 kHz and the magnitude of blue dashed line at about 2 kHz can be viewed as the influence from the lower frequency range, which are represented by two yellow arrows in Figure 4-16. The first pair of complex poles' peak magnitude could be approximated as the residue after removing the influence from the real pole and the first pair of complex zeros, which is about from -80 dB to -50 dB. Such prediction process is repeated until all of the following complex zeros and poles are initialised.

B. Parameters Iteration

As mentioned, since the parameters of the transfer function are sequentially processed, the effect from poles/zeros at higher frequency to the ones at lower frequency is not considered at parameter initialisation stage.

Only through iteration the correct parameters of the transfer function can be finalised. After the parameter initialisation stage, in the following iterations, the influence from higher frequency can be approximated using the parameters at higher frequency computed from the former iteration. The parameters are updated, with the mutual influence updated iteration by iteration. An accurate transfer function expression can be approached. Only a couple of iterations are needed in our experience.

C. Reallocation

A problem may appear when the resonance and antiresonance are located too close. For example, when a pair of complex zeros is very close to a pair of complex poles, the magnitude of points around the resonance caused by complex poles will drop due to the influence of the antiresonance caused by complex zeros. If the peak point drops more than its neighbor point, the original magnitude difference between them may be compensated. As a result, the point with the highest magnitude will change, and the resonance shifts. In order to solve this problem, the estimated traces are compared with the measured traces. When the locations of resonance or antiresonance do not match, the points which are used as the resonance or antiresonance points in the former iteration will shift accordingly. For instance, when the estimated resonance frequency is higher than the measured resonance frequency, the point used to calculate the poles will be shifted to a lower position by one data point in the next iteration.

Actually, due to the resolution limitation, it is found that instead of matching a specific point, it is better to match the region between two measurement points, which contains the actual resonance or antiresonance.

In the developed program, 20 iterations would be sufficient to eliminate the mutual influence between complex parameters. Instead of setting a termination condition, a total iteration number is set for simplification. Altogether 100 iterations are used, considering the fact normally the actual location of a resonance point will not shift more than 5 data bits away from the local extreme point. This iteration number satisfies the needs when dealing with FRA data currently investigated. The iteration number could be changed according to different needs.

Finally, the constant k is achieved by multiplying each constant coefficient from every zero and pole together.

D. Parameters Out Of Concerned Range

Parameters may exist with resonance/antiresonance lying outside the higher boundary of the concerned frequency range. Nevertheless, they may still have an influence on the magnitude within the concerned frequency range, which means their magnitude rises from 0 dB before the extreme points. If the last parameter within the measurement frequency range is a pair of complex zeros, a pair of complex poles can be added outside the frequency range, vise verse. The added pair of complex poles outside the frequency range can be calculated using a manually defined resonance frequency, and the magnitude and frequency of an arbitrary point on the frequency response, according to Equation 4-17 to Equation 4-19. By adding this pair of complex poles, the boundary frequency point's magnitude of the measured FRA can be matched with the estimated transfer function.

Another method is to use the measured resonant point out of the concerned range, when the measured data is available, to achieve a better estimation result. One more resonant point out of the frequency range is used when applying the Extreme Points Identification Algorithm in this section.

4.3.2 Application on Artificial Data

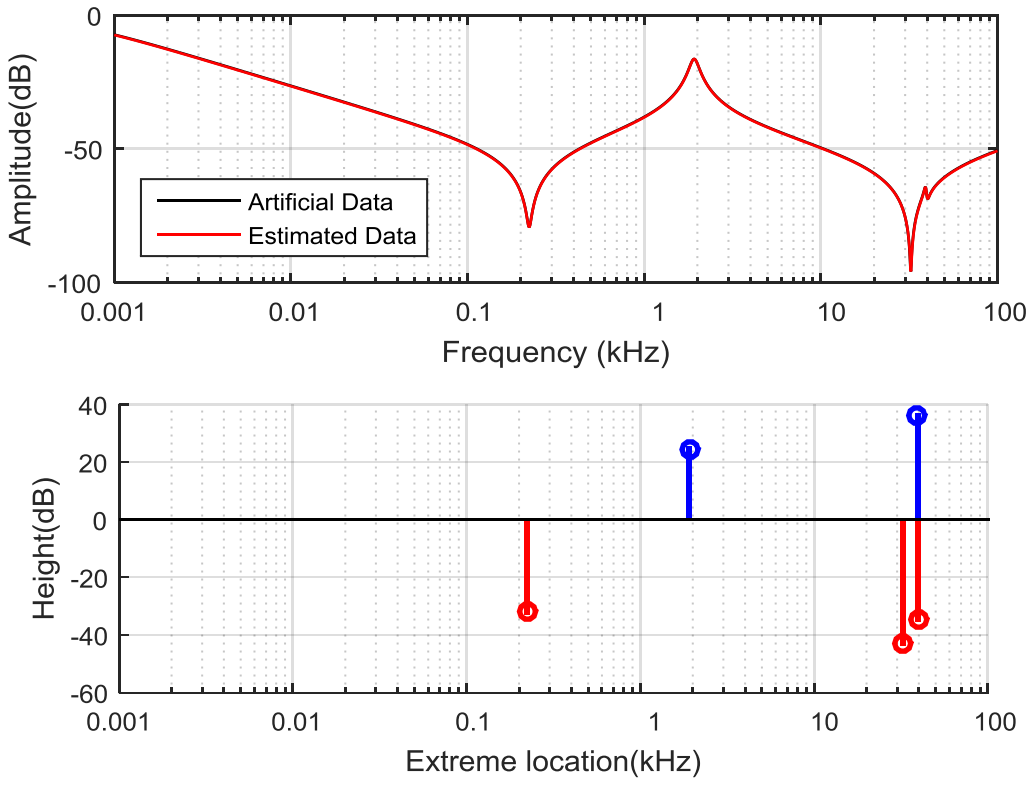
The artificial constructed frequency response plotted in Figure 4-16 consists of 2000 data points, which are logarithmically evenly distributed from 1 Hz to 100 kHz. The comparison of the artificially constructed frequency response and its estimated transfer function is shown in Figure 4-17, which illustrates a satisfactory result.

The original parameters, which are used to artificially construct the frequency response, and the estimated parameters are listed in Table 4-10. Their percentage difference in the real and imaginary parts of parameters are calculated and shown in Table 4-11. It can be seen that for the imaginary part, the percentage difference is very small, while the difference in the real part is comparatively larger. This is corresponding to the fact that the imaginary part is more dominant than the real part. Four numerical indices are used to describe the similarity between the original and estimated data in Table 4-12. The average Absolute Magnitude Difference is very small, only 0.1503 dB, with a Standard Deviation of 0.0574 dB. Among all the 2000 data points to be compared, the Maximum Absolute Magnitude Difference is as small as 1.0937 dB, suggesting a satisfactory result in terms of local deviation. Besides, the Correlation Coefficient shows a good match in the overall

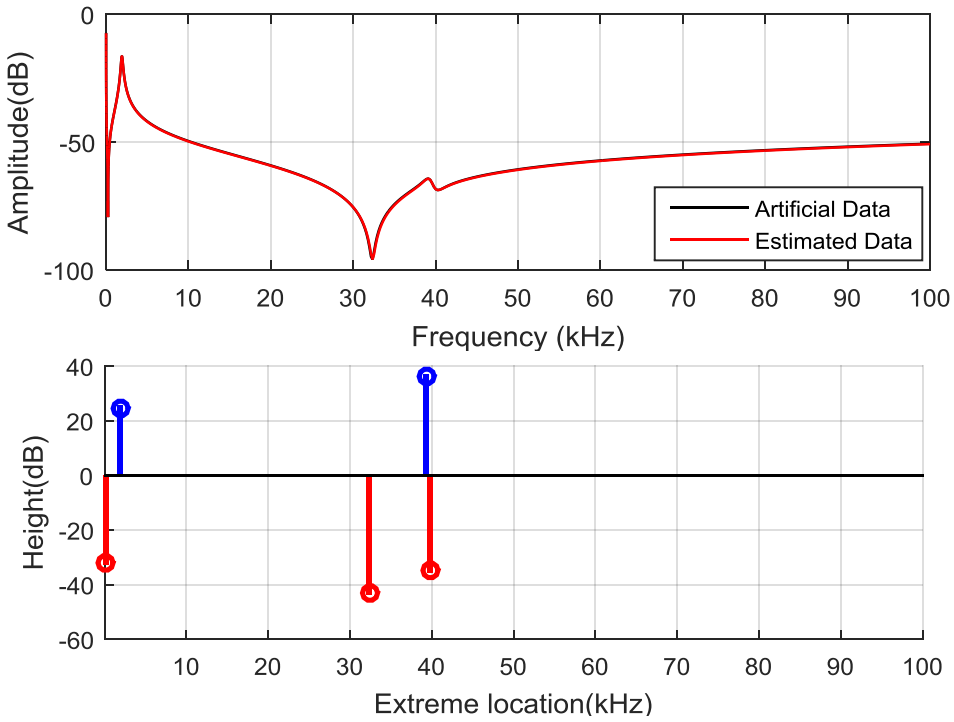
shape between the original and estimated data, since the closer the Correlation Coefficient to 1, the higher similarity there exists.

The estimated complex zeros and poles are also plotted in Figure 4-17, in both logarithmic and linear frequency scales, corresponding to the plotted frequency responses. The poles are plotted below the x axis in red line while the zeros are plotted above the x axis in blue line. The x axis is the location of the complex parameters, as shown in Figure 4-17, and the y axis is the peak/trough height of the complex parameters as shown in Equation 4-13.

Noticeably, sensitivity study has shown that the data resolution has an influence on the final result estimated. The higher resolution leads to a more accurate estimation result. However, since the measured data have a fixed resolution, it is hard to further improve in this aspect unless we use interpolation to increase the resolution.



(a) Logarithmic



(b) Linear

Figure 4-17 Estimation of Artificial Frequency Responses

Table 4-10 Estimated Transfer Function Parameters of Artificial Frequency Response

Original Parameters		**Estimated Parameters**	
k=5.2460e-09		k=5.1484e-09	
Zeros	**Poles**	**Zeros**	**Poles**
-35+1403i	-3	-36+1403i	-3
-1381+202869i	-721+12028i	-1450+203210i	-709+12039i
-4588+250217i	-4063+247247i	-4637+250065i	-3931+247191i

Table 4-11 Comparison of Estimated and Original Parameters of Artificial Frequency Response

Constant Coefficient Difference			
1.86%			
Zeros		**Poles**	
Difference of real part	**Difference of imaginary part**	**Difference of real part**	**Difference of imaginary part**
2.9%	0.0%	0.0%	-
5.0%	0.2%	1.7%	0.1%
1.1%	0.1%	3.2 %	0.0%

Table 4-12 Numerical Indices Comparing Original and Estimated Frequency Responses

Average absolute magnitude difference	0.1503 dB
Standard deviation of magnitude	0.0574 dB
Maximum absolute magnitude difference	1.0937 dB
Correlation Coefficient of magnitude	1.0000

4.3.3 Application on Measured Data

The proposed method is tested both on the windings with high and low series capacitance, i.e., the Multiple Layer winding and Plain Disc winding. All FRA traces are supposed to start with a real pole, and the real pole is followed by sequentially appeared pair of complex poles and pair of complex zeros. Examples of the use of the developed algorithm to estimate the transfer functions for FRA plots from 400/275/13kV auto-transformers are given in this section.

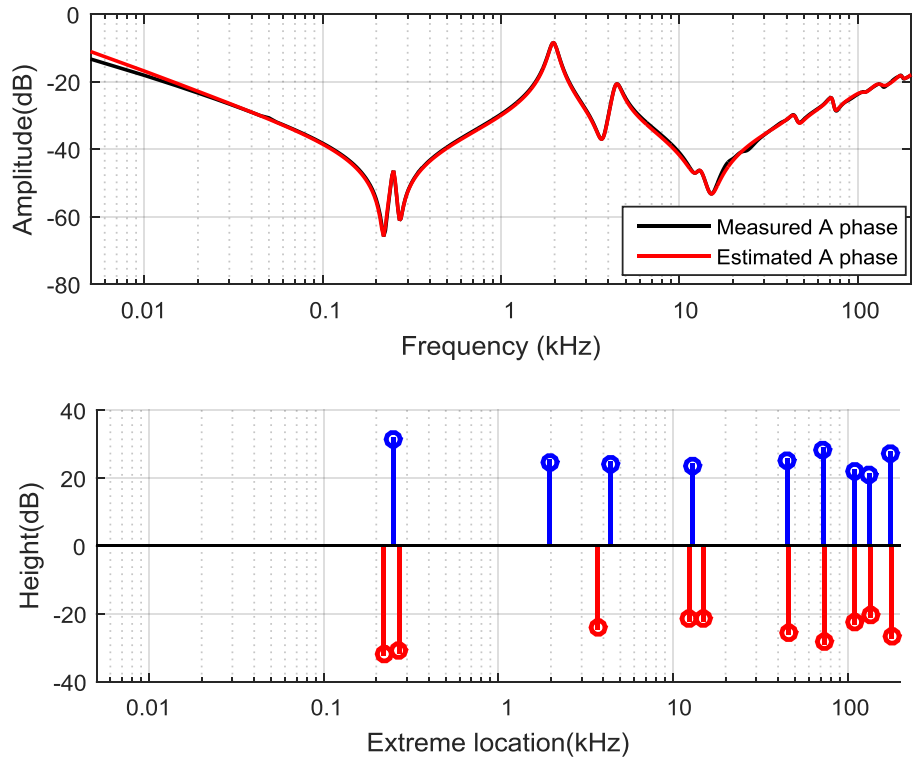
A. High Series Capacitance Winding Type

1) Application example

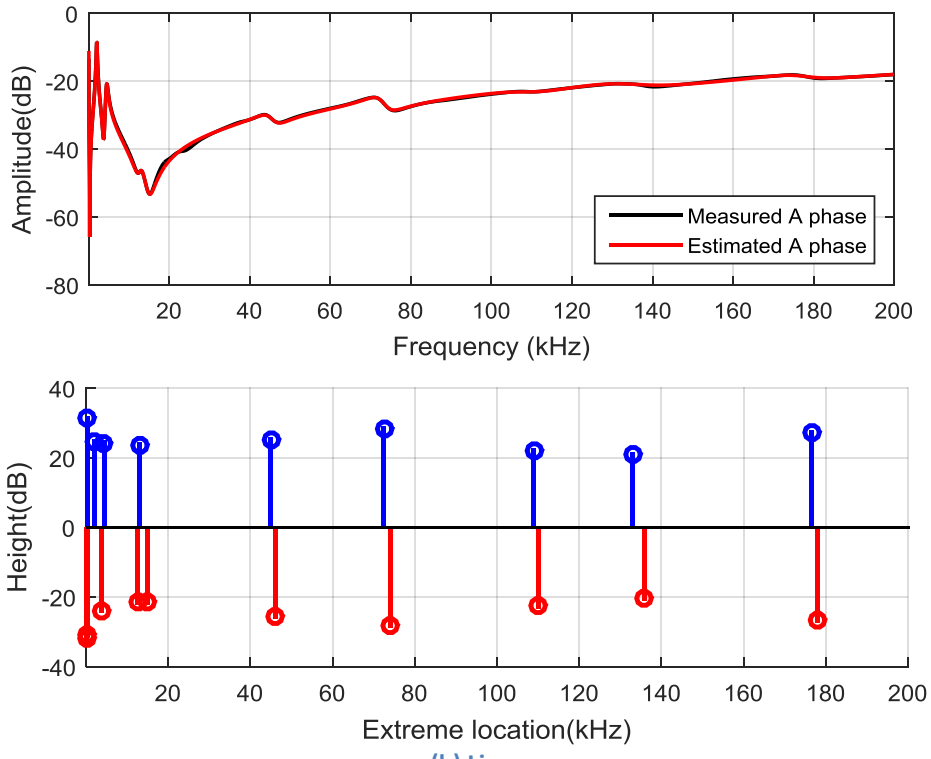
Multiple Layer windings have high series capacitance. In the frequency region which is dominated by winding properties, approximately from 20 kHz to a hundreds kHz, the magnitude normally rises at 20 dB per decade with fluctuations. The proposed algorithm is applied on the frequency response in Figure 4-1. Figure 4-18 shows the measured and the estimated frequency response.

The estimated transfer function is in the format of Equation 4-7, without any real zeros. The parameters of estimated transfer function, zc and pc, are plotted in Figure 4-18 as well. The estimated parameters are listed in Table 4-13. The pair of complex poles outside the measured frequency range is not shown in Figure 4-18 to give a clearer view.

Altogether 10 pairs of complex zeros, 10 pairs of complex poles and 1 real pole are used to describe this frequency response, which can significantly reduce the quantity of stored data, from a few thousands FRA measurement data points to a few dozens of transfer function parameters. The negative real pole, -9, is related to the low frequency range controlled by the transformer core. This real pole can be influenced by whether demagnetization is carried out before FRA measurement, thus it may not be as important as the complex parameters. All real parts of the complex zeros and complex poles are forced to be negative, which correspond to the resistance of the electrical winding circuit. All zeros and poles are forced to appear in sequence, because the fact that between two local maximum there must exist a local minimum and between two local minimum there must exist a local maximum.



(a) Logarithmic



(b) Linear

Figure 4-18 Comparison of Measured and Estimated Multiple Layer Winding Frequency Response

Table 4-13 Estimated Parameters of Multiple Layer Winding by Extreme Points Identification Method

	k= 2.8727e+05			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	-	-9	-	-
2	-36+1383i	-44+1571i	0.220	0.250
3	-47+1697i	-722+12368i	0.270	1.965
4	-1493+23296i	-1719+27699i	3.700	4.400
5	-6594+78503i	-5434+81862i	12.450	13.000
6	-7811+94258i	-15815+283185i	14.950	45.000
7	-15335+289433i	-17669+455873i	46.000	72.500
8	-17923+465301i	-55751+687133i	74.000	109.000
9	-51803+693089i	-77460+839246i	110.000	133.000
10	-81742+858414i	-48724+1110052i	136.000	176.500
11	-51323+1119584i	-814725+1577551i	178.000	215.000

Table 4-14 lists four numerical indices which evaluate how well the estimated transfer function matches the measured frequency response. The Average Absolute Magnitude Difference is 0.3390 dB. The Correlation Coefficient is 0.9995, which suggests a good match between the estimated and measured data as well.

Although most of the complex zeros and poles are identified, some closely located zeros and poles such as the pair of complex poles and the pair of complex zeros around 20 kHz, do not cause extreme values on the magnitude trace and are ignored deliberately to simplify the processing. This unavoidably contributes to the average difference and especially local difference between both magnitude and phase responses. Unlike the dominant zeros and poles which cause local extreme values on the FRA trace, such kind of recessive zeros and poles could be identified by the change in the rising or decreasing rate of magnitude response. Ideally when the acceleration of magnitude response changes its sign, a pair of complex zeros or a pair of complex poles is indicated. Nevertheless, due to the resolution limitation and noise in the measured data, it still remains difficult to accurately identifying recessive complex zeros and poles.

Table 4-14 Numerical Indices Comparing Measured and Estimated FRA Traces

Average absolute magnitude difference	0.3390 dB
Standard deviation of magnitude	0.3189 dB
Maximum absolute magnitude difference	2.2194 dB
Correlation Coefficient	0.9995

2) Comparison with other methods

In Figure 4-19, the estimated result by the proposed method is compared with the

estimations produced by other algorithms, i.e., Vector Fitting Method and Feature Extraction Method. The measured frequency response is also plotted as the reference. A high level of similarity can be seen visually between the measured data and the estimation results by the above 3 methods.

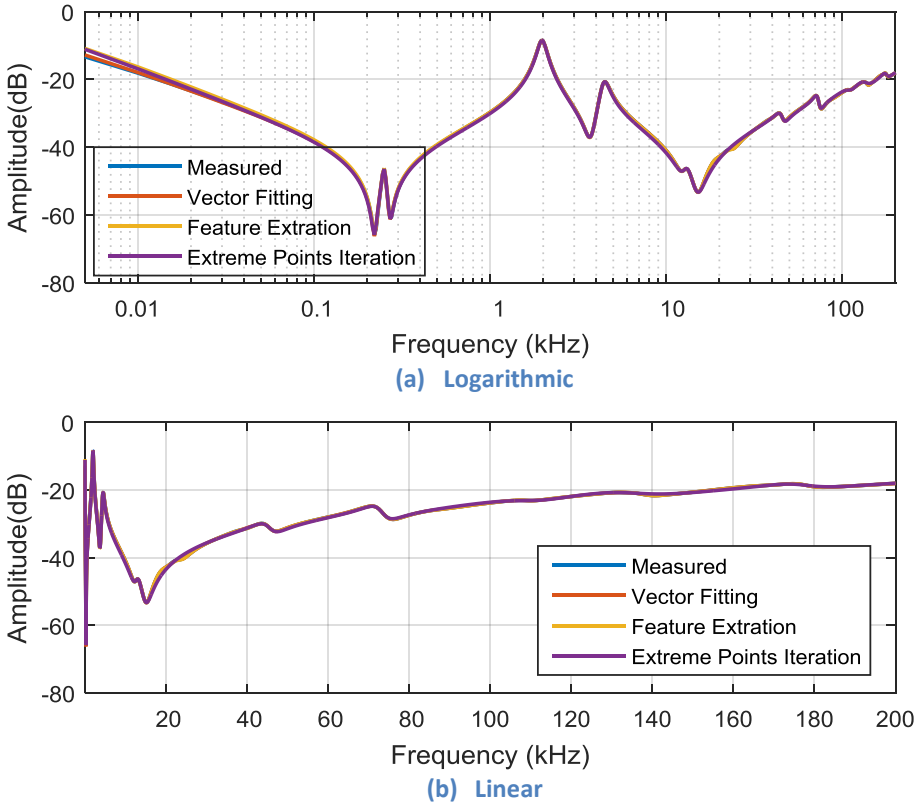


Figure 4-19 Comparison of Estimated Frequency Responses by Different Methods

Figure 4-20 plots the complex parameters estimated by the Vector Fitting Method without the parameters which are out of the measurement frequency range, as well the corresponding estimated frequency response. The real poles, real zeros and constant are listed in Table 4-15.

Figure 4-21 plots the complex zeros and poles estimated by the Feature Extraction Method in combination with the estimated frequency response. All the estimated parameters are shown in Table 4-16. The pairs of complex zeros and the pair of complex poles around 153 kHz are with positive real part, nevertheless they can roughly cancel with each other.

The numerical indices comparing the measured and estimated FRA traces using Vector Fitting and Feature Extraction are listed in Table 4-17. The Correlation Coefficients between the measured data and the estimated data by Vector Fitting Method, Feature Extraction Method and Extreme Points Iteration Method are 1.0000, 1.0000 and 0.9994 separately. The newly proposed method has the smallest correlation coefficient, which is understandable as the recessive zeros and poles are neglected. The Extreme Points

Identification method performs well globally regardless of acceptable local variation.

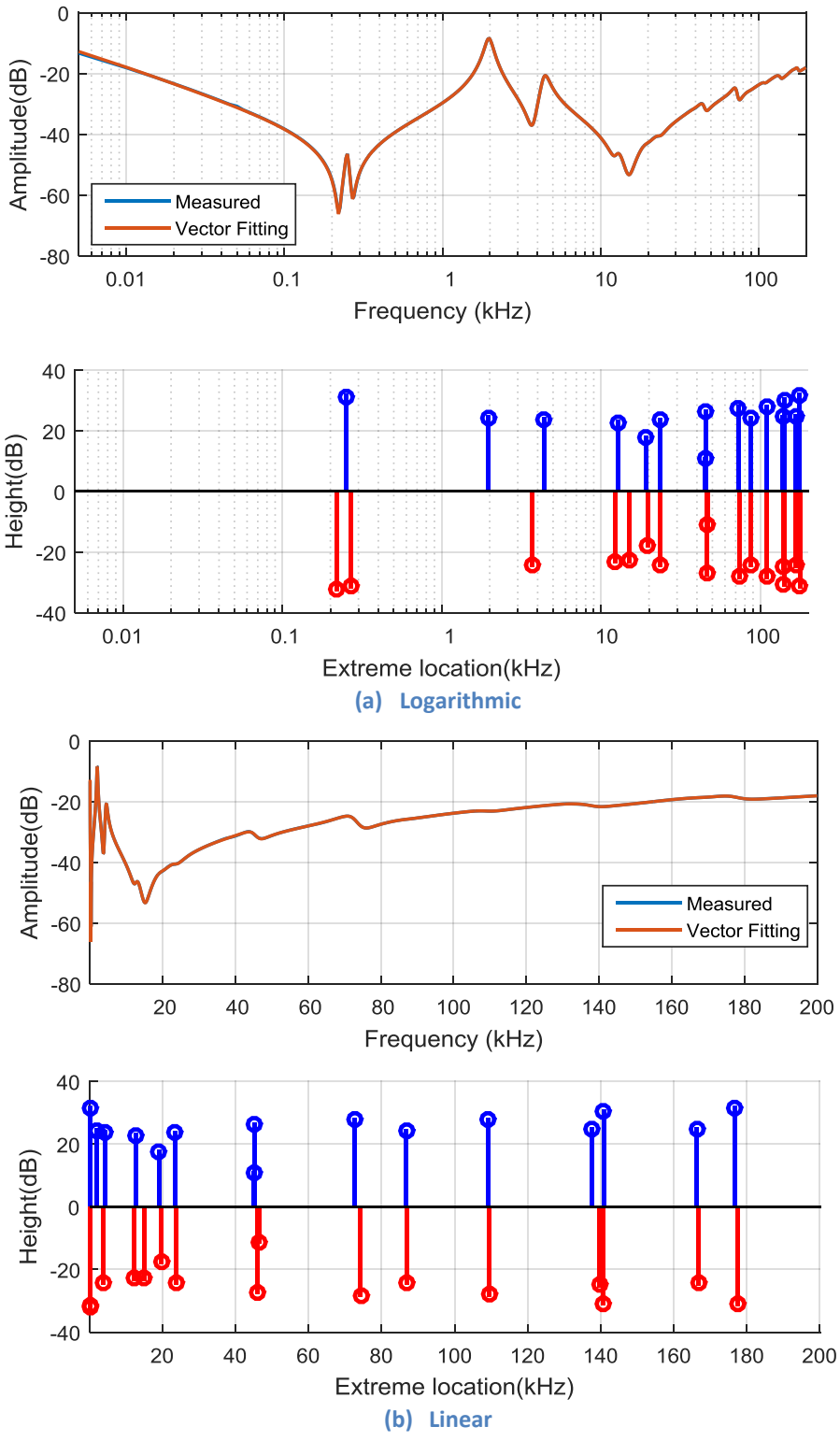
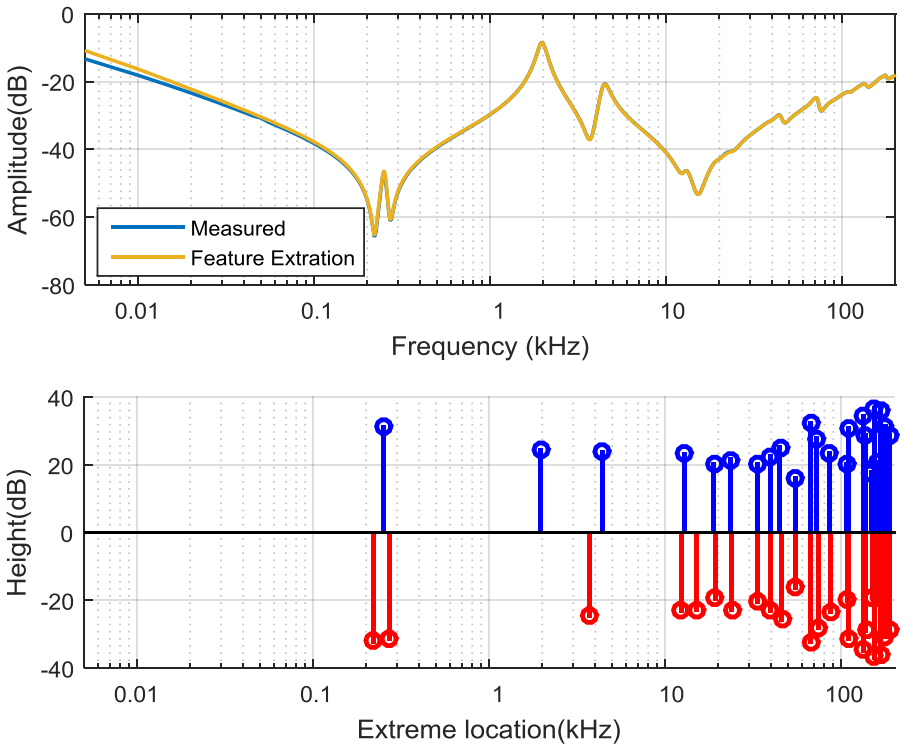


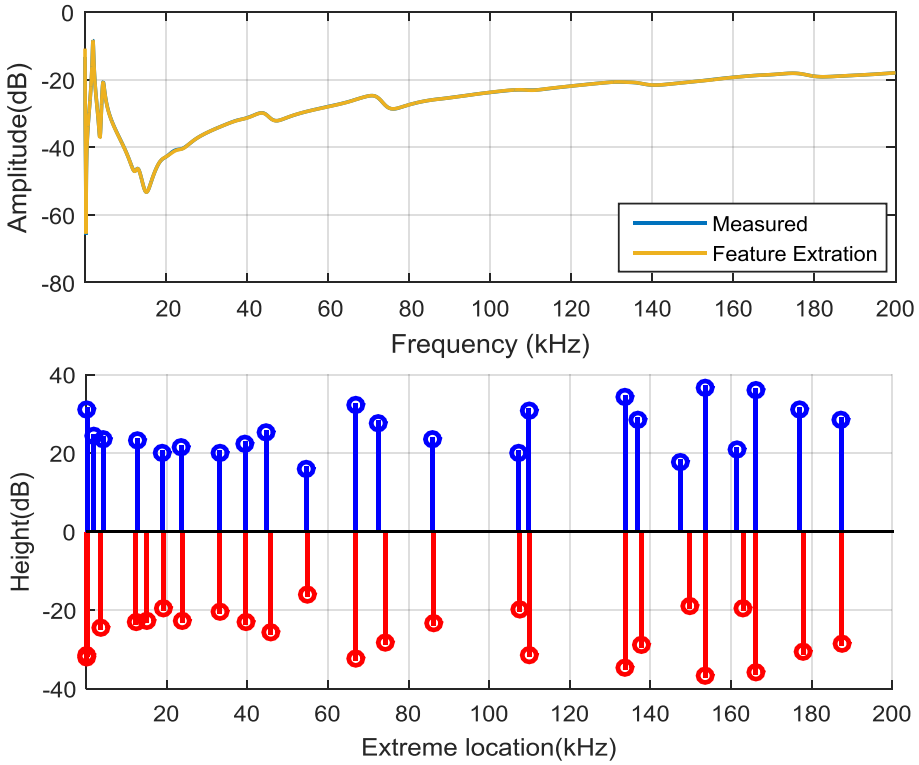
Figure 4-20 Complex Parameters of Estimated Transfer Function for Multiple Layer Winding Type Using Vector Fitting method

Table 4-15 Estimated Parameters of Multiple Layer Winding by Vector Fitting Method

	k=1.7900e+6			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	-	-128940	-	-
2	-199	-3492	-	-
3	-3427	-220	-	-
4	-106128	-43	-	-
5	-34+1387i	-6	0.221	-
6	-46+1692i	-42+1570i	0.269	0.250
7	-1426+23284i	-753+12360i	3.699	1.964
8	-5552+77571i	-1813+27691i	12.314	4.398
9	-6825+94796i	-5861+80677i	15.048	12.806
10	-16641+125021i	-16094+121039i	19.721	19.093
11	-9170+148756i	-9722+147245i	23.630	23.384
12	-12808+289612i	-13928+284838i	46.048	45.279
13	-93317+306454i	-96476+298702i	46.457	44.992
14	-18262+466876i	-18862+456420i	74.249	72.579
15	-33850+547286i	-33978+546153i	86.937	86.755
16	-27211+688358i	-28018+686277i	109.470	109.133
17	-49748+878733i	-51289+866718i	139.630	137.701
18	-25611+883900i	-26880+885548i	140.618	140.874
19	-63372+1050470i	-60331+1046281i	166.883	166.244
20	-31153+1115851i	-28786+1111372i	177.524	176.821
21	-	-2280067+3452004i	-	412.504



(a) Logarithmic



(b) Linear

Figure 4-21 Complex Parameters of Estimated Transfer Function for Multiple Layer Winding Type Using Feature Extraction Method

Table 4-16 Estimated Parameters of Multiple Layer Winding by Feature Extraction Method

	k=-2.0200e-14			
	Zero	Pole	fz(kHz)	fp(kHz)
1	4532795	-	-	-
2	-66860	4375	-	-
3	4569	-14	-	-
4	-35+1388i	-44+1572i	0.221	0.250
5	-46+1697i	-752+12358i	0.270	1.963
6	-1419+23289i	-1815+27680i	3.700	4.396
7	-5558+77852i	-5627+80919i	12.359	12.847
8	-6997+94680i	-11974+119328i	15.028	18.896
9	-13111+121599i	-12563+148621i	19.240	23.569
10	-11129+150072i	-20650+209600i	23.819	33.197
11	-20409+209517i	-18723+248948i	33.187	39.509
12	-17766+249904i	-15528+281827i	39.673	44.786
13	-15282+288369i	-56173+348483i	45.831	54.738
14	-56739+350470i	-10298+419804i	55.043	66.794
15	-10364+419721i	-18769+456256i	66.780	72.554
16	-18407+466723i	-36677+541811i	74.223	86.034
17	-36980+543362i	-67440+677163i	86.278	107.238
18	-69935+680017i	-19771+690359i	107.654	109.829
19	-18717+691146i	-16245+840392i	109.959	133.728
20	-15643+840503i	-31949+860657i	133.747	136.883
21	-31030+866844i	-121870+934182i	137.874	147.409
22	-107544+946340i	14240+965707i	149.639	153.680
23	14138+965763i	-93507+1018347i	153.689	161.390
24	-110916+1030659i	-16631+1043771i	163.082	166.100
25	-16810+1043808i	-30634+1112738i	166.106	177.031
26	-32779+1118201i	-43905+1178032i	177.891	187.359
27	-44779+1178994i	-233596+1330736i	187.507	208.505
28	-278749+1340177i	-	208.631	-

Table 4-17 Numerical Indices Comparing Measured and Estimated Frequency Responses

	Vector Fitting	Feature Extraction
Average absolute magnitude difference	0.0306 dB	0.0732 dB
Standard deviation of magnitude	0.0610 dB	0.1473 dB
Maximum absolute magnitude difference	0.7193 dB	2.4725 dB
Correlation Coefficient Of magnitude	1.0000	1.0000

The quantities of different types of parameters estimated by the 3 methods are shown in Table 4-18 Feature Extraction has the largest quantity of estimated parameters. Among all the methods, Extreme Points Iteration Method owns the smallest quantity of parameters. Only one real pole is produced by this method while the other two methods produce several real zeros and poles. For the proposed method, all of the complex zeros and poles estimated are dominant and they are found to coincide with part of those parameters estimated by the other two methods. For both Vector Fitting Method and Feature Extraction Method, redundant zeros and poles may appear, and some zeros and poles may not appear in sequence. For Feature Extraction Method some zeros and poles may not have physical meaning (with positive real part). If the cancellation criteria of Feature Extraction Method are relaxed, more zeros and poles can be viewed as identical.

Table 4-18 Comparison of Amount of Estimated Parameters

Number	Vector Fitting	Feature Extraction	Extreme Point
Real zero	3	3	0
Real pole	5	2	1
Complex zero pairs	16	25	10
Complex pole pairs	16	24	10

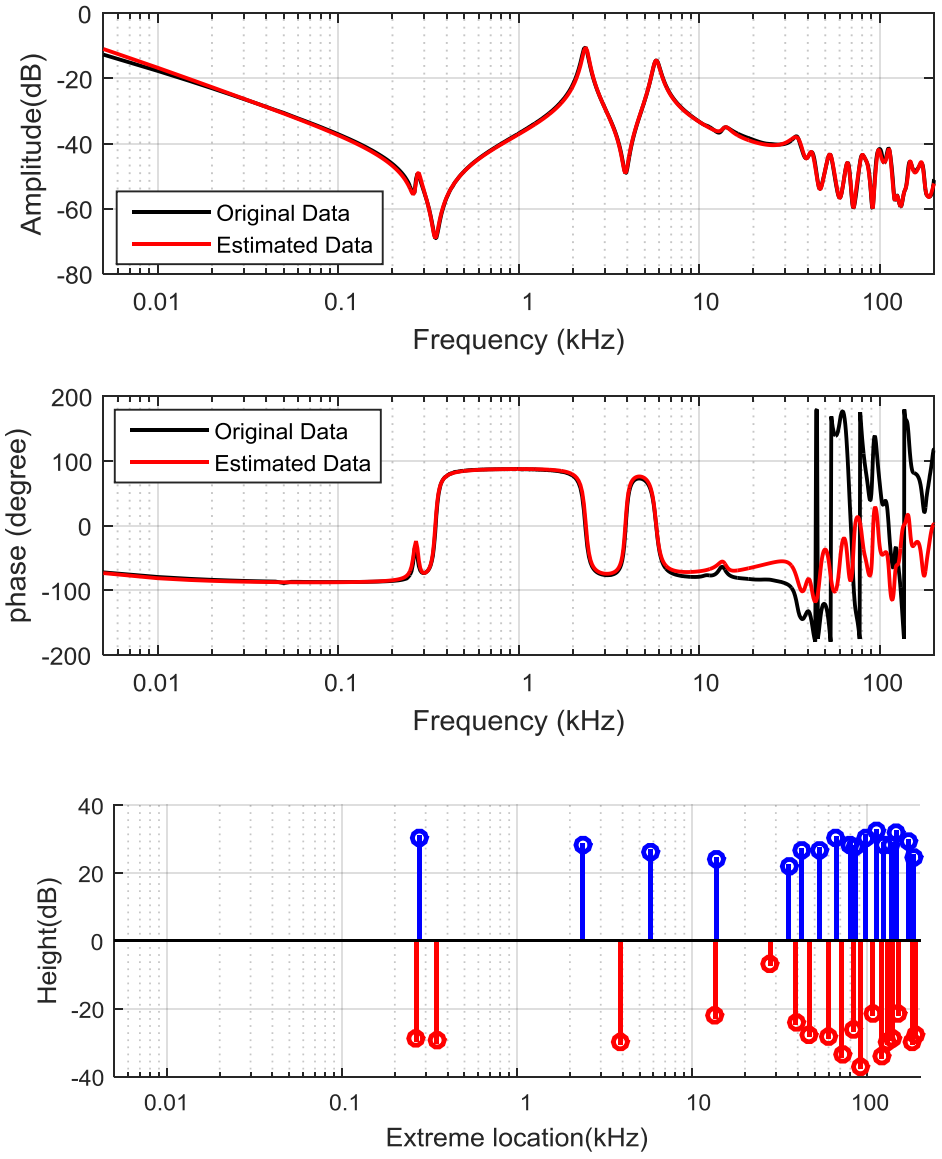
All of the 3 methods need to consider zeroes or poles outside the measured frequency range, i.e., beyond 200 kHz.

B. Low Series Capacitance Winding Type

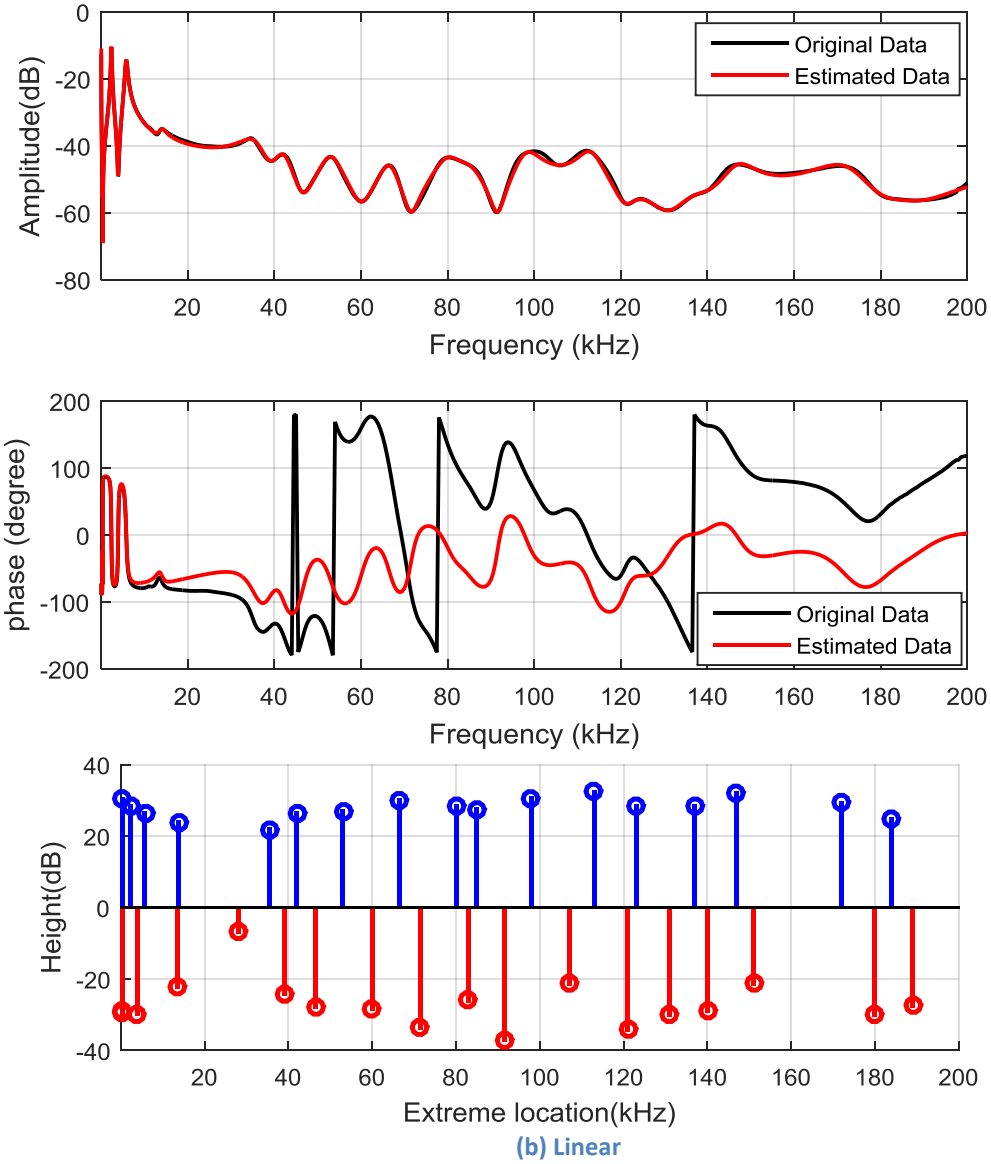
Plain Disc Winding has relatively low series capacitance. In the frequency region controlled by winding properties, Plain Disc Winding's frequency response has the feature of the camel humps, with either rising or flat magnitude trend.

An example of the frequency response with flat magnitude trend is plotted in black line in Figure 4-22, which is tested on C phase, Plain Disc Winding, HV to LV terminals of a 400/275/13 kV autotransformer. In this case, when the phase oscillates between ±180 degrees, the Vector Fitting method will fail to do the estimation, whilst the other two methods can estimate the transfer functions. Using the newly proposed method, the estimation result is plotted in red line in Figure 4-22. The parameters of the estimated transfer function are also plotted in Figure 4-22, and their exact values are given in Table 4-19. Two pairs of recessive complex zeros and poles are added manually to eliminate the evident difference between the measured and estimated data, i.e. -26312±522168i and -

22390±534540i located at 83 kHz and 85 kHz, and -32990±861428i and -32424±880243i located at 137 kHz and 140 kHz .Despite the small deviation caused by the other recessive parameters, the overall magnitude shows good agreement with the measured frequency response. However, the phase comparison shows a significant difference. The estimated phase plot is the minimum phase shift function. And this deviation is caused by the time lag caused by travelling wave during the FRA measurement.



(a) Logarithmic



(b) Linear

Figure 4-22 Comparison of Measured and Estimated Plain Disc Winding FRA Traces C phase HV to LV

Table 4-19 Estimated Parameters of Plain Disc Winding by Feature Extraction Method

	k=-2.4039e+6			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1		-9		
2	-60+1666i	-53+1729i	0.265	0.275
3	-72+2169i	-580+14777i	0.345	2.350
4	-797+24517i	-1747+36171i	3.900	5.750
5	-6670+85398i	-5599+86889i	13.550	13.800
6	-160553+238177i	-18365+223808i	28.000	35.500
7	-15325+245523i	-12602+264195i	39.000	42.000
8	-12119+292419i	-15082+333350i	46.500	53.000
9	-14757+377280i	-13113+418038i	60.000	66.500
10	-9341+449345i	-19509+503033i	71.500	80.000
11	-26312+522168i	-22390+534540i	83.000	85.000
12	-8195+574970i	-18225+616022i	91.500	98.000
13	-59021+674887i	-17283+710210i	107.000	113.000
14	-15060+760415i	-29466+773393i	121.000	123.000
15	-26655+823529i	-32990+861428i	131.000	137.000
16	-32424+880243i	-23411+923925i	140.000	147.000
17	-81617+952265i	-37675+1081364i	151.000	172.000
18	-36333+1131557i	-67383+1158068i	180.000	184.000
19	-50090+1188578i	-275901+1440388i	189.000	225.000

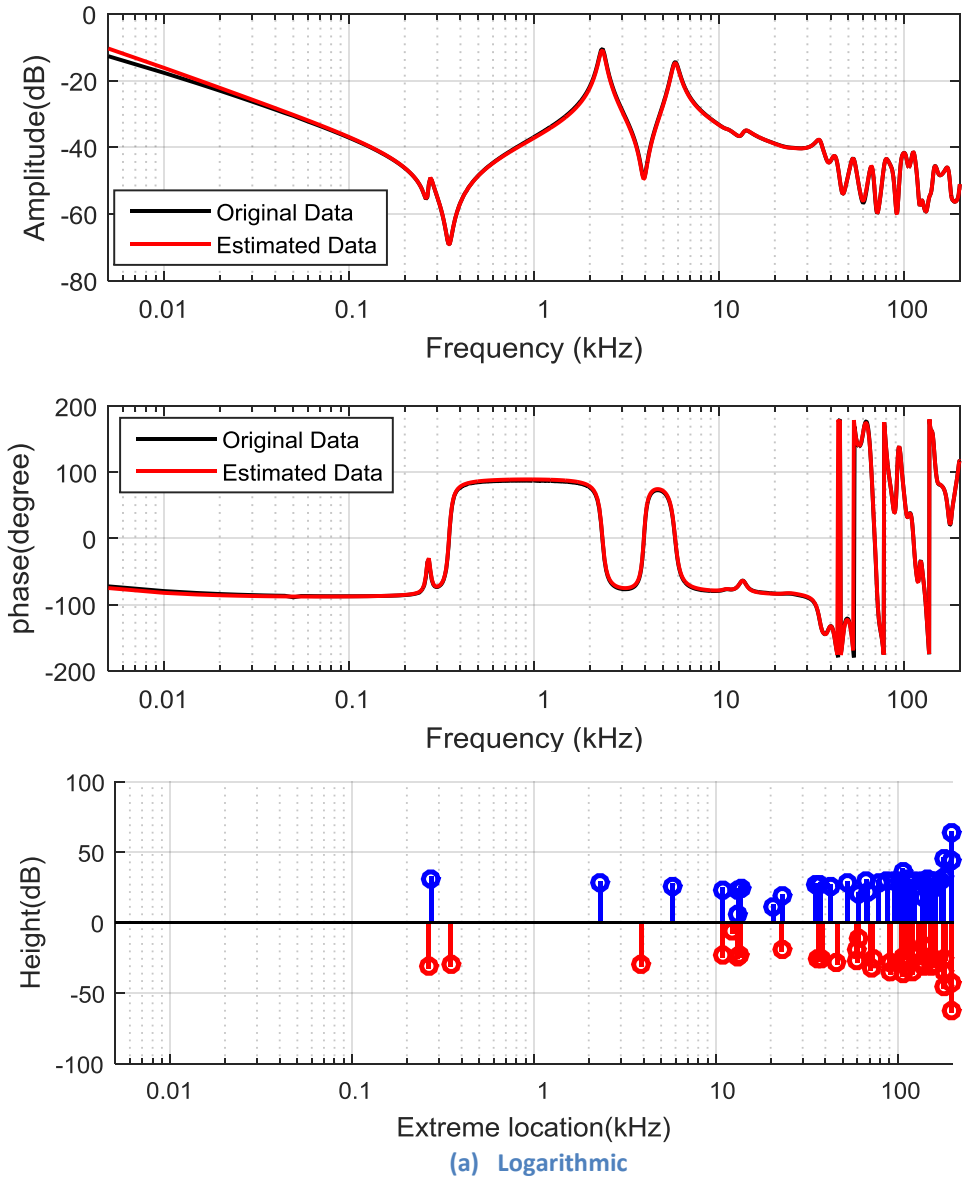
The numerical indices from Table 4-20 are used to reflect the similarity between the measured and estimated data. Noticeably, the Average Absolute Phase Difference is only 35.7549 degrees, which may seem to be small when considering the large frequency span from 20 kHz to 200 kHz. This is because that there are 760 points from 5 Hz to 20 kHz whose phase data match perfectly, and the amount of those data points is considerable.

In Figure 4-23, the estimation result from Feature Extraction method is compared with the measured data. It can be seen that not only the magnitude, but also the phase matches perfectly. The estimated parameters are plotted in Figure 4-23 and listed in Table 4-21. In Table 4-12 four numerical indices are used to evaluate the results. The Maximum Absolute Phase Difference is very large, reaching 358.7255 degrees. This variation is at 44.0 kHz, where the phase oscillates between ±180°. Actually this 358.7255 degrees difference could be wrapped into 1.2745 degree. Generally, the estimation result is satisfactory.

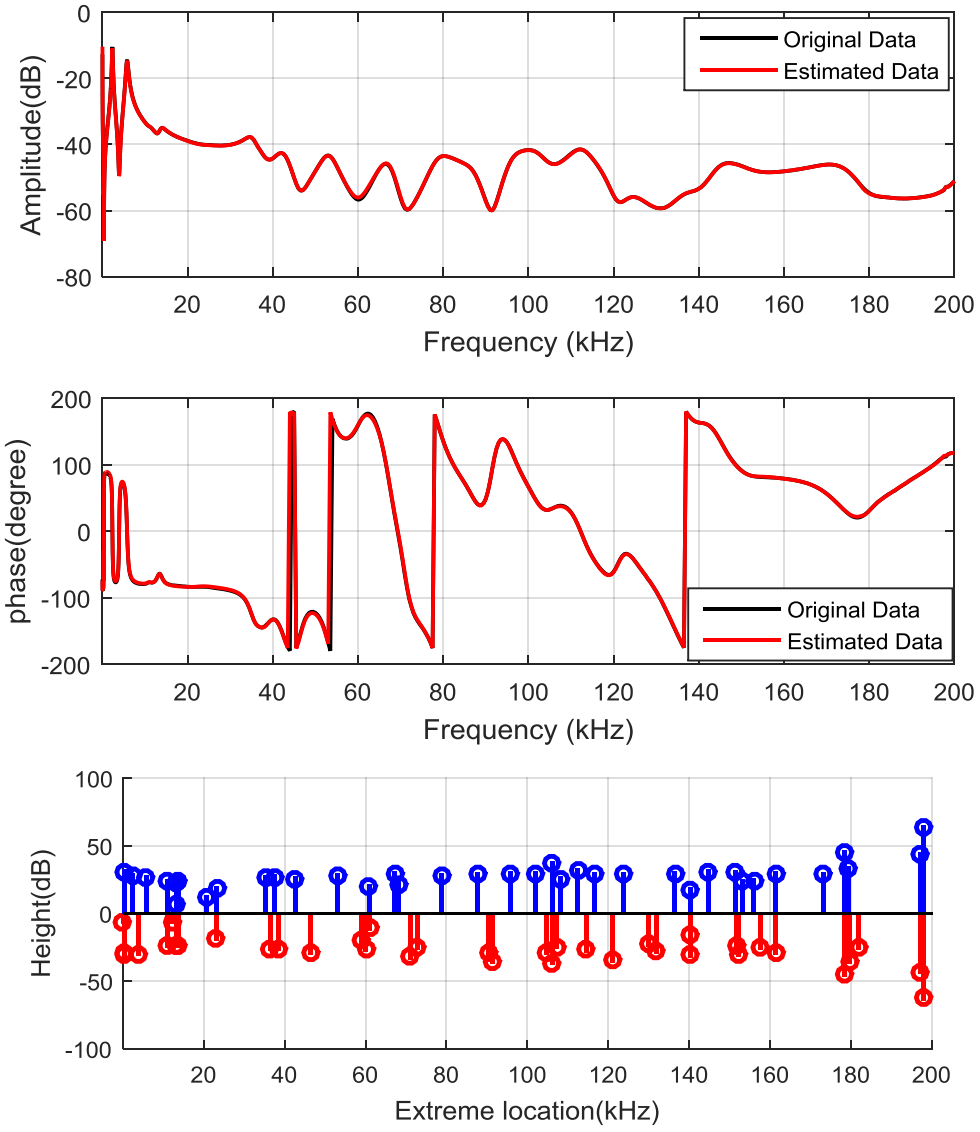
The Feature Extraction Method uses 2 real zeros, 2 real poles, 39 pairs of complex zeros and 39 pairs of complex poles to describe the given FRA trace. However, the Extreme Points Iteration method uses only 1 real pole, 18 pairs of complex zeros and 18 pairs of complex poles, producing satisfactory magnitude matching result and minimum phase shift.

Table 4-20 Numerical Indices comparing Measured and Estimated FRA Traces Using Extreme points Identification Method

	magnitude	phase
Average absolute difference	0.3259 dB	35.7549 degree
Standard deviation	0.2637 dB	54.2685 degree
Maximum absolute difference	1.7288 dB	294.7913 degree
Correlation Coefficient	0.9994	0.7048



(a) Logarithmic



(b) Linear

Figure 4-23 Comparison of Measured and Estimated Plain Disc Winding FRA Traces C phase HV to LV Using Feature Extraction Method

Table 4-21 Estimated Parameters of Plain Disc Winding by Feature Extraction Method

	k=-7.9880e-19			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	-79938	5432	-	-
2	5812	-8	-	-
3	-47+1662i	-48+1710i	0.264	0.272
4	-72+2180i	-586+14666i	0.347	2.332
5	-786+24624i	-1760+36243i	3.917	5.761
6	-4623+68599i	-136214+40091i	10.893	-
7	-5287+82626i	-4655+68855i	13.123	10.933
8	-5807+85593i	-6028+83681i	13.591	13.284
9	100235+126615i	-5238+86046i	12.312	13.669
10	-17179+146114i	95042+126585i	23.094	13.307
11	-212572+188635i	-17715+146694i	-	23.176
12	-11124+229155i	-10627+222156i	36.428	35.317

13	-11796+241182i	-10709+235302i	38.339	37.411
14	-10893+291387i	-14066+268224i	46.343	42.630
15	-39143+373273i	-13693+333833i	59.081	53.086
16	-16999+377849i	-38308+384467i	60.076	60.885
17	123910+403104i	-14440+421981i	61.050	67.121
18	10909+446627i	-35194+430223i	71.062	68.243
19	-25246+458429i	-19638+495414i	72.851	78.786
20	-20644+568390i	-19552+550898i	90.402	87.623
21	-9784+574740i	-20044+602409i	91.459	95.823
22	-24141+659285i	-22452+641981i	104.858	102.112
23	9849+667057i	9913+666988i	106.154	106.143
24	-38761+674983i	-35857+680159i	107.250	108.100
25	-31897+720580i	-19026+706891i	114.572	112.464
26	-13477+760694i	-26601+733329i	121.049	116.636
27	-59754+818736i	-27726+777903i	129.958	123.728
28	31566+828966i	-27894+858030i	131.838	136.487
29	-25271+882108i	-124854+890481i	140.334	140.324
30	-147257+893517i	-25832+910010i	140.263	144.774
31	62747+956066i	-30655+952266i	151.835	151.479
32	-28674+956625i	64757+965335i	152.183	153.292
33	54314+992590i	60503+982212i	157.739	156.027
34	35833+1015194i	33625+1015296i	161.472	161.501
35	6140+1121205i	-38831+1088809i	178.443	173.179
36	-19091+1128380i	6368+1121062i	179.562	178.420
37	-62231+1145301i	-25875+1127401i	182.011	179.384
38	8484+1239662i	8223+1239578i	197.294	197.281
39	-924+1243669i	-861+1243685i	197.936	197.939
40	-32682+1259527i	-16824+1268729i	200.393	201.907
41	-893810+1900451i	446842+1484454i	266.926	225.300

Table 4-22 Numerical Indices Comparing Measured and Estimated FRA Traces using Feature Extraction Method

	magnitude	**phase**
Average absolute difference	0.1717dB	1.7337 degree
Standard deviation	0.1794 dB	15.1192 degree
Maximum absolute difference	2.2950 dB	358.7255 degree
Correlation Coefficient	1.0000	1.0000

4.3.4 Application on Faulty Winding

The proposed Extreme Points Identification Algorithm is applied for diagnosis of faulty winding. As mentioned, the complex parameters out of concerned frequency range can be estimated using an artificially defined resonant point, to accurately represent the measured

data at high frequencies on the concerned frequency range.

This method is applied to the frequency responses of 275/132/11 kV autotransformer T2305. The B phase tap winding of this transformer is axially collapsed.

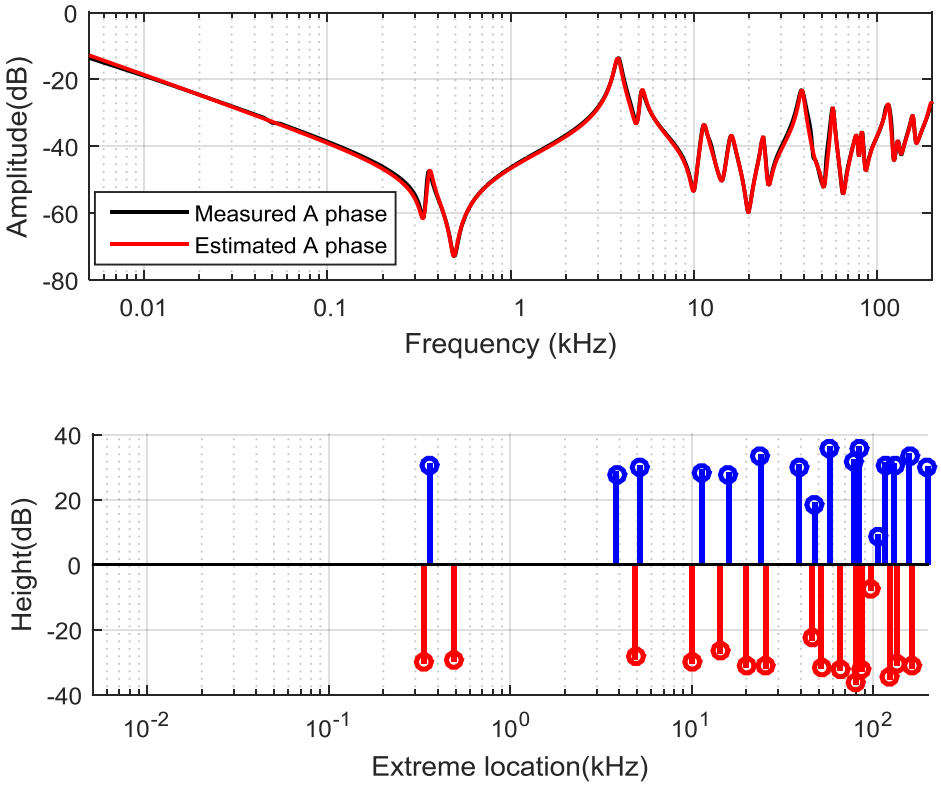
The measured frequency responses on common windings of A phase and B phase, N to LV (Tap 19) are estimated, and the comparison is illustrated in Figure 4-24 and Figure 4-25. The corresponding parameters are listed in Table 4-23 and Table 4-24. Noticeably, all of the resonant points are located at the measured frequency point, under the assumption of the algorithm that the resonance and antiresonance are located at a measured frequency point.

The estimated frequency responses match well with the frequency responses measured for both A phase and B phase. The original resonance located at 39 kHz, which should be the same as the one in the frequency response of A phase, is shifted to 35 kHz in the frequency response of B phase after the fault, which is circled in Figure 4-26 (a). The shifting can be reflected by the alteration of estimated parameters. The pair of complex poles -7522±245160i becomes -8016±220058i. In section 4.2.3, the two corresponding pairs of complex poles, estimated by the Feature Extraction Method, are -9696±244344i and -8797±220568i. The differences between their imaginary parts are 0.33% and 0.23%, whilst the differences between the real parts are 22.42% and 8.87%. The imaginary parts are very similar. The large deviation in the real parts is understandable, since the real part does not play an important role as the imaginary parts in determining the resonant frequency, and the real part is influenced by the neighbour resonant points.

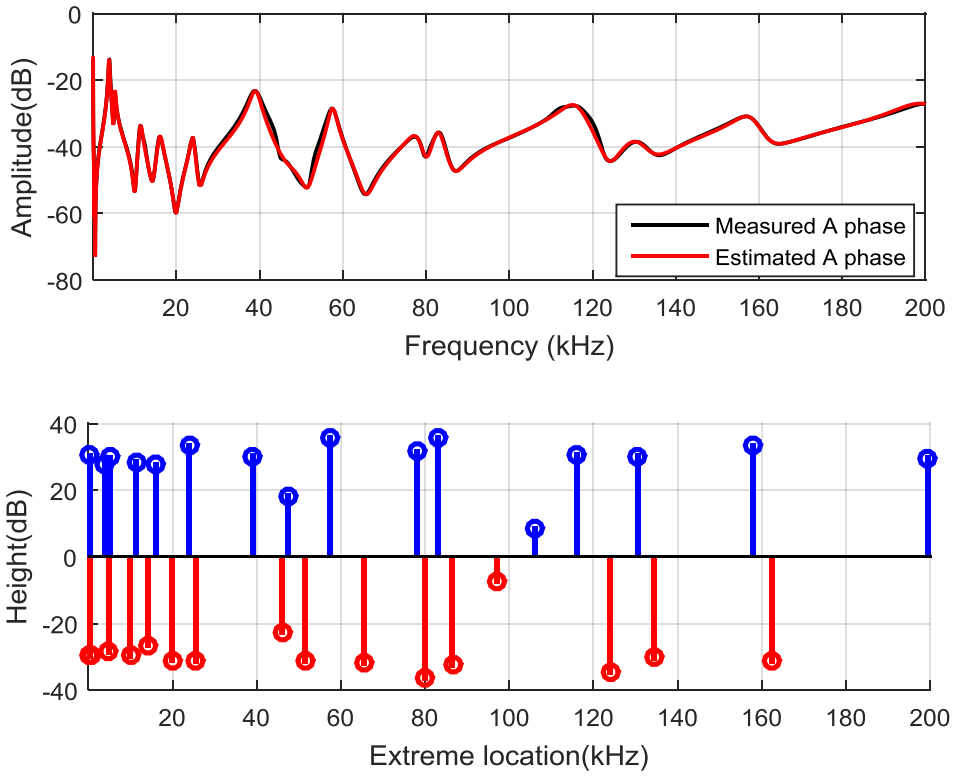
There are 17 pairs of complex zeros and 16 pairs of complex poles used to describe the frequency response from A phase in this proposed method, whilst there are 18 pairs of complex zeros and 17 pairs of complex poles used to describe the frequency response from B phase. As well known, the frequency response of A phase has 2 resonance and 1 antiresonance in the core controlled frequency region, roughly below 2 kHz, while the frequency response of B phase has only 1 resonance, due to the difference in the magnetic reluctance paths. This means that due to the fault, there are 2 new pairs of complex zeros and 2 new pairs of complex poles appearing in the frequency response of B phase. The newly appeared resonant points are circled in dashed line in Figure 4-26 (b). One new pair of complex zeros and one new pair of complex poles appear around 80 kHz, i.e. -9332±490177i and -12765±496536i, while the other new pairs of complex zero and

complex poles appear around 120 kHz, i.e. -26758±741899i and -20957±704029i, as listed in Table 4-24.

The measured frequency response of A and B phase are plotted together and the estimated ones are plotted together in Figure 4-26. It can be concluded that the proposed method is able to give a satisfactory match of the measured frequency response, and the difference between the reference and diagnostic frequency response can be analysed through the parameters estimated.

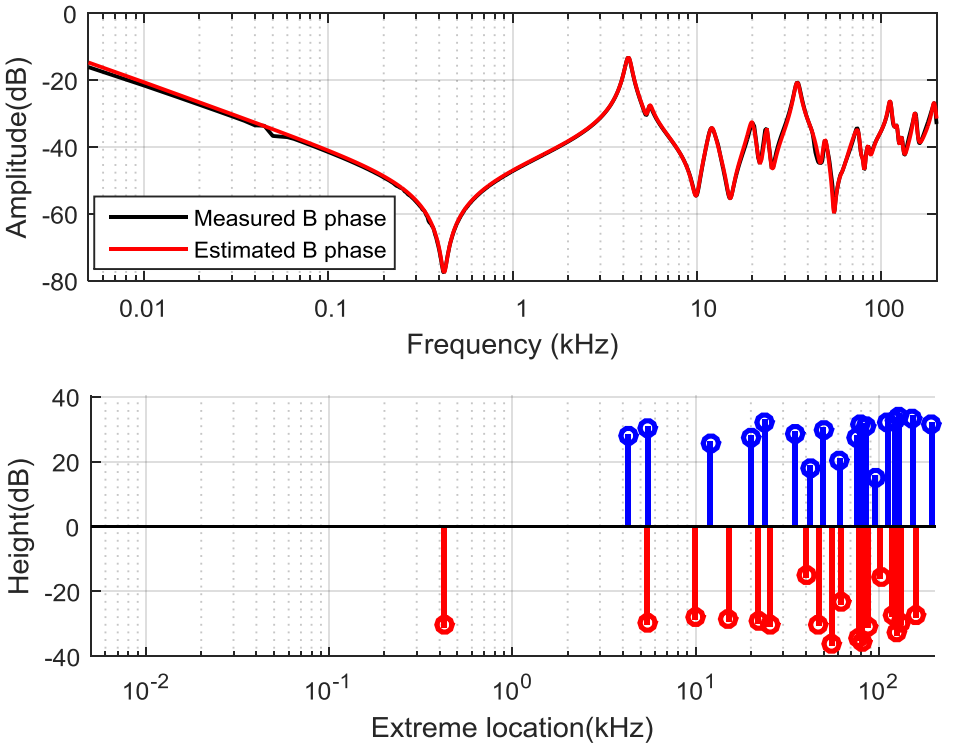


(a) Logarithmic

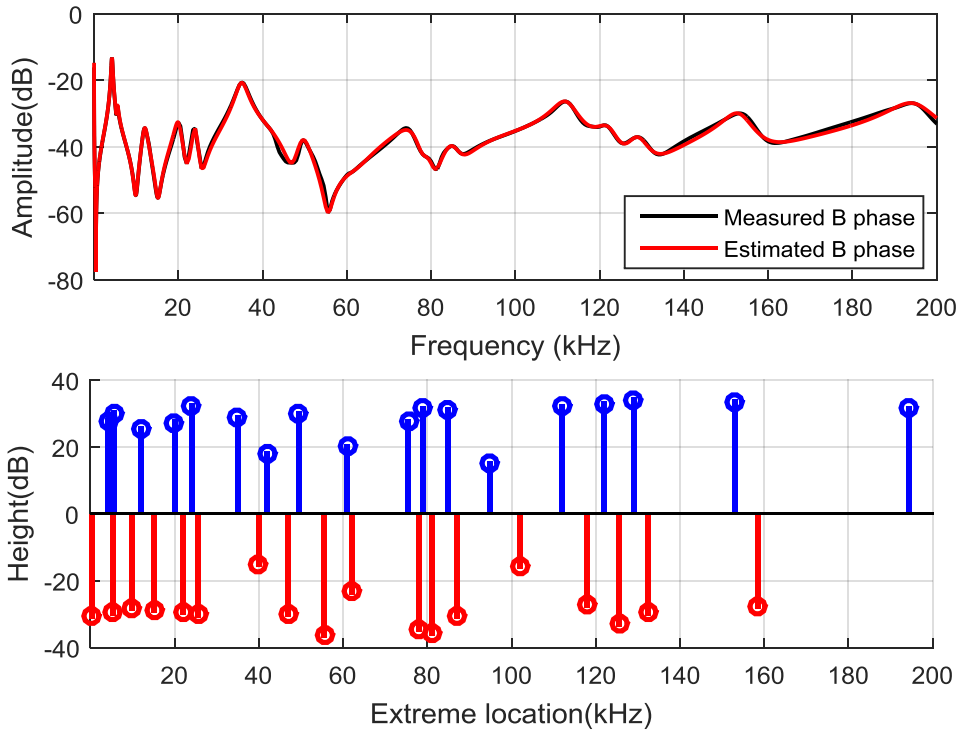


(b) Linear

Figure 4-24 Comparison of Measured and Estimated Frequency Response of N to LV A phase T2305



(a) Logarithmic



(b) Linear

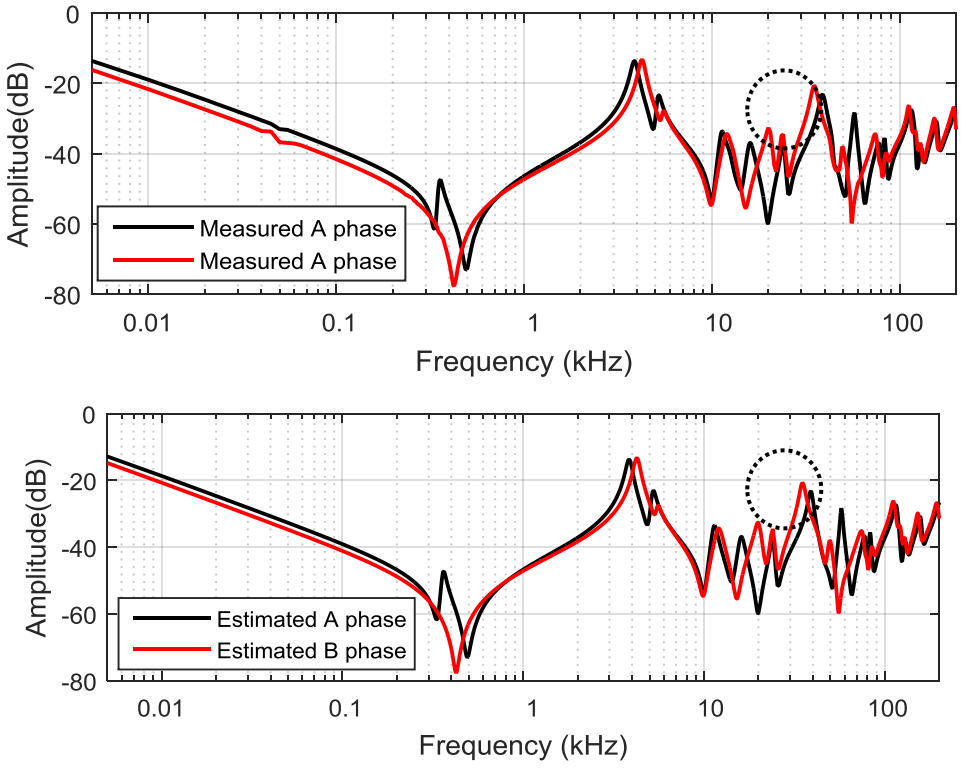
Figure 4-25 Comparison of Measured and Estimated Frequency Response of N to LV B phase T2305

Table 4-23 Estimated parameters of N to LV A Phase T2305

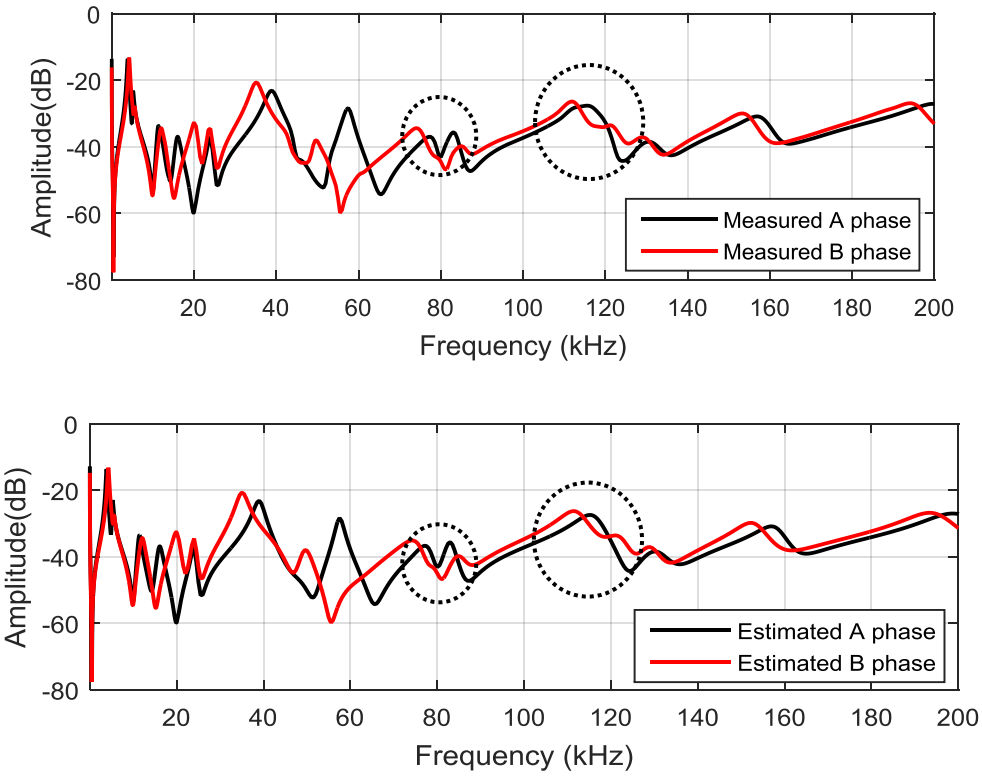
	k= 1.4752e-08			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	-68+2106i	-7	0.335	
2	-106+3081i	-68+2263i	0.490	0.360
3	-1170+30810i	-979+24210i	4.900	3.850
4	-2069+62866i	-1059+32690i	10.000	5.200
5	-4210+89634i	-2694+71365i	14.250	11.350
6	-3539+125085i	-4226+100620i	19.900	16.000
7	-4391+160281i	-3292+150832i	25.500	24.000
8	-21358+289815i	-7594+245162i	46.000	39.000
9	-8743+323702i	-37099+300748i	51.500	47.500
10	-10332+411678i	-5875+361331i	65.500	57.500
11	-7671+502713i	-12954+490260i	80.000	78.000
12	-424558+742767i	-8567+521575i	97.000	83.000
13	-13139+543654i	-340720+748111i	86.500	106.000
14	-14403+779248i	-21215+729158i	124.000	116.000
15	-25911+845486i	-24992+820336i	134.500	130.500
16	-27675+1021393i	-21235+992970i	162.500	158.000
17	-98007+1323104i	-41225+1254173i	210.000	199.500

Table 4-24 Estimated parameters of N to LV B Phase T2305

	k= 1.9089e-08			
	Zero	**Pole**	**fz(kHz)**	**fp(kHz)**
1	-79+2672i	-6	0.425	-
2	-1121+33948i	-1100+26726i	5.400	4.250
3	-2446+62252i	-1075+34574i	9.900	5.500
4	-3501+95255i	-4061+75508i	15.150	12.000
5	-4739+138311i	-5344+125777i	22.000	20.000
6	-4958+160298i	-3630+150840i	25.500	24.000
7	-46476+255588i	-8016+220058i	40.000	35.000
8	-9491+295462i	-32884+265935i	47.000	42.000
9	-5787+348765i	-9608+311166i	55.500	49.500
10	-32668+390925i	-44710+385873i	62.000	61.000
11	-9332+490177i	-18519+474742i	78.000	75.500
12	-8732+509013i	-12765+496536i	81.000	79.000
13	-17217+546908i	-14246+534261i	87.000	85.000
14	-26758+741899i	-20957+704029i	118.000	112.000
15	-16195+788706i	-18831+766780i	125.500	122.000
16	-22799+832834i	-17394+810718i	132.500	129.000
17	-35755+996527i	-22831+961598i	158.500	153.000
18	-65950+1289740i	-35660+1222600i	205.000	194.500



(a) Logarithmic



(b) Linear

Figure 4-26 Comparison of Measured and Estimated Frequency Response of T2305

4.3.5 Summary

A novel method has been developed to estimate the transfer function of frequency response. It is simple to apply and easy to understand, while giving a unique solution to a specific FRA trace; Regardless of which kind of measurement device is used, which may have different accuracy on the magnitude, i.e. -80dB/-100dB, and how the data are sampled as different devices can be set with different sampling frequency range, with different distribution on either a logarithmic or linear frequency scale, the large quantity of stored FRA data (normally a few thousands) could be shrunk into several dozens of zeros, poles and a constant.

The resolution does have an influence on the estimation result using this proposed method, especially when zeros and poles are closely located. Three methods to estimate transfer functions are compared. And their advantages and disadvantages are listed in Table 4-25. The similarity in the shape of traces can be reflected mathematically by the parameters of transfer functions estimated. When any changes happen to the FRA trace, such as the magnitude variation, shift in resonance or antiresonance, or changes in quantity of zeros and poles, the parameters of transfer function could also be changed which can be used as

indicators of fault.

Application of this method to diagnosis transformer windings is promising. The unique solution of transfer function can guarantee an accurate interpretation. Further work needs to be done to complete this method in terms of recessive parameters.

Table 4-25 Comparison of 3 Estimation Methods

	Vector Fitting	Feature Extraction	Extreme Points Identification
Need magnitude spectrum only	×	×	✓
No redundant zeros/poles	×	×	✓
Zeros & poles in sequence	×	×	✓
Physically achievable Meaning	✓	×	✓
Process data with phase between ±90°	✓	✓	✓
Process data with phase between ±180°	×	✓	✓
Match phase data between ±180°	×	✓	×
Unique solution	×	×	✓

4.4 Conclusion

Two methods have been developed and they are introduced in this chapter. The first method is called Feature Extraction Method, is based on the MATLAB command 'invfreqs', which uses the Non Linear Least Square algorithm. Because this command is only capable to estimate a transfer function with limited order, the larger frequency range 5 Hz to 200 kHz are divided into several frequency regions. The key information, complex zeros and complex poles, are extracted from each frequency region and combined to form a Feature Transfer Function. Then a Difference Transfer Function is used to correct the deviation between the Feature Transfer Function and the measured data, and a Finalized Transfer Function can be produced. This method has been applied on different winding types, and satisfactory match for both magnitude response and phase response can be achieved. However, the complex poles and zeros of the finalised transfer function may not be physically achievable and may not appear in the right sequence.

Thus the second method, Extreme Points Iteration Method, is proposed. The height and frequency location of the resonance and antiresonance are related to the real and imaginary parts of the complex parameters. The mathematical relationships can be built to compute the complex parameters. Iterations should be conducted to eliminate the mutual influences between the complex parameters. This method has been successfully applied on the FRA traces from a Multiple Layer winding and a Plain Disc winding. Due to the existence of recessive parameters, the application on an Interleaved Disc winding is not as good as the

estimation on the other winding types, thus further work is required to identify the recessive parameters.

The Extreme Points Iteration Method is capable to produces a physically achievable minimum phase shift transfer function for the frequency response with phase data oscillating between ±180°.

Both of the proposed methods have been applied to estimate the transfer functions of the frequency responses measured on common windings of A phase and B phase, from N to LV (Tap 19) of a 275/132/11 kV autotransformer. The investigated transformer's tap winding of B phase axially collapsed and because there was no previous FRA measurements done on the investigated transformer, construction based comparison method was applied. The frequency response from A phase is used as reference measurement.

Both of the estimation methods are able to produce well-matched frequency responses of the measurement results. The parameters of the transfer functions estimated by both methods are able to reflect the alteration in the frequency responses. The Feature Extraction Method can describe the delicate features of the frequency responses well. The Extreme Points Identification Algorithm can generate a simpler transfer function expression, which is physically achievable. Each pair of the complex parameters, produced by Extreme Points Identification Algorithm, corresponds to a resonant point. Therefore, the analysis of alteration in the frequency response is easier using Extreme Points Identification algorithm.

Chapter 5 Identification of Winding Construction Types by Supervised Machine Learning Method

5.1 Introduction

Windings are important electrical components of a transformer, and the choice of winding construction type is greatly influenced by the transformer manufacturer's historic experience, the transformer's voltage and power rating. In general, for the same power rating, a higher voltage winding prefers to use the winding construction type which gives a larger winding series capacitance due to the requirement to withstand a stringent BIL level. It is well known that different winding construction types are susceptible to different modes of mechanical deformations, hence knowing the winding design information is helpful for transformer fault diagnosis. However in practice, transformer asset managers in utilities know it too well that a large number of transformers operating in their networks are lack of design information such as winding type, structure and dimension, as this information is the manufacturers' safe guarded know-how. Effective asset management, especially for those transformers without any technical support from the Original Equipment Manufacturer (OEM), calls for the development of non-intrusive winding type recognition techniques.

Frequency Response Analysis (FRA) has been developed as an effective and sensitive technique to identify winding mechanical movement such as displacement or deformation. The FRA characteristics at different frequency regions are determined by different parts of a transformer. For the 400/275/13 kV transformers to be investigated, the frequency range can be split into three regions 0.005-2 kHz, 2-20 kHz and 20-1000 kHz, which are dominated by the core, inter-winding interaction and the structure of the winding-under-test, respectively. The frequency region higher than 1 MHz is influenced by the FRA measurement setup such as the bushing and the earthing lead. For the frequency region of 20-1000 kHz, the ratio between winding series capacitance and shunt capacitance is the main shaping factor. Different winding types have different typical values of winding series capacitance, and thus different FRA characteristics in this frequency region. This lends itself to the winding type recognition.

Different artificial intelligence methods, especially pattern recognition techniques have been applied for the transformer fault diagnosis. Those methods have also been applied to

FRA for various purposes, such as estimation of parameters of transformer models and identification of winding. Pattern recognition techniques developed so far assumed that the winding type information is known beforehand and the solutions derived focused solely on that type. However, for different winding types, the characteristics of their FRA are different, thus the fault pattern may differ significantly. The identification of winding type is the cornerstone for further study of winding mechanical faults in order to establish a generic interpretation guide. On the other hand, as mentioned, different winding types are susceptible to different types of mechanical faults, thus knowing the design information such as the winding type can be helpful for transformer asset management. Therefore, there is an urgent need to identify winding types through FRA measurement data.

Support Vector Machine (SVM) method is a supervised machine learning method for classification problem. It finds an optimal hyperplane which separates two types of data in binary problems. Multiple binary SVM classifiers are used when the number of data types to be identified is larger than two. The SVM method has been employed successfully for the recognition of transformer winding fault types and degrees. The advantage of SVM is that it can deal with small size of samples, nonlinearity, and high dimensional issues. In addition, it can overcome the problem of local minimums in neural networks.

In this chapter, a novel SVM-based winding type recognition method is proposed. The SVM is trained and tested by FRA traces from the UK's National Grid FRA database of a group of three winding three phase 400/275/13kV auto transformers. Four winding types, including multiple layer, plain disc, interleaved disc and single helical are used in these transformers. Then, the SVM built is applied to FRA traces without winding type information to test its performance in winding type recognition. The test results and sensitivity studies confirm that the proposed method is comprehensive and can be used along with expert experience and forensic information to aid transformer winding fault interpretation and transformer asset management.

In the UK's National Grid FRA database, both the magnitude and phase are recorded, and only the magnitude frequency spectrum is used which proves to be sufficient for winding type recognition.

5.2 Support Vector Machine

A hard-margin classification model requires all features to be accurately classified, while a soft-margin classification model allows a certain amount of features to be misclassified.

Normally for soft-margin model, a regularization parameter can be used to avoid overfitting, by restricting the norm of the parameters. SVM was initially proposed for the two-type classification problem. Multiclass classification problems can also be solved by SVMs if using multiple binary classifiers. An SVM is a generalised linear classifier, which can be also applied to the nonlinear classification problem in combination with kernel method. In this study, only the hard-margin linear classifier is discussed.

The SVM algorithm finds a separation hyperplane from two groups of observations with known types and utilises it to categorize new examples. The hyperplane is also referred to as the decision boundary. The process of devising an SVM can be divided into two stages: training where the optimal hyperplane(s) is obtained from observing training data; and testing where the SVM is validated with a group of observations with known types. Upon completion, the SVM can be used to classify new observations.

5.2.1 Binary SVM Classifier

Binary SVM classifier finds a classification hyperplane, which separates two groups of observations with known types. The distances from this hyperplane to the nearest observations separately from the two data types should be equal. For example, in the two-dimensional space of Figure 5-1, an optimal hyperplane, represented by the bold line, leans neither to the nearest observation from type C1 (triangles) nor the nearest observation from type C2 (circles). Each observation, either a triangle or circle, is called a feature in SVM algorithm. The boundary features, which are circled in dash line in Figure 5-1, decide the hyperplane of binary classifiers, and they are the nearest observations to the classification hyperplane from each feature type. Such boundary features are called support vectors in SVM.

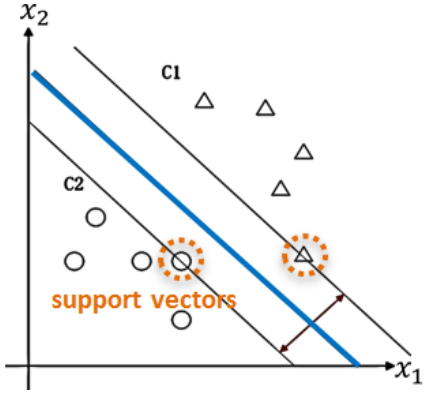


Figure 5-1 Binary SVM Classifier in Two-Dimensional Space

It is stated in details how to build a binary SVM model in Chapter 2.

5.2.2 Multiclass SVM Classifier

Since the binary SVM classifier can only distinguish two types of data, multiple binary SVM classifiers are needed for the multiclass classification problem. 'One-versus-one', 'one-versus-all' and 'binary tree' are three commonly adopted multiclass classification strategy. In this study, 'one-versus-one' method is used. This means that between every two classes, a decision is made on to which class the new feature is more similar, and the new feature is identified as the class which wins the most votes.

5.3 Winding Type Classification

The purpose of this study is to use SVM algorithm to classify the winding type of transformers. To obtain a well-functioned SVM model, the input features should be carefully selected. The statistic indices are not suitable as they are normally used for the comparison between two FRA traces, which means they indicate the similarity or dissimilarity between two sets of data, but not the characteristics of a single set of data. Though the transfer function can accurately describe the FRA traces in a mathematical way, one problem is that the parameters of a transfer function do not have a fixed quantity for both complex poles and zeros. This is because the number of peaks and troughs on different FRA traces from different winding types may vary. Besides, the quantity of real zeros and poles is also hard to be unified. Therefore, it is difficult to find appropriate input features for the SVM model using transfer function expression. As a result, in the study, the measured magnitude responses of different FRA traces, with unified format, are selected as the input features. FRA analysers on the market use different frequency and amplitude resolutions, which lead to the fact that the FRA trace may have different number of frequency points. In case such a scenario arises, pre-processing of data is needed and by applying transfer function estimation method, the measurement data in different frequency resolution can be expressed into a mathematical equation, which can reproduce FRA traces in the same desired format.

The FRA traces used for training and testing the SVM model are obtained from National Grid's FRA database with the same format, and all together 108 FRA traces are used in this study, including 30 multiple layer winding FRA traces, 36 plain disc winding FRA traces, 27 interleaved disc winding FRA traces and 15 single helical winding FRA traces.

Before training the SVM model, standardization should be applied to all 108 FRA traces as follows:

$$\boldsymbol{x}' = \frac{\boldsymbol{x} - x_a}{\sigma}$$ Equation 5-1

where $\boldsymbol{x}$ is the original vector consisting of the 1280 FRA amplitudes in dB for the frequency range from 5 Hz to 1000 kHz, x_a is the mean value of the original vector, σ is the standard deviation of original vector, and $\boldsymbol{x}'$ is the standardised dimensionless vector. The range of $\boldsymbol{x}'$ is roughly from -4 to 3. Though the classification difficulty will be increased after standardization, the focus can be put more on the traces' characteristics, rather than relying on the difference in magnitude.

There are altogether four winding types to be identified. Since the 'one-versus-one' strategy is adopted, a binary classifier is needed for every two winding types. Therefore, a total number of (4×3/2) =6 binary classifiers should be built. The input features are the standardized FRA magnitude response $\boldsymbol{x}'$. If there are altogether n FRA traces used as training feature, and there are 1280 points in each FRA magnitude response, the input of the SVM model to be built is an n ×1280 matrix.

Using the methodology introduced in Chapter 2, the weight vector $\boldsymbol{w}$ and bias b for each binary classifier can be computed with a given set of training data. By doing so, the multiclass classification SVM model can be built. Four labels are used for the four investigated winding types. The output of the SVM model is the label of the identified winding type, and the meaning of each label is as follows:

Label 1: Multiple Layer winding;

Label 2: Plain Disc winding;

Label 3: Interleaved Disc winding;

Label 4: Single Helical winding.

5.3.1 Cross Validation Process

Cross validation and bootstrap are two resampling methods, which can be used to evaluate the effectiveness of classification models when the quantity of available features is limited.

Bootstrap resamples randomly with replacement, which means duplicate features may be sampled. In a soft-margin SVM model, the weight of duplicate features changes, thus the parameters of SVM model can be influenced. A hard-margin SVM classification model does not notice the duplicate features.

The cross validation method without replacement is preferred. K-fold Cross Validation is therefore used to verify the applicability of SVM model under the small sample size

setting. It divides the original data into K roughly equal-sized folds. One fold is used as the testing data, and the rest K-1 folds are used as the training data. Each of the K folds should be used as the testing data once, which means K classification models are built. This makes the most of the available information. When the value of K is selected, the average accuracy can be used to evaluate the performance of the SVM model. With the change of K, the performance of the model may change. The model with the highest accuracy can be chosen.

The FRA traces from the A, B, and C phase of the same transformers should be all grouped into either the training or testing data group due to their similarity, in order to guarantee the credibility of the classification model. Considering there are 15 single helical FRA traces from 5 different transformers, the maximum amount of folds is 5. The 108 FRA traces investigated in this study are divided into 2, 3, 4, and 5 roughly equal-sized folds respectively. For example, for the 2 folds cross validation, there are 12 frequency responses from Multiple Layer winding, 18 frequency responses from Plain Disc winding, 18 frequency responses from Interleaved Disc winding and 6 frequency responses from Single Helical winding are used are training data, and the rest of the frequency responses are used as testing data. The number of training and testing traces when K changes are listed in Table 5-1. For each K-fold cross validation, different combination of training traces are selected randomly without repetition for 20,000 times. To guarantee that the sampling size 20,000 is large enough, the cross validation process is conducted twice. The corresponding average accuracies are listed in Table 5-1. It can be seen that the overall accuracy is satisfactory, which proves the feasibility of winding type classification through SVM method. Among all the SVM models built, the lowest accuracy is 38.89% when K=2. Though for this particular model, the training features used may not be appropriate, however the accuracy is still larger than 25% which can be resulted from random guess when distinguishing the four winding types. In this cross validation process, with the increase of number of training data, the accuracy of the SVM model increases.

Since the models of 2-fold cross validation have the lowest accuracy, one of the models built when K=2 with an accuracy of 100% is more of interest to investigate further in order to increase our understanding on when and why the SVM method can work well.

Table 5-1 Cross Validation Accuracy

	Number of Training Traces				Number of Testing Traces					
Winding type	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**Accuracy 1**	**Accuracy 2**
K=2	12	18	18	6	18	18	9	9	97.70%	97.75%
K=3	18	24	18	9	12	12	9	6	99.54%	99.56%
K=4	21	27	21	12	9	9	6	3	99.97%	99.97%
K=5	24	30	21	12	6	6	6	3	100.00%	100.00%

5.3.2 FRA Traces and Training Process

For the SVM model with an accuracy of 100%, which is built when K=2, altogether 54 FRA traces are used to train the SVM model, and another 54 FRA traces are used to test the model.

As mentioned before, the boundary frequencies of 2 kHz and 20 kHz are empirical; hence the whole range from 5 Hz to 1 MHz should be used in the following study of the SVM model. There exists a great similarity for FRA traces in the frequencies lower than 20 kHz for all the multiple layer, plain disc and interleaved disc windings, and this similarity is reasonable as the frequency region is dominated by the core and the interaction between windings. Inclusion of the low frequency regions would increase the complexity for classification but and enhance the confidence level of the SVM model developed.

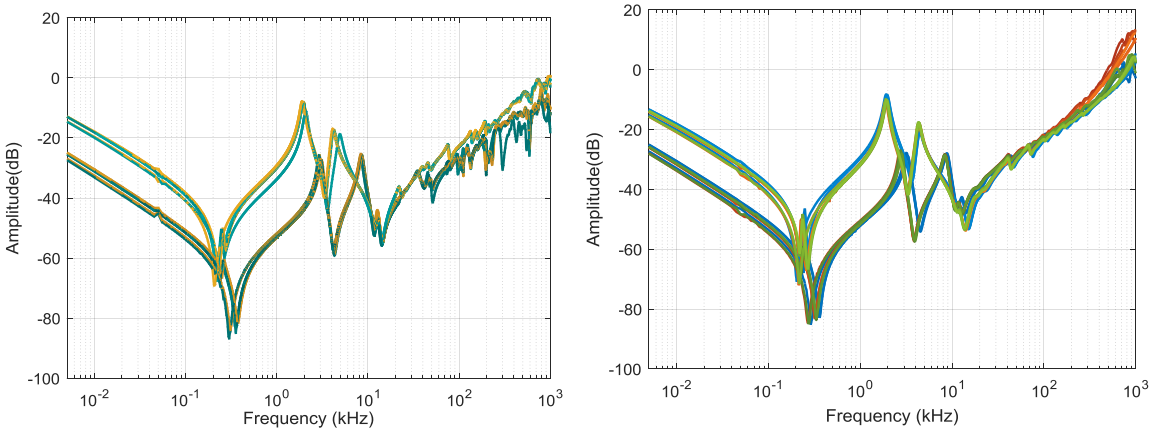
For the Multiple Layer windings, the training features are 12 FRA traces from two 500 MVA transformers, including the common and series windings from A, B, and C phases, as plotted in Figure 5-2(a). The testing features are 18 FRA traces from three 750 MVA transformers, including the common and series windings from A, B, and C phases, as plotted in Figure 5-2(b). It can be seen that for the Multiple Layer windings, training features have more obvious oscillations than the testing features in the frequency range 20 kHz to 1000 kHz.

For the Plain Disc windings, the training features are 18 FRA tracs from two 750 MVA and one 1000MVA transformers, including common and series windings from A, B, and C phases. The testing features are 18 traces from two 750 MVA and one 1000MVA transformers, including common and series windings from A, B, and C phases As mentioned, Plain Disc windings' FRA traces can either have a rising or flat trend in the frequency range 20 kHz to 1000 kHz. All the training and testing features own rising trend,

except that 3 testing FRA traces from the A, B, and C phases of one 750 MVA transformer's series windings own flat trend, as shown in Figure 5-3.

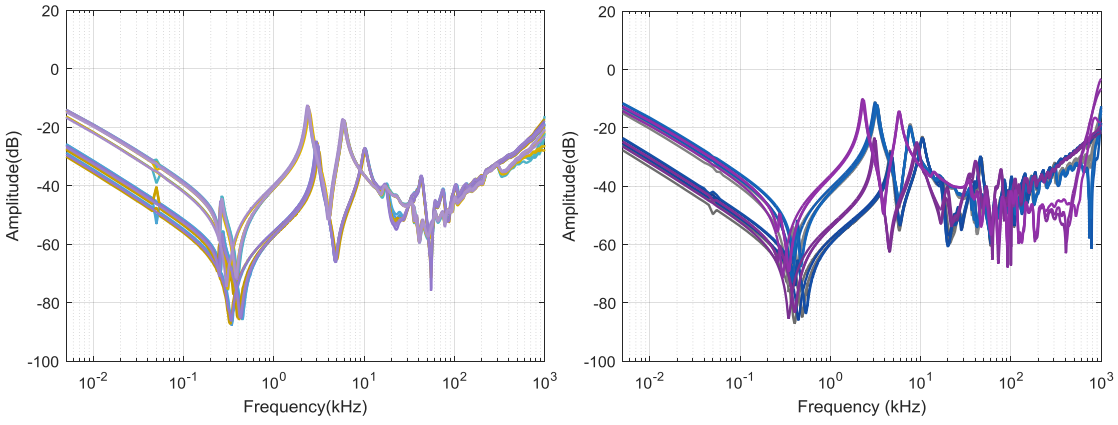
The choice of FRA traces of Interleaved Disc windings is limited, due to its small quantity of frequency response stored in National Grid database. For this winding type, there are 18 training FRA traces from three 750MVA transformers of same manufacture, from A, B, and C phase of common and series windings. The testing features are 9 FRA traces from A, B and C phases of the series windings of three 750 MVA transformers, as plotted in Figure 5-4.

For the Single Helical windings, 6 FRA traces from two 750 MVA transformers are used as training features while 9 FRA traces from two 1000 MVA and one 750MVA transformers are used as testing features, as plotted in Figure 5-5.



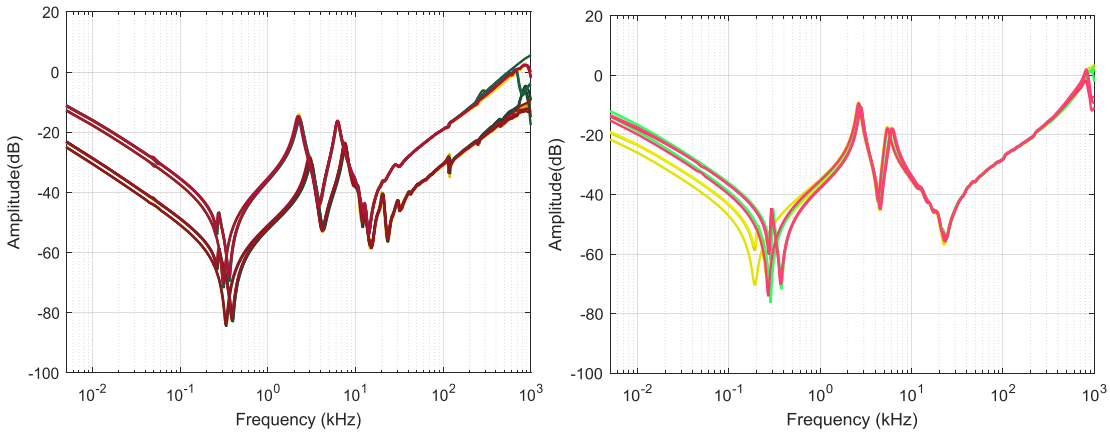
(a) Training Multiple Layer Windings (b) Testing Multiple Layer Windings

Figure 5-2 Multiple Layer Windings used for Training and Testing



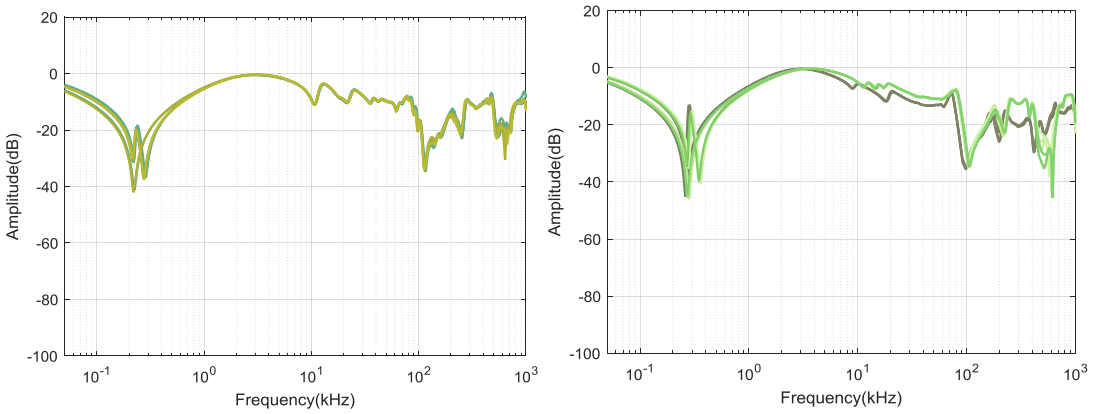
(a) Training Plain Disc Windings (b) Testing Plain Disc Windings

Figure 5-3 Plain Disc Windings used for Training and Testing



(a) Training Interleaved Disc Windings (b) Testing Interleaved Disc Windings

Figure 5-4 Interleaved Disc Windings used for Training and Testing



(a) Training Single Helical Windings (b) Testing Single Helical Windings

Figure 5-5 Single Helical Windings used for Training and Testing

Once the parameters of SVM model are computed, the distances from the training vectors to the hyperplanes of each binary classifier can be calculated. For example, the distances from 12 training multiple layer winding FRA traces (Label 1) to the following 3 binary classifiers, 1 VS 2, 1 VS 3 and 1 VS 4, are tabulated in Table 5-2. For each binary classifier, the closest training vectors to its hyperplane are the support vectors. The distances between the support vectors and the corresponding hyperplane should ideally be 1. In Table 5-2 to Table 5-5, the distances from all 54 training vectors to the hyperplanes of the relevant binary classifiers are listed. The range of these distances is from -1.5 to -1, and 1 to 1.5, which indicates that the distances between every two training vectors from the same winding type are small, due to their high similarity. Considering inevitable calculation errors, all the training vectors, whose distances to any classification hyperplanes range from 0.9990 to 1.0010, are considered as support vectors. All the support vectors are noted with '*' in Table 5-2 to Table 5-5. As stated, there may be more than one support vector for each winding type

It can be seen in Table 5-2 that for Multiple Layer VS Interleaved Disc classifier, traces 8, 9, and 11 are all very close to 1. They are plotted in Figure 5-6. It can be seen that Trace 9 and Trace 11 almost overlap with each other. However, trace 8 differs a lot from the other two traces. This means that not all the support vectors have a very high similarity in shape between each other, though they have the same distance to the classification hyperplane.

Table 5-2 Distances from the Input Multiple Layer Features to the Hyperplane of Binary Classifier

Multiple Layer training traces	1VS2	1VS3	1VS4
1	1.0243	1.0058	0.9997*
2	1.0173	1.1676	1.0013
3	1.0001*	1.0459	1.0027
4	1.4132	1.1755	1.0055
5	1.4006	1.5139	1.0000*
6	1.4301	1.2443	1.0022
7	1.0465	1.0139	1.0030
8	1.0986	1.0004*	1.0094
9	1.0151	1.0002*	1.0000*
10	1.4159	1.1933	1.0074
11	1.5086	1.0002*	1.0594
12	1.4092	1.2991	1.0000*

Table 5-3 Distances from Plain Disc Input Features to the Hyperplane of Binary Classifier

Plain Disc training traces	1VS2	2VS3	2VS4
1	-1.0615	1.0506	1.0078
2	-1.0658	1.1224	1.0003*
3	-1.0385	1.0362	1.0005*
4	-1.1457	1.1119	1.0213
5	-1.0673	1.0813	1.0002*
6	-1.1076	1.0831	1.0189
7	-1.0076	1.0148	1.0281
8	-1.0357	1.0859	1.0126
9	-0.9999*	0.9999*	1.0067
10	-1.1179	1.1222	1.0002*
11	-1.0548	1.0869	1.0000*
12	-1.0781	1.0566	1.0518
13	-1.0358	1.0260	1.0383
14	-1.0600	1.1123	1.0218
15	-1.0002*	0.9999*	1.0311
16	-1.0300	1.0003*	1.0826
17	-1.0000*	1.0216	1.0200
18	-1.0358	1.0146	1.0706

Table 5-4 Distances from Interleaved Disc Input Features To the Hyperplane of Binary Classifier

Interleaved Disc training traces	1VS3	2VS3	3VS4
1	-1.0943	-1.0155	1.0000*
2	-1.1479	-1.0431	0.9995*
3	-1.0674	-1.0614	1.0033
4	-1.0807	-1.4572	1.0067
5	-1.0006*	-1.4724	1.0000*
6	-1.0003*	-1.4434	1.0058
7	-0.9998*	-1.0001*	1.0055
8	-0.9996*	-1.0136	1.0084
9	-1.0614	-1.0054	1.0046
10	-1.0004*	-1.4887	1.0170
11	-1.0218	-1.4380	1.0110
12	-1.1275	-1.4767	1.0183
13	-1.1843	-1.0541	1.0035
14	-1.2479	-1.0548	1.0000*
15	-1.1119	-1.0263	1.0060
16	-1.0390	-1.4586	1.0077
17	-1.0177	-1.4654	1.0017
18	-1.1013	-1.4587	1.0066

Table 5-5 Distances from Single Helical Input Features to the Hyperplane of Binary Classifier

Single Helical training traces	1VS4	2VS4	3VS4
1	-0.9994*	-1.0000*	-0.9996*
2	-1.0160	-1.0340	-1.0138
3	-1.0004*	-1.0005*	-0.9998*
4	-1.0218	-1.0605	-1.0189
5	-1.0042	-1.0177	-1.0073
6	-1.0045	-1.0003*	-1.0007*

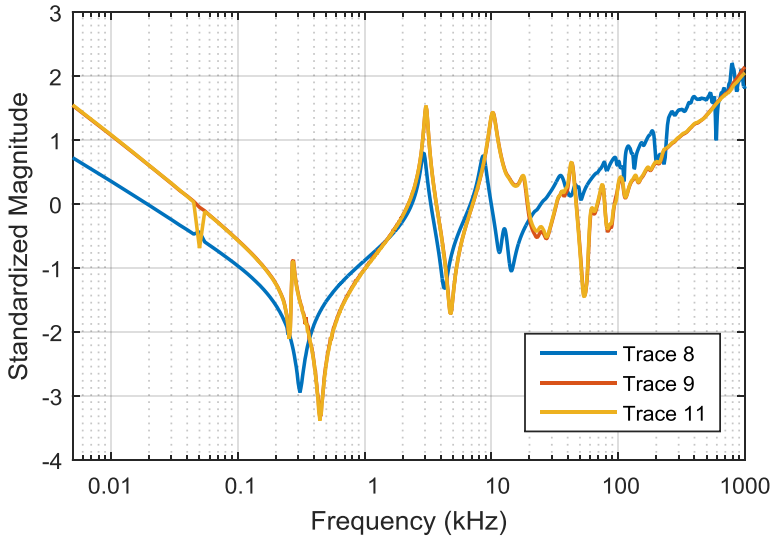


Figure 5-6 Comparison of Multiple Layer Traces 8, 9, and 11

5.3.3 Testing Process

Two testing features are selected as an example to show the process of winding type prediction. One testing feature is of the multiple layer winding, measured on A Phase, common winding, and the other testing feature is of the plain disc winding, measured on A Phase, common winding.

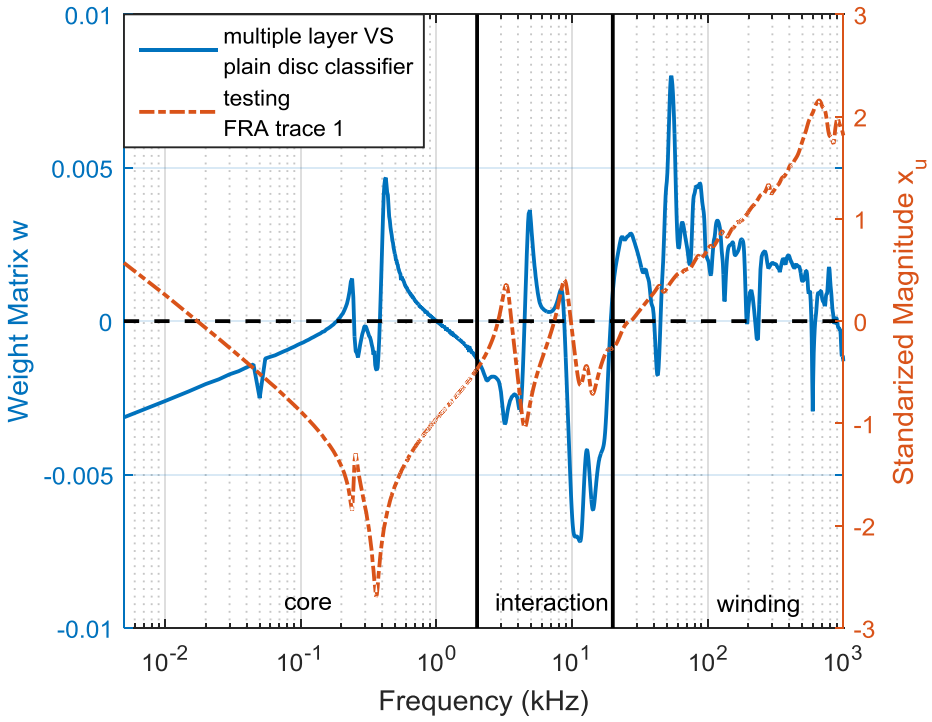
For the Multiple Layer VS Plain Disc classifier, its weight matrix $\boldsymbol{w}$ is plotted against in frequency solid lines in both Figure 5-7(a) and Figure 5-7(b). The bias factor b is -0.1956. The first testing feature $\boldsymbol{x}_{u1}$ to be identified is plotted in Figure 5-7(a) after standardization, and the second testing feature $\boldsymbol{x}_{u2}$ is plotted in Figure 5-7(b) , both in dash lines. The accumulated values $\sum_{i=1}^{n} w_i \cdot x_{ui} + b$ for the two testing features $\boldsymbol{x}_{u1}$ and $\boldsymbol{x}_{u2}$ are plotted against frequency in Figure 5-7(c) and Figure 5-7(d). The final value of $g(\boldsymbol{x}_u)$ can be calculated: $g(\boldsymbol{x}_{u1})$ =1.0954−0.1956=0.8998 and $g(\boldsymbol{x}_{u2})$ =-8392−0.1956=-1.0347. For the first testing feature, the positive number 0.8998 advises that it should be a Multiple Layer winding. For the second testing feature, the negative value suggests that it should be Plain Disc winding. The higher the absolute value of $g(\boldsymbol{x}_u)$ is, the farer the testing feature is from the classification hyperplane, and the higher the prediction confidence is. In Figure 5-7(c), the value of first testing feature $\sum_{i=1}^{n} w_i \cdot x_{ui} + b$ starts from a negative number and then becomes positive in the frequency region 2-20 kHz dominated by winding interaction. Its absolute value, i.e., prediction confidence, grows rapidly from 0.1184 to 0.8998, in the frequency domain controlled by winding type. Such behaviour suggests that in this example, weight factor of the magnitude points on the frequency region from 20 kHz to 1000 kHz plays the most important role in the decision process.

In Figure 5-7(d), the value of second FRA trace $\sum_{i=1}^{n} w_i \cdot x_{ui} + b$ remains negative over whole frequency range. For the frequency region controlled by winding properties, from 20 kHz to 200 kHz, there exists the camel humps characteristics and the absolute value of $g(\boldsymbol{x}_{u2})$ increases to 1.3070 as well as the prediction confidence. From 200 kHz to 1000 kHz, the camel humps disappear and the rising trend with oscillation becomes similar to the characteristics of the multiple layer winding type, thus the absolute value of $g(\boldsymbol{x}_{u2})$ decreases to 1.0350, as well as the prediction confidence. The prediction results from all 6 binary classifiers are listed in Table 5-6. For the first testing vector, the multiple layer winding type wins the most votes, 3, which suggests that the winding corresponding to $\boldsymbol{x}$u1 is has the multiple layer winding type. Similarly, the winding with regard to the second

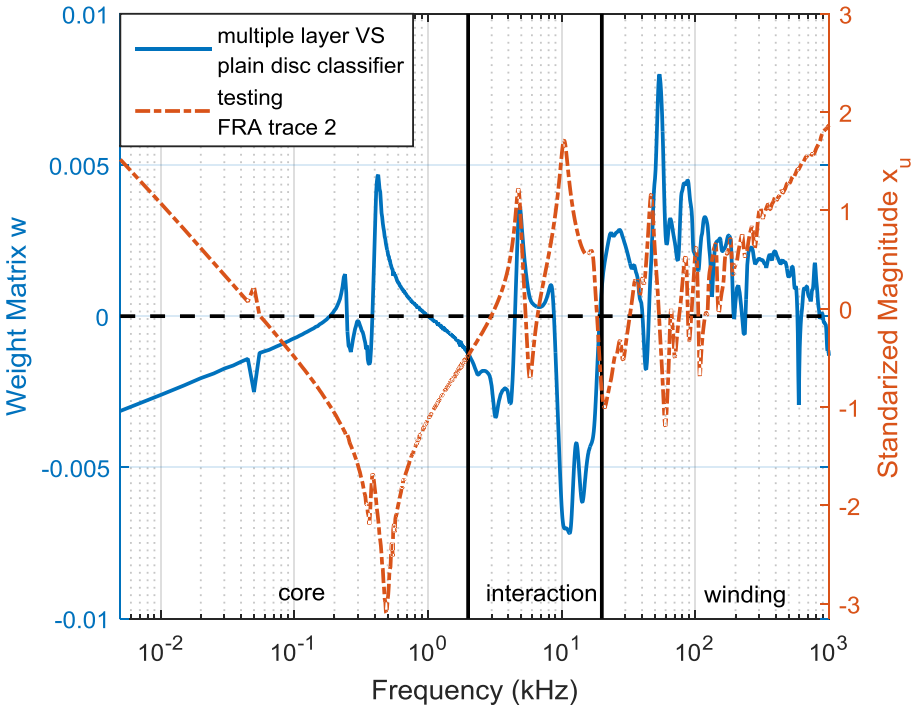
testing feature $\boldsymbol{x}$u2 is classified as having the plain disc winding type. The proposed method successfully identifies the correct winding type in both cases.

Noticeably, even though the testing feature $\boldsymbol{x}_{u1}$ belongs neither to the plain disc winding nor the interleaved winding, the plain disc VS interleaved winding classifier determines it having the interleaved disc winding type. It indicates that the testing feature $\boldsymbol{x}_{u1}$ is more analogous to the feature of an interleaved disc winding than that of a plain disc winding. This conclusion corresponds to the fact that the FRA traces of multiple layer and interleaved disc windings share some similarity in the rising magnitude trend from 20 kHz to 1000 kHz, though their oscillation levels are different.

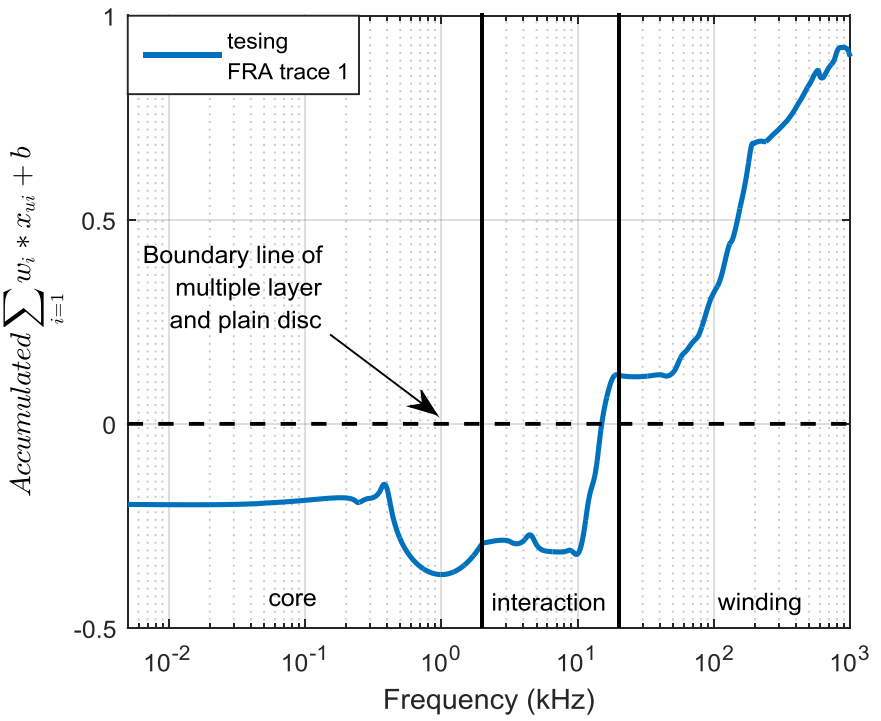
All the 54 testing features are identified correctly, as shown in Table 5-7. For plain disc windings, although all the training FRA traces have rising magnitude trend, the three testing FRA traces with flat trend are classified into the right winding type. The classification result is encouraging, in that the overall characteristics of FRA traces can be identified by considering not the magnitude but the trend. The testing process confirms that the developed SVM model can correctly classify FRA traces that are different from training data and is ready for winding type recognition.



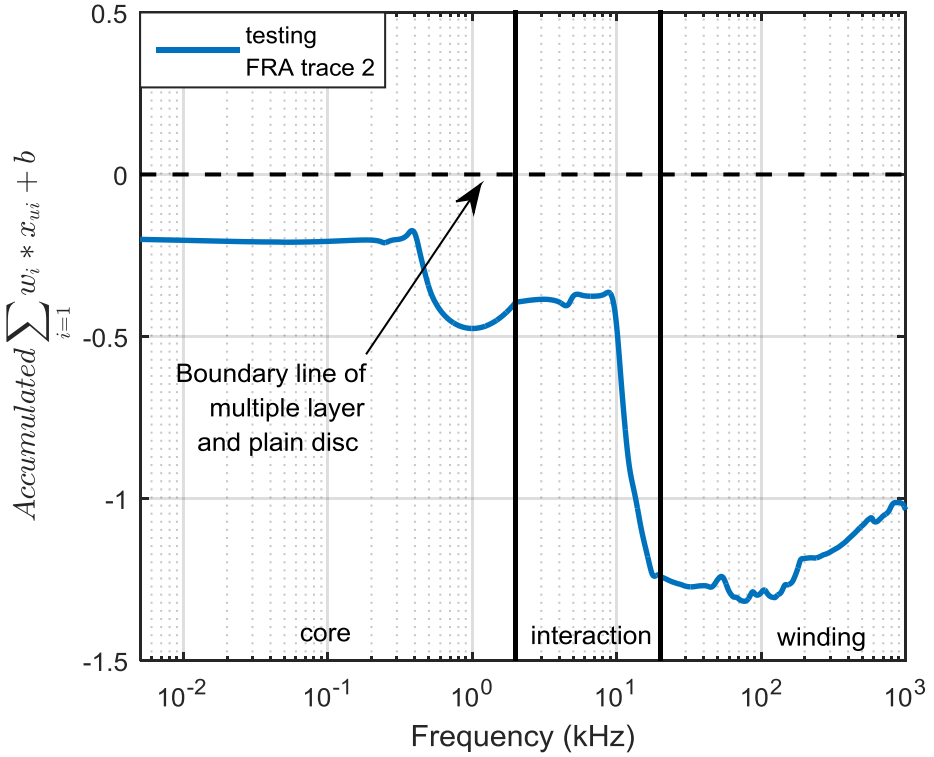
(a) Multiple Layer VS Plain Disc Classifier Weight Matrix and Testing FRA trace 1



(b) Multiple Layer VS Plain Disc Classifier Weight Matrix and Testing FRA trace 2



(c) Accumulated Value $\sum_{i=1}^{n} w_i \cdot x_{ui} + b$ of Testing FRA Trace 1



(d) Accumulated Value $\sum_{i=1}^{n} w_i \cdot x_{ui} + b$ of Testing FRA Trace 2

Figure 5-7 Classification of Testing Feature Example

Table 5-6 SVM Voting Result for Testing Feature x_u

	Testing trace 1					Testing trace 2				
Classifier	$g(x_{u1})$ Value	**1**	**2**	**3**	**4**	$g(x_{u2})$ Value	**1**	**2**	**3**	**4**
1 VS 2	0.8998	✓	×	-	-	-1.0347	×	✓	-	-
1 VS 3	1.1701	✓	-	×	-	1.5322	✓	-	×	-
1 VS 4	1.0469	✓	-	-	×	0.4519	✓	-	-	×
2 VS 3	-0.8589	-	×	✓	-	0.9999	-	✓	×	-
2 VS 4	1.2167	-	✓	-	×	0.7495	-	✓	-	×
3 VS 4	1.0598	-	-	✓	×	0.5478	-	-	✓	×

Table 5-7 Testing Result of SVM Model

Winding type	Training set	Testing set	Accurate number identified	Accuracy
Multiple Layer	12	18	18	100%
Plain Disc	18	18	18	100%
Interleaved Disc	18	9	9	100%
Single Helical	6	9	9	100%

5.3.4 Application

A total of 51 FRA traces, from ten 400/275/13 kV transformers, without winding type information are classified by the SVM model built. The 51 traces are plotted in Figure 5-8 after standardisation. Among all the 51 FRA traces, 30 traces correspond to series windings and the rest are obtained from common windings. In parallel, the winding types of the traces are also manually identified based on expert experience, which are used as the reference to validate the SVM prediction results.

Winding classification by expert experience suggests that 6 of them are of Multiple Layer winding type, 36 of them are of Plain Disc winding type and 9 of them are of Interleaved Disc winding type. The Plain Disc winding type was widely used for 400/275/13 kV auto transformers before 1960s before the Interleaved Disc winding type was invented, thus it takes up the largest number. The FRA traces of Single Helical windings are not included here because the measurement information indicates they belong to tertiary windings and the majority of tertiary windings use the Single Helical winding type. However, this does not mean that the FRA traces of Single Helical windings can be excluded from training data of the SVM model, since the measurement information may sometimes be missing.

Tests show that the SVM model generates the same classification results as the expert experience. As shown in Table 5-8, according to the two classification methods, 6 traces are both classified as Multiple Layer winding type, 36 traces are both classified as Plain Disc winding type and 9 traces are both classified as Interleaved Disc winding type. The proposed method achieves 100% success rate in winding type recognition.

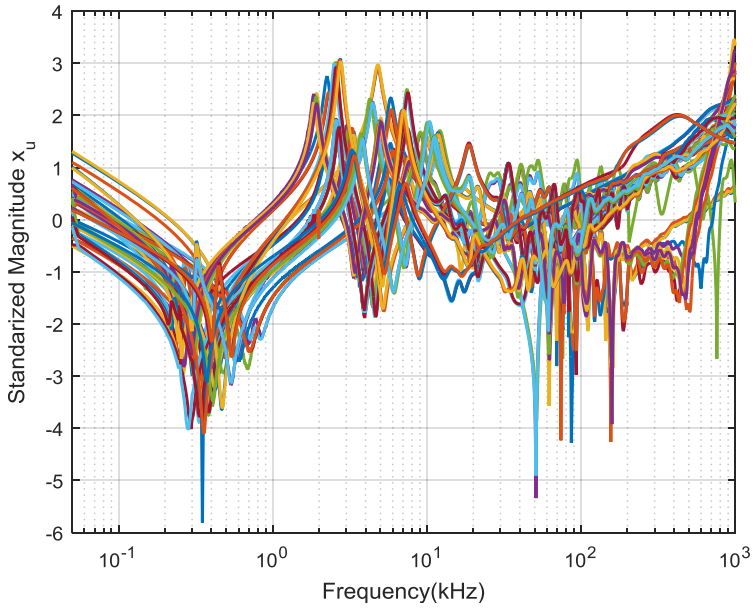


Figure 5-8 51 Unknown FRA Traces to be Classified

Table 5-8 Application of SVM to FRA Traces with Unknown Winding Types

Manually identified winding Type		SVM identified winding type		Corresponding rate
Multiple Layer	6	**Multiple Layer**	6	100%
Plain Disc	36	**Plain Disc**	36	100%
Interleaved Disc	9	**Interleaved Disc**	9	100%

5.4 Sensitivity Study

Training features determine the structure of SVM model and parameters of SVM classifiers. Therefore, the features chosen for training the SVM play an important role on the performance of SVM. It is crucial to ensure the choice of training features does not severely deteriorate the accuracy of winding type recognition. To this end, three sensitivity studies are carried out to investigate the impact of training data selection on the SVM performance. First of all, Multiple Layer winding type features initially used for testing are chosen to train the SVM. The study is devised to see if the SVM design can be easily swayed by training data as these features are similar to another winding type. In SVM theory, training features can be classified into two categories: non-support vectors and support vectors based on their distances to the classification hyperplane. The other two sensitivity studies focus on the impact of removing these two types of vectors from training data on the SVM design and performance.

Figure 5-9 plots the distances from all training features with both Multiple Layer winding type and Plain Disc winding type to the Multiple Layer VS Plain Disc classifier as discussed. It can be seen from the figure that the features with these two different winding types are separated by a large distance of 2, compared to the distances between features with the same winding type, the largest of which is around 0.5. This indicates that the training features with the same winding type are located closely but not necessarily share similarity in their FRA trace shapes.

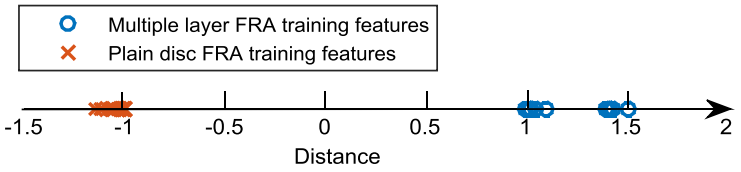


Figure 5-9 Distance Scatter Plot of Multiple Layer and Plain Disc Training Features to Multiple VS Plain Disc Classifier

5.4.1 Exchange Training and Testing Features

As shown in Figure 5-2, the magnitude responses of the training features and testing features for the Multiple Layer winding type are different, in the frequency region between 20 kHz and 1000 kHz. The testing features shown in Figure 5-2 (b) are smoother than the training features shown in Figure 5-2, which means that the testing features are more similar to those of the interleaved disc winding type.

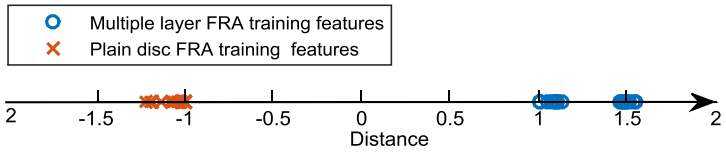
When the training features and testing features are exchanged for the Multiple Layer winding type, the classification result will change, as shown in Table 5-9. Taking multiple Layer features and Plain Disc features for example, Figure 5-10(a) and Figure 5-10(b) show the distances from the new training data and the new testing data to the new Multiple Layer VS Plain Disc classifier, respectively. The classification result during the testing process remains the same for Multiple Layer features and Plain Disc features after the exchange. However, it can be seen from Figure 5-10(c) that one Interleaved Disc winding type testing feature is wrongly classified as Multiple Layer type. The voting results for this Interleaved Disc winding type feature are shown in Table 5-10. Due to the change of Multiple Layer training features, all three classifiers related to this winding type are affected. When such classifiers are applied to the Interleaved Disc winding type feature, the values of g($\boldsymbol{x}$) change. Most importantly, the sign of g($\boldsymbol{x}$) alters for the classifier 1 VS 3, leading to incorrect classifying it as Multiple Layer winding type. Eventually, the feature is falsely identified as Multiple Layer type because it wins the most votes according to the classification criterion in a multiclass SVM problem. However, the classification confidence is not high for this specific classifier, being 0.1886 which is smaller than 0.3717 when the original SVM classifier is used. Consequently, there is a high probability for it to cross the classification hyperplane, once the training features change again. The exchange of training and testing features for Multiple Layer winding reduces the dissimilarity between the Multiple Layer training features and Interleaved Disc training features.

Table 5-9 Sensitivity Study: Exchange of Training and Testing Data

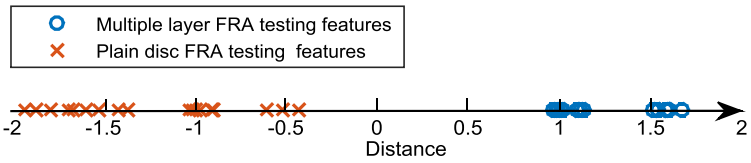
Winding Type	Training set	Testing set	Accurate Number identified	Accuracy
Multiple Layer	18	12	12	100%
Plain Disc	18	18	18	100%
Interleaved Disc	18	9	8	89%
Single Helical	6	9	9	100%

Table 5-10 Voting Results for Wrongly Classified Testing Feature

	Before exchange		After exchange	
Classifier	**g(x) Value**	**decision**	**g(x) Value**	**decision**
1 VS 2	0.6421	1	0.7914	1
1 VS 3	-0.3717	3	0.1886	1
1 VS 4	0.9198	1	0.9332	1
2 VS 3	-0.8038	3	-0.8038	3
2 VS 4	1.4215	2	1.4215	2
3 VS 4	0.8386	3	0.8386	3



(a) Distance Scatter Plot of Multiple Layer and Plain Disc Training Features to Multiple Layer VS Plain Disc Classifier



(b) Distance Scatter Plot of Multiple Layer and Plain Disc Testing Features to Multiple Layer VS Plain Disc Classifier

(c) Distance Scatter Plot of Multiple Layer and Interleaved Disc Testing Features to Multiple Layer VS Interleaved Disc Classifier

Figure 5-10 Distance Scatter Plot after Exchange of Training and Testing Multiple Layer Features

When the training data and testing data from the four winding types are swapped, the accuracy of the SVM prediction model drops to 67%.

5.4.2 Delete Non-Support Vectors

This section examines the impact of deleting non-support vectors from training data on the performance of the proposed SVM. After the SVM model is built, the support vectors can be identified. All the support vectors are noted by '*' in Table 5-2 to Table 5-5. Instead of using all features, this time only the support vectors are used for training the SVM. The result shows that the newly built SVM model is identical to the original SVM model.

Figure 5-11 shows the change in the quantity of training features. Only 27 FRA traces are used for the newly constructed SVM model, which is half of the amount of the original 54 training FRA traces. If any feature of the rest 27 training features is deleted, the parameters of the original SVM model will change, though the classification results might remain the same for the testing features.

A comparison is done in Figure 5-11 on the weight matrixes of Multiple Layer VS plain Disc classifier, before and after the non-support vectors are deleted. It shows that the two weight matrixes exactly overlap with each other. The correlation coefficient and average difference are used to describe the similarity between the parameters of SVM before and after the change of training data. All the correlation coefficients for the 6 binary classifiers are 1.0000, with 0 mean differences. The biases of the 6 binary classifiers also remain unchanged, with a difference value of 0. The bias of 1 VS 2 classifier before and after is listed in Table 5-11.

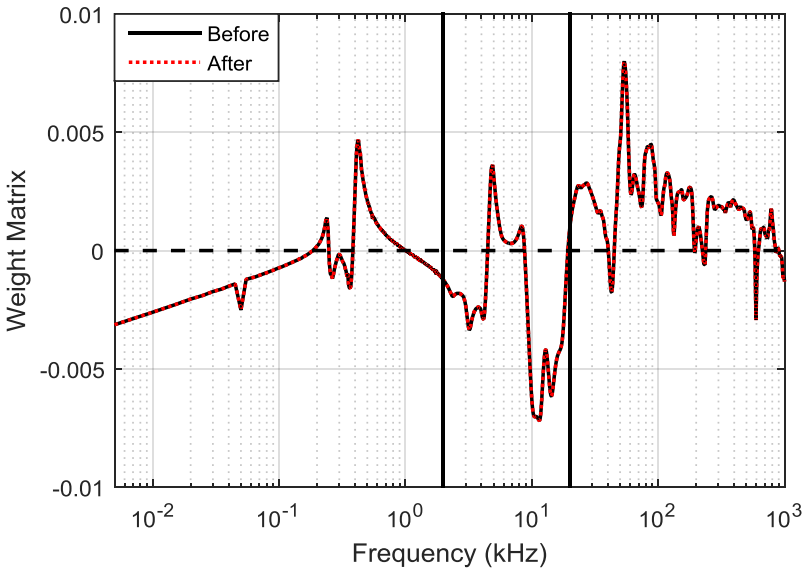


Figure 5-11 Comparison of 1 VS 2 Classifier before and after Deletion of Non-Support Vectors

Table 5-11 Bias Factor before and after Exchange

	Original	New
Bias factor	-0.1956	-0.1956

5.4.3 Delete Support Vectors

This section investigates the effect of deletion of support vectors from training data on the SVM performance of winding type recognition. The sensitivity study is carried out on the Multiple Layer VS Plain Disc winding type binary classifier.

As tabulated in Table 5-2, 12 FRA traces of multiple layer type windings are used as training vectors in the original SVM model. These 12 FRA traces are from A, B and C phases of two transformers. It can be found that the support vector for this binary classifier from multiple layer winding type is the 3rd training vector in Table 5-2. FRA traces from three phases of the same transformer series/common winding are similar in shape and their features are close to each other. To accelerate and simplify the process, the training features corresponding to the other two phases are removed together with the support vector to be deleted. After deletion the SVM is trained with the rest features and a new support vector can be identified. Performance of the newly built SVM is assessed. Afterwards the new support vector is deleted together with two features from the same transformer series/common winding. Such process to remove support vectors iterates until only 3 vectors from the same transformer winding are left. The weight matrices in all the iterations are plotted in Figure 5-12. The values of the corresponding biases and the performance of SVM are tabulated in Table 5-12 and Table 5-13, respectively.

On the first iteration, the weight matrix deviates slightly from the original. The deleted features are from a common winding. The SVM model functions well, identifying all testing features correctly. This suggests that the deleted support vectors are similar to a number of remaining training vectors and their deletion does affect the performance of the SVM. On the second iteration, the weight matrix differs from the previous weight matrix considerably. The deleted features are also from the common winding and the remaining features are all from series windings. This leads to the significant change in the weight matrix because it is determined only by the series windings rather than the combination of common and series windings. The FRA trace of multiple layer series windings is similar to that of interleaved windings, so the SVM trained only by the former can have difficulties to distinguish these two winding types. At this stage 9 multiple layer windings are wrongly classified as interleaved disc windings and the prediction accuracy drops to 83%. On the third iteration, only 3 features from series windings are used as training vectors. The classification accuracy remains at 83%. Slight changes can be observed in the weight matrix. This is because FRA traces obtained from series windings share great similarity.

For the 9 wrongly classified multiple layer windings, the g(x) values of Multiple Layer VS Interleaved Disc classifier are listed in Table 5-14. It can be seen from the table that the sign of g(x) changes from positive to negative after the first iteration. Hence, the SVM wrongly classify the features as interleaved disc type windings.

The study demonstrates the importance of the support vectors in the training of SVMs. The performance of the proposed winding type recognition method depends on if the suitable support vectors are included in training. However, non-support vectors are indispensable because the support vector is a relative concept and there always exists a support vector, but not necessarily a suitable one. It is found that completely removing training features from either common or series winding can deteriorate the performance of the SVM.

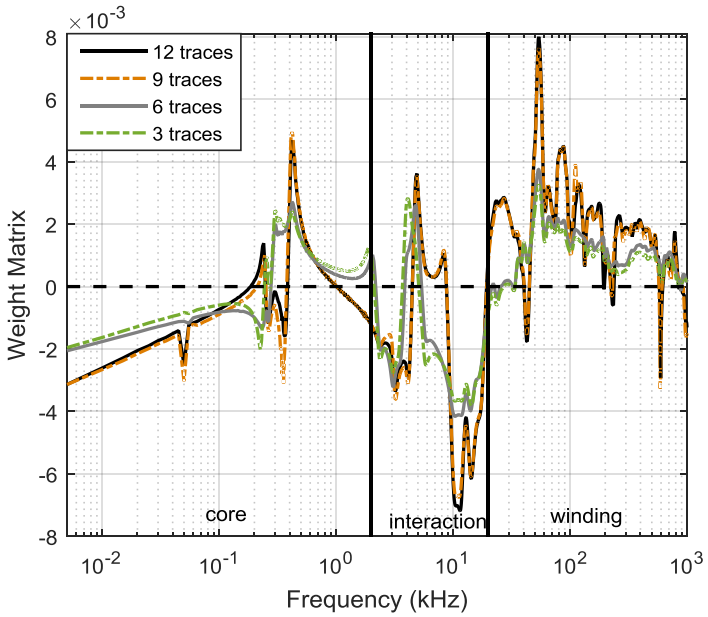


Figure 5-12 Weight Matrix Changing Process

Table 5-12 Bias Factors Changing Process

	Original	Step 1	Step 2	Step 3
Bias factor	-0.1956	-0.2226	-0.3663	-0.3235

Table 5-13 Accuracy Changing Process

	Training set	Testing set	Accurate number	Accuracy
Original	12	54	54	100%
Step 1	9	54	54	100%
Step 2	6	54	45	83%
Step 3	3	54	45	83%

Table 5-14 Multiple Layer VS Interleaved Disc Classifier g(x) Value of Wrongly Classified Multiple Layer Features Changing Process

Multiple Layer features	Original	Step 1	Step 2	Step 3
1	1.1701	1.1699	-2.4613	-2.1110
2	1.4254	1.4253	-2.2660	-1.9860
3	1.1953	1.1952	-2.4509	-2.0963
4	1.0599	1.0598	-2.1795	-2.0031
5	1.5126	1.5126	-1.8913	-1.8511
6	1.0774	1.0774	-2.1864	-2.0041
7	1.1416	1.1414	-2.3361	-2.1157
8	1.5241	1.5241	-2.0551	-1.9632
9	1.0846	1.0843	-2.3573	-2.1289

5.5 Conclusion

In this study, a novel SVM-based method is proposed for transformer winding type recognition using FRA data. The SVM model is built with FRA traces of 400/275/13kV auto transformers from the UK's National Grid Database. This model is trained using FRA traces with design information, and later tested by different FRA traces with a 100% accuracy rate. Examples are given to show the prediction process and to analyze the influence of different frequency regions on the final classification result. The frequency region from 20 kHz to 1000 kHz controlled by the winding structure plays an important role in the winding type recognition. Subsequently, the SVM is applied to 51 FRA traces with unknown winding type, and the prediction result is validated with the classification made by expert experience. The proposed method successfully identifies the correct winding type in all cases, which demonstrates the satisfactory performance of the SVM-based method. Sensitivity studies are carried out to investigate the impact of training data selection on the performance of SVM. It is concluded that the performance is mainly affected by support vectors. When small changes occur to the support vectors, the SVM model might still produce correct prediction results for the original testing vectors. However, the prediction accuracy will drop once the support vectors change significantly. It is important to identify and include the critical FRA traces in the training data for accurate identification. To ensure the suitable support vectors are used to build the classification model, it is suggested to use as many training FRA traces as possible. Meanwhile, expert judgement and practical experience can be exercised and referred to when training the SVM model

With the changes in the voltage ratio, power rating, winding types, etc., the frequency region dominated by the properties of winding-under-test, may be different. The linear SVM model built in this paper would be only applicable with confidence to the 400/275/13 kV autotransformers with different power ratings (500MVA, 750MVA and 1000 MVA). Further study should be carried out on winding type recognition of transformers with a variety of voltage levels, such as a mixed database of 275/33 kV, 275/132/13 kV, and 400/275/13 kV transformers.

Chapter 6 Identification of Winding Construction Types by Unsupervised Machine Learning Method

6.1 Introduction

Over the years utilities have accumulated a large number of measured frequency responses whilst the transformers' design information such as winding construction types may or may not be necessarily known.

Different winding types own different equivalent electrical parameters, i.e. capacitance and inductance. For instance, the Interleaved Disc winding has higher series capacitance whilst the Plain Disc winding has lower series capacitance. Resonance and anti-resonance are caused by the conjunct effect of inductance and capacitance. They normally appear as local maximum and minimum, i.e. peak and trough on FRA traces. The variation in their equivalent electrical parameters, especially winding series capacitance, leads to the different characteristics on their FRA traces. As a result, unalike features are caused at specific frequency regions of frequency response, especially the high frequency region which is dominated by winding-under-test. Consequently it is possible to correlate FRA characteristics with known design features.

Same faults may occur to same winding types and result in similar distortions of FRA features. In the UK forensic examination are routinely carried out on retired transformers, and knowledge can then be acquired for asset management. Therefore, any technique which helps to identify the unknown winding types of transformers is desirable when considering the fact that a significant number of transformers' design information is unknown to the utilities.

The unsupervised machine learning method, Hierarchical Clustering, is applied on the FRA traces of 400/275/13 kV autotransformers in this chapter for the identification of winding types. The frequency responses investigated in this chapter are from five winding types, i.e. Intershielded Disc winding, Plain Disc winding, Interleaved Disc winding, Multiple Layer winding, and Single Helical winding. The Hierarchical Clustering groups similar frequency responses together according to the distances between them. Clustering using different frequency regions is also investigated, considering the fact that different frequency regions are controlled by different physical factors.

With the technique employed in this study, in combination with expertise knowledge and forensic information accumulated, the utility will be able to develop a strategy to manage similar type of transformers and achieve effective asset management. Besides, once the winding construction type is known, it is also helpful for the fault diagnosis of transformer windings.

6.2 Hierarchical Clustering

Hierarchical clustering method is an unsupervised machine learning algorithm which groups similar objects together as introduced in Chapter 2. Different from supervised machine learning method, it does not require any training process using input data with known classifications. For a given group of observations, each observation is initially treated as a separate cluster, and the algorithm repeatedly identifies the closest two clusters and merges them, until all the clusters are merged into one cluster. An example using Hierarchical clustering method to cluster six dataset will be given.

As shown in Table 6-1, six datasets, A – F, the distances between every two datasets are listed in Table 6-1 (a). Hierarchical Clustering algorithm finds the smallest distance in this distance matrix, which is 10 between E and F in Table 6-1 (a), thus E and F are clustered together first. Then, the smaller distance between A to E and A to F is adopted as the distance from A to the new cluster {E, F} which is 77. This way of calculating the new distance between the newly merged cluster and other clusters is called Single Linkage. The new distance matrix is recalculated in Table 6-1 (b), where the smallest distance is 16 between A and B. The algorithm keeps merging clusters until only one cluster exists, as shown in Table 6-1 (c)-(d).

Table 6-1 Distance Matrix Example

	A	B	C	D	E
B	16				
C	47	37			
D	72	57	40		
E	77	65	30	31	
F	79	66	35	23	10

(a)

	A	B	C	D
B	16			
C	47	37		
D	72	57	40	
E+F	77	65	30	23

(b)

	A+B	C	D
C	37		
D	57	40	
E+F	65	30	23

(c)

	A+B	C
C	37	
D+E+F	57	30

(d)

	A+B
C+D+E+F	37

(e)

Dendrogram in Figure 6-1 shows the corresponding process of merging. The horizontal axis is the dataset, and vertical axial is the distance between the datasets.

The clustering result can be decided either by the pre-set final cluster number or by the cut off distance. When the final cluster number is set as 2, then two clusters {A, B} and {C, D, E, F} are divided. When the cut off distance is set at the height of the dash line in Figure 6-1, clustering result remains the same.

In fact, the distance between two observations reflects the similarity between them. Euclidean Distance is used in this study. Weighted Linkage uses the average of two distances between a specific cluster and the two observations to be merged, as the new distance. In the given example in Table 6-1, using weighted linkage, the distance of newly merged cluster {E,F} to cluster A is (77+79)/2=78.

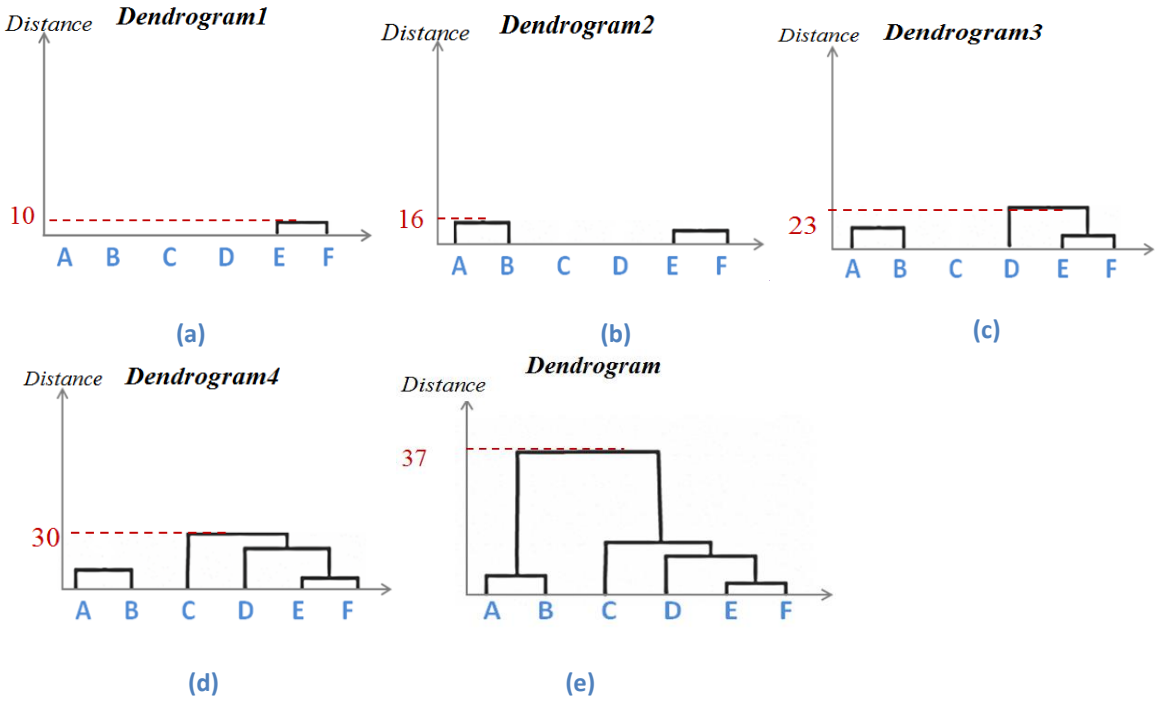


Figure 6-1 Dendrogram of Datasets A – F in Figure 6-1

6.3 Application of Hierarchical Clustering

As illustrated in Figure 3-9, Single Helical winding has a higher magnitude roughly from 5 Hz to 100 kHz. For the frequency region of 20 kHz to 1000 kHz, which is believed to be dominated by winding properties, Multiple Layer winding's FRA magnitude trace rises about 20dB per decade with obvious oscillations. With a higher series capacitance, Interleaved Disc winding has the same rising speed but the trace is smoother. Camel

humps usually appear on Plain Disc winding's frequency response. The FRA characteristic of Intershielded Disc winding is not unique; it looks like Plain Disc winding. For both winding types, the smooth rising trend would occur after camel humps feature.

Different frequency regions of frequency response are influenced by different transformer parts. For the 400/275/13 kV autotransformers investigated, transformer core, inter-winding influence, properties of winding-under-test and measurement setup are believed to dominate the bandwidths < 2 kHz, 2-20 kHz, 20-1000 kHz and >1000 kHz respectively. Actually, the boundary points of bandwidth, 2 kHz, 20 kHz, and 1000 kHz, are empirical. The concerned FRA characteristic in one frequency region can move up or down to another frequency region when the voltage and power ratings of winding or winding type changes. For clarity, to avoid the influence of measurement setup, the measured FRA data up to 200 kHz, are used to test the unsupervised machine learning methods in this study.

All together 28 measured FRA traces from 400/275/13 kV autotransformers are studied, and their winding type information are listed in Table 3-11 and FRA traces are plotted in logarithmic scale in Figure 3-22. There are 4 frequency responses from Intershielded Disc winding with label {1-4}, 6 frequency responses from Plain Disc winding with label {5-10}, 4 frequency responses from Multiple Layer winding with label {11-14}, 6 frequency responses from Interleaved Disc winding with label {15-20}, and 8 frequency responses from Single Helical winding with label {21-28}. The frequency responses {5-7, 12, 13, 15-18} are from series windings, the frequency responses {1-4, 8-10, 13, 14, 19, 20} are from common windings, and the frequency response {21-28} are from tertiary windings.

For each frequency response, there are 1120 data points for the frequency range from 5 Hz to 200 kHz. Euclidean Distance and Weighted Linkage are used for clustering. The Euclidean Distance between two frequency responses is calculated as:

$$d = \sqrt{\sum_{i=1}^{n=m} (X_a^i - X_b^i)^2} \qquad \text{Equation 6-1}$$

where $X_a{}^i$ and $X_b{}^i$ are the magnitudes at frequency point i on FRA traces a and b, in the unit of dB.

For Weighted Linkage, cluster K is produced by merging two clusters M and N, the distance between the new cluster K and another cluster L, $d(K, L)$, can be calculated as the

mean value of the distances $d(M,L)$ between Clusters M and L and the distance $d(N,L)$ between Clusters N and L:

$$d(K,L) = \frac{d(M,L) + d(N,L)}{2} \qquad \text{Equation 6-2}$$

The clustering Dendrogram using Euclidean Distance and Weighted Linkage is shown in Figure 6-2.

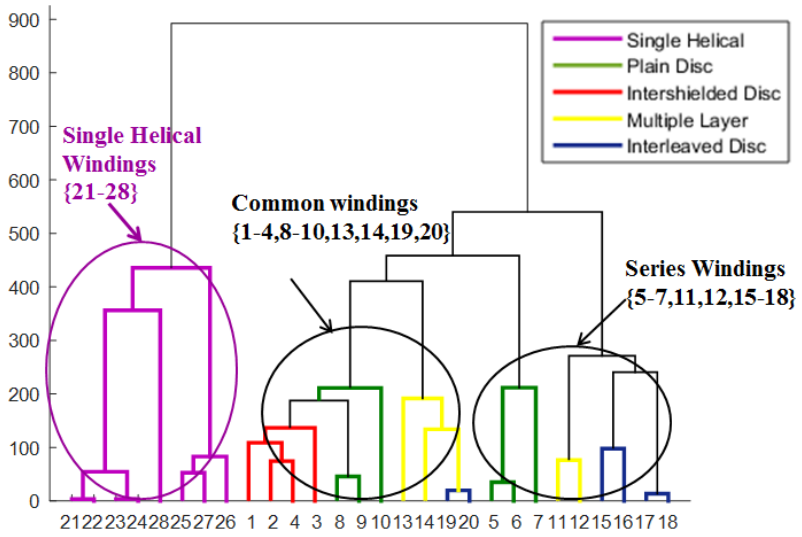


Figure 6-2 5 Hz to 200 kHz Clustering Dendrogram using Euclidean Distance Weighted Linkage

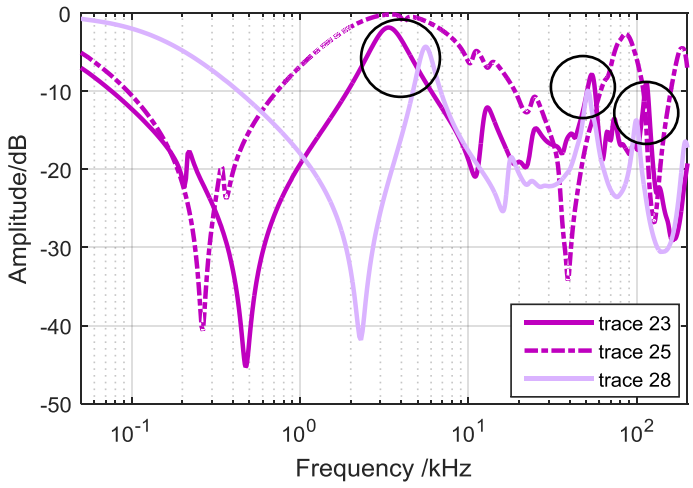
Suppose the final cluster number is set as 2, the left cluster contains 8 traces of Single Helical Winding type, and the right cluster contains traces from the rest winding types.

For the left hand side cluster, 8 traces {21- 28} from Single Helical winding are grouped due to their high magnitude compared with other frequency response at most frequencies in the whole range. The frequency responses {21-24} are tested on the same transformer, and they are clustered together firstly. The frequency responses {25-27}, with distances around 100, are from another transformer and thus they are clustered together. The frequency response {28} is from a third transformer, and its distance to the other Single Helical Traces from other transformers is about several hundreds. The frequency responses {23, 25, 28} are plotted in Figure 6-3 (a). Though the three frequency response all have high magnitude compared with the series and common windings, there still exist obvious difference in them. The locations of their antiresonance, which drop to around -40dB, are 0.265 kHz, 0.475 kHz and 2.3 kHz separately. The resonance of frequency response 28, circled in solid black lines, located close to that of frequency response 23. It can be seen that the similarity between the shapes of frequency responses 23 and 28 is higher than that

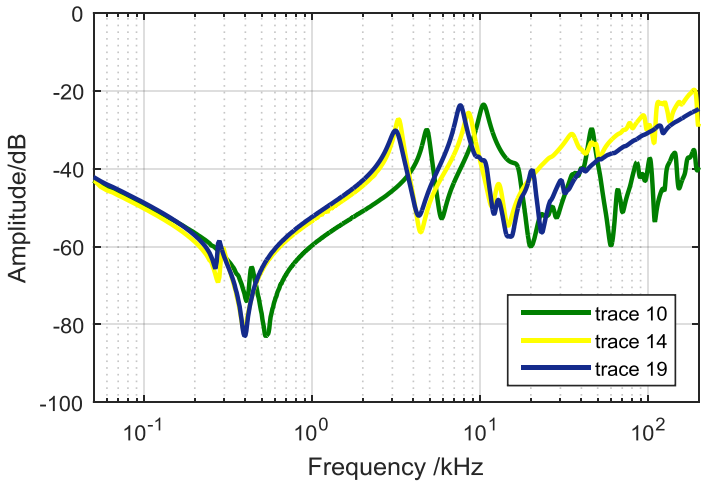
between frequency responses 25 and 28. Thus frequency response 28 is clustered in to {21-24} rather than {25-27}, leading to the clustering result shown in Figure 6-2.

For the right hand side cluster, all common windings {1-4, 8-10, 13, 14, 19, 20} are cluster together into a sub-group due to similarity between them. Series windings {5-7} are clustered into the sub-group of common windings afterwards. Finally the rest of series windings traces {11, 12, 15-18} are merged with the common windings and the series windings mentioned above. The common windings 9, 10 and 14 are plotted together in Figure 6-3 (b) to investigate the clustering process. Before 0.4 kHz, their FRA magnitude spectra almost overlap with each other. Before about 3 kHz, they all have a rising trend, and the magnitude difference among themselves is much smaller compared with the difference between them and the series/tertiary windings. Therefore, the common windings are clustered together, mainly due to their similarity in the low frequency region before several kHz.

However, this result is not desirable, since it does not distinguish winding types, but perhaps overemphasized the differences in magnitude.



(a) Comparison between Traces 23, 25, and 28



(b) Comparison between Traces 10,14,19

Figure 6-3 Examples of Frequency Responses from 400/275/13 kV Transformers

As known the frequency range 5 Hz to 200 kHz can be split into three regions and the relation among the distances exists as:

$$d^2 = d_1^2 + d_2^2 + d_3^2$$ Equation 6-3

where d, d1, d2, d3 are the Euclidean distances of 1120 points from 5 Hz to 200 kHz, 400 points from 5 Hz to 2 kHz, 360 points from 2 kHz to 20 kHz and 360 points from 20 kHz to 200 kHz. It should be noted that the distance between two frequency responses is not only affected by the distance at each measured frequency but also the amount of measured frequency points.

It is worth to study the distance in each frequency region to understand their individual contribution to the final clustering classification results.

6.3.1 Bandwidth Dominated by Core and by Winding Interaction

Using frequency response in the 5 Hz – 2 kHz region which is dominated by the core, the dendrogram is shown in Figure 6-4. Set the final cluster number as 3 and a clear clustering result can be obtained, i.e. the tertiary windings on the left, common windings in the middle and series windings cluster on the right. Clearly d1 are sensitive to the voltage levels. i.e. the number of turns, and is good at classifying tertiary, common and series windings. The classification result is reasonable as frequency response is dominated by the equivalent core in this frequency region and the core inductance is proportional to the voltage level (number of turns) squared.

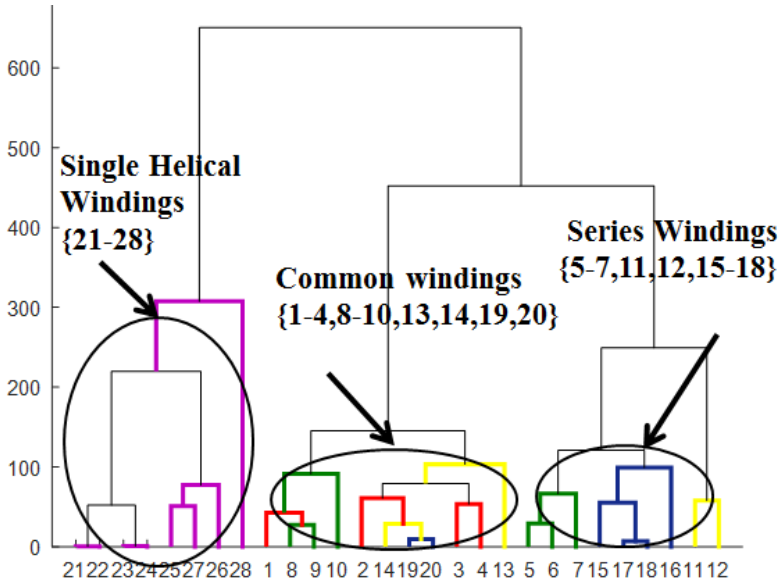


Figure 6-4 5 Hz to 2 kHz Clustering Dendrogram using Euclidean Distance Weighted Linkage

The winding interaction dominated frequency region is from 2 kHz to 20 kHz. Due to its complexity, the clustering results, using frequency response at this frequency region only, are messy, as shown in Figure 6-5. And thus discussion will not be included in this section. Due to frequency response 3 is from Intershielded Disc winding while frequency response 10 belongs to Plain Disc winding, the line connecting them in Figure 6-5(b) are in black colour.

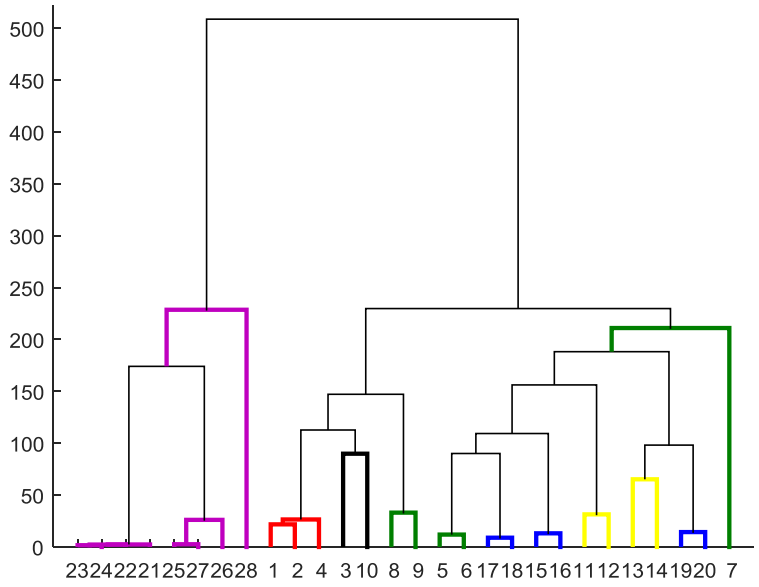


Figure 6-5 2 kHz to 20 kHz Clustering Dendrogram using Euclidean Distance Weighted Linkage

6.3.2 Bandwidth Dominated by Winding Properties

For winding-property controlled frequency region 20 Hz to 200 kHz, the clustering results for the common and series windings are shown in Figure 6-6. Two clusters are clearly

formed. Multiple Layer and Interleaved Disc windings are clustered together whilst Plain Disc and Intershielded Disc windings are gathered together as the other cluster. Multiple Layer windings and Interleaved Disc windings are both the type of windings with high series capacitance, and Plain Disc windings and Intershielded Disc windings are the type of windings with low series capacitance.

Discussions on d3 are detailed here. For the left hand side cluster {11-20}, traces {14, 15, 16, 19, 20} are assembled as one small group, and traces {11, 12, 13, 17, 18} are assembled as another small group. Interleaved Disc traces {15, 16} are merged with a distance as tiny as 4, and Interleaved Disc traces {19, 20} are merged with a distance of 10. The distance between small cluster {15, 16} and cluster {19,20} rises to 68, mainly due to the large distances between Multiple Layer trace 14 and Interleaved Disc traces {19, 20}, which are 104 and 108. Because of the same reason, the distance between cluster {14, 15, 16, 19, 20} and cluster {11, 12, 13, 17, 18} rises to 150, and this distance 150 is largely influenced by the distance between the two most dissimilar traces members in the cluster. The final distance between the last two clusters in Figure 6-6 of the 20 trance is 356. Actually the distance between clusters is not only influenced by the largest distance (which is 503 between Plain Disc winding {5} and Interleaved Disc winding {17}), but also every distance between every two traces. In the right hand side cluster {1-10}, the Intershielded Disc windings are clustered firstly, and then the Plain Disc windings are merged with the Intershielded Disc windings.

In summary, from 20 kHz to 200 kHz, the difference in winding series capacitance causes the magnitude difference in the FRA traces and leads to this clustering result.

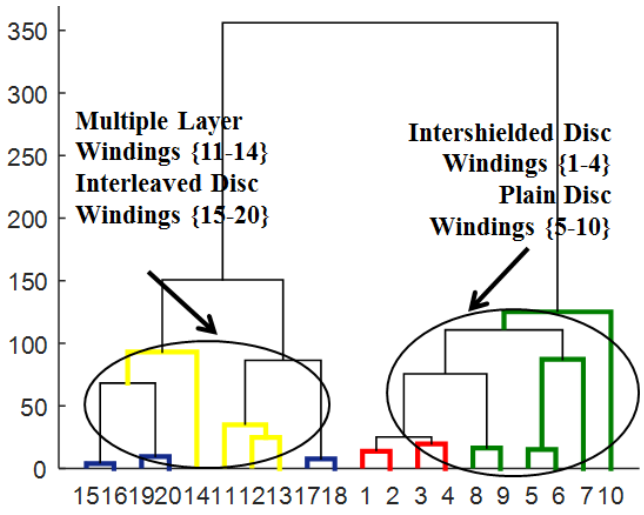


Figure 6-6 20 kHz to 200 kHz Clustering Dendrogram using Euclidean Distance Weighted Linkage

6.4 Summary

In this chapter, the unsupervised machine learning method Hierarchical Clustering is used to process 28 frequency responses measured on 400/275/13 kV autotransformers, with the aim to identify winding types.

Initially each frequency response is viewed as a cluster. The frequency responses are clustered according to the distances between each other. The 28 clusters are merged until the predefined final cluster number or the cut off distance is met.

When using the frequency response from frequency range 5 Hz to 200 kHz, 8 Single Helical Windings can be easily picked out due to their high magnitude in most frequencies, especially at the low frequency region of 5 Hz to 100 kHz. The frequency responses measured on common and series windings can also be roughly clustered together.

Due to the particular FRA measurement device used, for the frequency region <20 kHz, a larger amount of data points, 760, is used; this is in contrast to the amount of data points, 360, at the higher frequency region from 20 kHz to 200 kHz. This means more weight is given to the low bandwidth, and thus the difference between Single Helical windings and other windings are emphasised. The weight can be adjusted according to different needs, by reproducing frequency response using transfer function estimation method.

The frequency regions 5 Hz to 2 kHz and 20 kHz to 200 kHz, which are controlled by the core and the property of winding-under-test, are investigated. Using the frequency region controlled by the core, which is 5 Hz to 2 kHz, the tertiary, common, and series windings can be clustered and identified correctly. Using the frequency region controlled by the winding properties, which is 20 kHz to 200 kHz, the windings with high series capacitance (Multiple Layer winding and Interleaved Disc winding) can be clustered together, while the windings with low series capacitance (Plain Disc winding and Intershielded Disc winding) can be clustered together.

When a proper clustering method is adopted, the frequency response with unknown winding type can be correctly identified by being clustered into a group of frequency responses together with a frequency response with known winding type. Once the frequency responses are in the same cluster, it is in default to think the windings should share the same winding construction type.

Chapter 7 Conclusions and Future Work

7.1 Conclusions

FRA technique has been widely applied in the diagnosis of transformer winding related mechanical faults. This PhD thesis focuses on the study of a group of UK National Grid 400/275/13 kV autotransformers and their FRA measurement records from the FRA database, to develop the methods for transfer function estimation and winding type identification.

The current FRA diagnostic method for winding mechanical integrity assessment is largely subjective. As the best practice in the utilities; experts rely on their understating of FRA curves, transformer design, prior knowledge of system short-circuits or incident reports during transportation/installation, to make diagnostic decisions; and these decisions are often yes/no for a question such as if there is winding deformation/displacement or not. Different from the subjective diagnosis according to experts' experience, objective diagnosis attempts to use mathematical parameters to indicate and quantify the differences between the reference and measured frequency responses. Other than numeric indices, one of the mathematic methods is transfer function representation, which uses either real or complex zeros and poles and a constant coefficient to mathematically describe the frequency response. The changes in the frequency responses can be reflected by the alteration of the parameters of transfer functions, and objective diagnostic rules can be then made according to the relationships between these mathematic parameters and their changes with the type, severity and location of the winding movement.

In the UK National Grid FRA database, there exists a large amount of FRA measurement traces, whose transformer design information are not known to the utility, such as the winding construction types. Knowing the winding construction type is helpful for the diagnosis of transformer winding through FRA technique, as well as for the effective asset management. The specific frequency regions of FRA traces are influenced by the different factors including the winding construction type, though the exact and specific boundary of this frequency region can be influenced by the transformer's voltage ratio, power rating and winding type used. Different frequency response features are caused by different

winding construction. Thus it is feasible to identify the winding construction types according to the characteristics of frequency responses.

The main work conducted in this PhD study can be divided into two parts. One part is on the development of transfer function estimation methods for accurate representation of frequency responses. Two estimation methods are developed. The other part is on the winding type classification using the measured FRA data. The supervised machine learning method, Support Vector Machines (SVM) and the unsupervised machine learning method, Hierarchical Clustering, are employed. Detailed conclusions will be given in the following sections.

7.1.1 Conclusions on Transfer Function estimation

Frequency responses of 400/275/13 kV autotransformers from 5 Hz to 200 kHz are used for the demonstration of developed methodologies for transfer function estimation.

The first method, Feature Extraction Method, developed in this PhD study is based on the Non Linear Least method. The key information, complex zeros and poles are extracted from frequency regions to construct a Feature Transfer function, and the Difference Transfer Function is used to correct the difference between the Feature Transfer Function and the measured frequency responses. A MATLAB program is developed to automatically process the input frequency responses. The proposed method has been tested on eight transformers. For each transformer, 6 frequency response from both series and common windings of A, B, and C phases, are processed, which means a total amount of 48 FRA traces are given their transfer function representation. Four winding types are included, i.e. Multiple Layer winding, Plain Disc winding type, Interleaved Disc winding type and Intershielded Disc winding type. For each FRA trace, the time length of running the estimation program is about from 10 to 15 seconds, on a computer with Inter-i7-4770 3.4GHz processor, 16 GB RAM and 64-bit operating system. This method is proved to be very accurate and efficient.

The estimation result is well matched with the measured FRA data for both magnitude and phase data. This developed method is especially good at the matching of delicate features of FRA traces. The cancellation criteria, when extracting the complex zeros and poles from different frequency regions, can be adjusted to control the estimation accuracy, or to avoid the over fit of noise on the FRA traces in different situations. However, one problem with this method is that the estimated complex zeros and poles may not appear in sequence and

they may have positive real part, which indicates that they are not physically achievable, i.e. the transfer function could not be given an electrical circuit through e.g., filter design software. The choice of order for each frequency region, and the cancellation criteria both have an influence on the Finalised Transfer Function, which mean the solution may not be unique.

Considering the shortcoming of the first method, the second method, Extreme Points Identification Algorithm, is developed. Different from the first method which requires both magnitude and phase frequency response data, the second method needs the magnitude response only for the estimation of transfer function. A real pole is used to describe the decreasing magnitude trend before the first trough. The induced height at each peak and trough on the frequency response, by complex poles and zeros, are calculated, and the height along with the location of the peak or trough can be used to compute the corresponding complex parameters. The mutual influence between different parameters can be eliminated by iterations. In addition to the fact that the complex zeros and poles are enforced to appear in sequence, this method also forces the real part of the complex parameters to be negative. A unique solution which is physically achievable thus can be guaranteed. Similar program running time to the first method is required by the second method, on the same computer. The time length may increase when more iterations are needed.

This method has been verified on two typical winding types, i.e., the Multiple Layer winding type and Plain Disc winding type, and those two types of windings own obvious magnitude oscillations and the local minimum and maximum are easy to identify. Compared with other estimation methods, this proposed method uses the smallest amount of parameters to describe the FRA traces, which assures the simplicity of the mathematical expression. No redundant zeros or poles appear and only one real pole is needed, while other methods may have several pairs of redundant zeros and poles along with a few real poles and zeros.

A problem with the second method is that when the complex zeros and poles are located too close, the local minimum or maximum magnitude may be hidden, thus a pair of complex zeros and poles might be neglected. Another problem is also found on the estimation of phase data. Though phase data is not used, the phase data between ±90° can be accurately matched. However, for the frequency response with phase data which oscillates between ±180°, the estimated transfer function produces FRA phase data which

oscillates between ±90° to guarantee the physically achievable transfer function using the estimated parameters. This situation occurs when the original Frequency response is not a driving point transfer function.

Both of the proposed methods have been applied for the diagnosis of a 275/132/11 kV autotransformer. The tap winding of B phase from this transformer is axially collapsed. The parameters produced by both methods can reflect the shifting of the resonant points in the measured frequency response as compared with the reference. However the second method which is the Extreme Points Identification Algorithm, is preferred, because its complex parameters, i.e. the complex zeros and poles, are corresponding to the antiresonance and resonance points of the frequency response. This means it is advantageous to use this method when interpreting the differences between the reference and diagnostic frequency responses.

7.1.2 Conclusions on Winding Construction Type Classification

Both the supervised and unsupervised machine learning methods are employed on frequency response for the identification of the winding types. All the frequency respones come from the 400/275/13 kV autotransformers.

For the supervised machine learning method, Support Vector Machine (SVM), a particular classification model is studied, where a multiple type classification model, using 54 frequency responses with known winding types. When testing this SVM model using 54 frequency responses with known winding types, the classification accuracy reaches 100%. Then the model is applied for winding type identification after testing. For another group of 51 FRA traces, with unknown winding types, both subjective classification using expert knowledge and the SVM classification are conducted, to yield to same prediction results. Therefore, the classification ability of SVM model is verified.

Sensitivity studies are carried on the key features, or support vectors, which decides the parameters of the built SVM model. It is found that when the support vectors remain unchanged, no matter how many the non-support vectors are used, the SVM model remains the same. If the support vectors are removed from the training data of the SVM model, as a result, the parameters of the model will change. However, it is hard to accurately locate support vectors in the large amount of input features. Therefore, it can be concluded that to build a proper SVM model, adequate input features are needed. This

means a higher probability to contain the proper support vectors, and to construct a well functioned SVM model.

The unsupervised machine learning method Hierarchical Clustering is also used for the identification of winding types of frequency responses. 28 frequency responses are clustered into different groups according to the distances between each another. The clustering result shows that tertiary, common, and series windings can be roughly clustered together. The influences from different frequency regions are investigated. It is possible to identify frequency responses from the tertiary windings, common windings and series winding, using 5 Hz to 2 kHz frequency response, whose magnitudes are related to the numbers of winding turns. Windings with high series capacitance, i.e. Multiple Layer winding type and Interleaved Disc winding type, and windings with low series capacitance, i.e. Intershielded Disc winding type and Plain Disc winding type, can be clustered together, using 20 kHz to 200 kHz frequency response.

It is found that actually the amount of data points has an influence on the clustering result. The quantity of data points can be viewed as a weight factor, related to the measured frequency range, which has influence on the distances between FRA traces. This weigh can be adjusted by reproducing frequency response using transfer function estimation methods. When a proper clustering method is adopted, the identification of unknown frequency response can be achieved by clustering frequency responses with known and unknown winding types together. The winding type of the unknown frequency response is suggested by the known frequency response in the same clusters.

7.2 Future Work

This PhD study focuses on the transfer function estimation and winding type classification using measured frequency response of transformers. Further study is still needed to advance the developed methods. The detailed suggestion future work will be given in the following section.

7.2.1 Future Work on Transfer Function Estimation

As mentioned, the first method developed is able to obtain satisfactory estimation results. Nevertheless, the selection of order of each frequency region, as well as the cancellation criteria for a pair of redundant zeros or poles, has influence on the Finalised Transfer Function. The future work plan for this method is suggested as follows:

The criteria for the selection of order of transfer function to be estimated for a frequency region should be improved. It should allow different orders for the numerator and the denominator according to the amount of peaks of troughs contained in a frequency region. A better estimation result is expected with this improvement, which should significantly reduce the possibility of producing parameters without physical meaning. However, even with this improvement, the solution may not be unique, since that the quantity of real zeros and poles cannot be controlled with the inbuilt MATLAB command.

For the cancellation criteria, study should be conducted to give an adjustment guide for different situations, such as for a different winding construction type, i.e. Single Helical winding type, or for transformers with another voltage ration. Also, the criteria should balance between guaranteeing the estimation accuracy and avoiding the measurement noise. Recommendation should be made after further study. Noticeably, the division of frequency regions might change accordingly.

For the second method developed, the following study is suggested:

The subtle features should be considered by either identifying the change in the rate of magnitude increase/decrease with frequency or by using the more sensitive phase data. Once the closely located complex zeros and poles are estimated, the accuracy of the proposed method can be further improved. This method should be improved and applied on winding types with subtle features, such as Interleaved Disc winding.

7.2.2 Future Work on Winding Construction Type Classification

For the supervised machine learning method SVM, the following further work is suggested:

The method should be applied on frequency responses of transformers with other voltage ratios, such as 275/132kV transformers and 275/33kV transformers. For each voltage ratio, a different SVM model should be built. For example, the Multiple Layer winding from two transformers with different voltage ratio can be viewed as two difference classifications in the SVM model.

The input of the current SVM model is selected as the measured frequency responses after standardisation, which makes the most of the known information. An alternative way is to use the parameters extracted from the frequency response as the input of the model.

Currently, the linear SVM model is sufficient for the investigated winding type classification. The nonlinear SVM model, which pre-processes the frequency responses, may be used for more complex classification challenge such as identification of winding types using the mixed groups of frequency responses from transformers with different voltage ratio as mentioned above. The nonlinear SVM model requires the design of a proper kernel function to pre-process the input data.

For the unsupervised machine learning method, Hierarchical Clustering, the following further work is suggested:

Application on transformers with different voltage ratios should be carried out. For Hierarchical Clustering method, the selection of frequency range of FRA data should be careful, since the frequency range controlled by the winding properties may vary with the voltage ratios and power ratings of transformers.

The transfer function estimation method can be used to generate any arbitrary number of data in any frequency region to investigate a proper weighting factor for different frequency regions, in order to develop a more appropriate clustering method.

For both methods, the following future work is suggested:

The classification of winding types may be combined with the transfer function estimation, by using the parameters of transfer function as the input to both machine learning methods. However, the order of the transfer function should be defined to guarantee the uniform format of the input feature for the SVM model and to enable the calculation of distance between two sets of transfer function parameters.

Reference

[1] M. Heathcote, *J & P transformer book*: Elsevier, 2011.

[2] CIGRÉ Working Group, "An international survey on failures in large power transformers in service," *Electra,* vol. 88, 1983.

[3] M. Wang, A. J. Vandermaar, and K. D. Srivastava, "Review of condition assessment of power transformers in service," *Ieee Electrical Insulation Magazine,* vol. 18, pp. 12-25, Nov-Dec 2002.

[4] O. Grechko and N. Kalacheva, "Current trends in the development of in-service monitoring and diagnostic systems for 110-750 kV power transformers (A survey)," *APPLIED ENERGY-NEW YORK-C/C OF IZVESTIIA-ROSSIISKAIA AKADEMIIA NAUK ENERGETIKA,* vol. 34, pp. 84-97, 1996.

[5] S. Ghosh, *Electrical Machines*: Pearson Education India, 2012.

[6] H. Z. El-hajjar, "Identification of transformer mechanical faults using frequency response analysis," 2008.

[7] S.-S. Group, "SGB-SMIT Group - Transformatorenhersteller aus Europa," *https://Sgb-smit.com,* pp. Available: https://www.sgb-smit.com/. [Accessed: 13- Sep- 2019]. 2019.

[8] Ž. Janic, Z. Valkovic, and Ž. Štih, "Helical winding's magnetic field in power transformers," *Electrical engineering,* vol. 91, p. 161, 2009.

[9] M. Bagheri, B. T. Phung, and M. S. Naderi, "Impulse voltage distribution and frequency response of intershield windings," *IEEE Electrical Insulation Magazine,* vol. 32, pp. 32-40, 2016.

[10] S. G. Fletcher, "Transformer winding," ed: Google Patents, 1948.

[11] S. Morita, K. Ohta, and K. Kurita, "Shielded conductor for disk windings of inductive devices," ed: Google Patents, 1974.

[12] R. A. Hinton, K. W. Doughty, and W. N. Kennedy, "Electrostatic shielding of disc windings," ed: Google Patents, 1977.

[13] G. E. Sauer, "Impulse voltage distribution improving partial-turn electrostatic shields for disc windings," ed: Google Patents, 1977.

[14] J. J. Kelly, "Transformer Fault Diagnosis by Dissolved-Gas Analysis," *IEEE Transactions on Industry Applications,* vol. IA-16, pp. 777-782, 1980.

[15] N. A. Gómez, H. M. Wilhelm, C. C. Santos, and G. B. Stocco, "Dissolved gas analysis (DGA) of natural ester insulating fluids with different chemical compositions," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 21, pp. 1071-1078, 2014.

[16] S. A. Boggs, "Partial discharge: overview and signal generation," *IEEE Electrical Insulation Magazine,* vol. 6, pp. 33-39, 1990.

[17] "IEC 60270, High-voltage test techniques - Partial discharge measurements," ed, 2000.

[18] M. Yaacob, M. A. Alsaedi, J. Rashed, A. Dakhil, and S. Atyah, "Review on partial discharge detection techniques related to high voltage power equipment using different sensors," *Photonic sensors,* vol. 4, pp. 325-337, 2014.

[19] T. Jia, N. Zheng, A. Sun, P. Li, Q. Yu, and L. Luo, "Partial Discharge Localization through a UHF Signal Amplitude Strength Attenuation Approach," in *IOP Conference Series: Materials Science and Engineering*, 2019, p. 012123.

[20] S. Tenbohlen, A. Pfeffer, and S. Coenen, "On-site experiences with multi-terminal IEC PD measurements, UHF PD measurements and acoustic PD localisation," in *2010 IEEE International Symposium on Electrical Insulation*, 2010, pp. 1-5.

[21] J. Lapworth and T. McGrail, "Transformer failure modes and planned replacement," in *IEE Colloquium on Transformer Life Management (Ref. No. 1998/510)*, 1998, pp. 9/1-9/7.

[22] G. K. Supramaniam, Z. F. Hussien, and M. A. T, "Application of Frequency Domain Spectroscopy (FDS) in assessing dryness and ageing state of transformer insulation systems," in *2008 IEEE 2nd International Power and Energy Conference*, 2008, pp. 55-61.

[23] J. Alff, V. D. Houhanessian, W. S. Zaengl, and A. J. Kachler, "A novel, compact instrument for the measurement and evaluation of relaxation currents conceived for on-site diagnosis of electric power apparatus," in *Conference Record of the 2000 IEEE International Symposium on Electrical Insulation (Cat. No.00CH37075)*, 2000, pp. 161-167.
[24] T. K. Saha and P. Purkait, "Investigation of polarization and depolarization current measurements for the assessment of oil-paper insulation of aged transformers," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 11, pp. 144-154, 2004.
[25] G. Bertagnolli, "Short-circuit duty of power transformers," 2013.
[26] S. D. Mitchell and J. S. Welsh, "Methodology to locate and quantify radial winding deformation in power transformers," *High Voltage,* vol. 2, pp. 17-24, 2017.
[27] G. Bertagnolli, *The ABB Approach to Short-circuit duty of power Transformers*, 2007.
[28] M. Florkowski and J. Furgal, "Detection of transformer winding deformations based on the transfer function - measurements and simulations," *Measurement Science and Technology,* vol. 14, pp. 1986-1992, Nov 2003.
[29] W. Lech and L. Tyminski, "Detecting transformer winding damage-the low voltage impulse method," *Electrical Review,* vol. 179, pp. 768-772, 1966.
[30] P. T. M. Vaessen and E. Hanique, "A New Frequency-Response Analysis Method for Power Transformers," *Ieee Transactions on Power Delivery,* vol. 7, pp. 384-391, Jan 1992.
[31] F. Marek and F. Jakub, "Detection of transformer winding deformations based on the transfer function—measurements and simulations," *Measurement Science and Technology,* vol. 14, p. 1986, 2003.
[32] T. Leibfried and K. Feser, "Monitoring of power transformers using the transfer function method," in *IEEE Power Engineering Society. 1999 Winter Meeting (Cat. No.99CH36233)*, 1999, p. 1046 vol.2.
[33] E. Rahimpour, J. Christian, K. Feser, and H. Mohseni, "Transfer function method to diagnose axial displacement and radial deformation of transformer windings," *IEEE Transactions on Power Delivery,* vol. 18, pp. 493-505, 2003.
[34] E. Rahimpour and D. Gorzin, "A new method for comparing the transfer function of transformers in order to detect the location and amount of winding faults," *Electrical Engineering,* vol. 88, pp. 411-416, 2006.
[35] J. Christian and K. Feser, "Procedures for detecting winding displacements in power transformers by the transfer function method," *IEEE Transactions on Power Delivery,* vol. 19, pp. 214-220, 2004.
[36] J. Christian, K. Feser, U. Sundermann, and T. Leibfried, "Diagnostics of power transformers by using the transfer function method," in *1999 Eleventh International Symposium on High Voltage Engineering*, 1999, pp. 37-40 vol.1.
[37] J. Christian, K. Feser, U. Sundermann, and M. Loppacher, "Practical use and limits of the transfer function method for diagnoses of power transformers," in *12th Int. Symp. on High Voltage Engineering, Bangalore, Paper 6-1*, 2001.
[38] E. P. Dick and C. C. Erven, "Transformer Diagnostic Testing by Frequuency Response Analysis," *IEEE Transactions on Power Apparatus and Systems,* vol. PAS-97, pp. 2144-2153, 1978.
[39] A. J. Ghanizadeh and G. B. Gharehpetian, "ANN and cross-correlation based features for discrimination between electrical and mechanical defects and their localization in transformer winding," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 21, pp. 2374-2382, 2014.
[40] K. Pourhossein, G. Gharehpetian, E. Rahimpour, and B. Araabi, "A probabilistic feature to determine type and extent of winding mechanical defects in power transformers," *Electric Power Systems Research,* vol. 82, pp. 1-10, 2012.
[41] M. H. Samimi, S. Tenbohlen, A. A. S. Akmal, and H. Mohseni, "Improving the numerical indices proposed for the FRA interpretation by including the phase response," *International Journal of Electrical Power & Energy Systems,* vol. 83, pp. 585-593, 2016.

[42] P. M. Nirgude, D. Ashokraju, A. D. Rajkumar, and B. P. Singh, "Application of numerical evaluation techniques for interpreting frequency response measurements in power transformers," *IET Science, Measurement & Technology,* vol. 2, pp. 275-285, 2008.
[43] M. Wang, A. Vandermaar, and K. Srivastara, "Evaluation of frequency response analysis data," in *14th international symposium on high voltage engineering*, 2001, pp. 904-907.
[44] E. Rahimpour, M. Jabbari, and S. Tenbohlen, "Mathematical Comparison Methods to Assess Transfer Functions of Transformers to Detect Different Types of Mechanical Faults," *IEEE Transactions on Power Delivery,* vol. 25, pp. 2544-2555, 2010.
[45] K. Jong-Wook, P. ByungKoo, J. Seung Cheol, K. Sang Woo, and P. PooGyeon, "Fault diagnosis of a power transformer using an improved frequency-response analysis," *IEEE Transactions on Power Delivery,* vol. 20, pp. 169-178, 2005.
[46] M. Heindl, S. Tenbohlen, A. Kraetge, M. Krüger, and J. Velásquez, "Algorithmic determination of pole-zero representations of power transformers' transfer functions for interpretation of FRA data," in *Paper D-26, 16th int. symp. on high voltage engineering, Johannesburg*, 2009.
[47] R. Wimmer, S. Tenbohlen, M. Heindl, A. Kraetge, M. Krüger, and J. Christian, "Development of algorithms to assess the FRA," in *Proceedings of 15th Int. Symp. High Voltage Engineering, Ljubljana, Slovenia*, 2007.
[48] J. Bak-Jensen, B. Bak-Jensen, and S. D. Mikkelsen, "Detection of faults and ageing phenomena in transformers by transfer functions," *IEEE Transactions on Power Delivery,* vol. 10, pp. 308-314, 1995.
[49] "IEEE Guide for the Application and Interpretation of Frequency Response Analysis for Oil-Immersed Transformers," *IEEE Std C57.149-2012,* pp. 1-72, 2013.
[50] "IEC 60076 Power transformers - Part 18: Measurement of frequency response," ed, 2012
[51] "CIGRE A2.26 Technical Brochure: Mechanical-Condition Assessment of Transformer Windings Using Frequency Response Analysis (FRA) " Apr. 2008.
[52] J. A. S. B. Jayasinghe, Z. d. Wang, P. N. Jarman, and A. W. Darwin, "Winding movement in power transformers: a comparison of FRA measurement connection methods," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 13, pp. 1342-1349, 2006.
[53] https://www.doble.com/product/m5400/.
[54] "Sweep Frequency Response Analyzer for Power Transformer Winding Diaganosis User Manual, OMICRON," ed.
[55] https://www.omicronenergy.com/en/products/franeo-800/.
[56] "Frequency Response Analysis on Winding Deformation of Power Transformers," *DL/T 911,* 2004.
[57] S. A. Ryder, "Diagnosing transformer faults using frequency response analysis," *IEEE Electrical Insulation Magazine,* vol. 19, pp. 16-22, 2003.
[58] J. Seung Cheol, K. Jong-Wook, P. PooGyeon, and K. Sang Woo, "A pattern-based fault classification algorithm for distribution transformers," *IEEE Transactions on Power Delivery,* vol. 20, pp. 2483-2492, 2005.
[59] S. Birlasekaran and F. Fetherston, "Off/on-line FRA condition monitoring technique for power transformer," *IEEE Power Engineering Review,* vol. 19, pp. 54-56, 1999.
[60] M. H. Samimi and S. Tenbohlen, "The Numerical Indices Proposed for the Interpretation of the FRA Results: a Review," in *VDE High Voltage Technology 2016; ETG-Symposium*, 2016, pp. 1-7.
[61] M. H. Samimi, S. Tenbohlen, A. A. S. Akmal, and H. Mohseni, "Evaluation of numerical indices for the assessment of transformer frequency response," *IET Generation, Transmission & Distribution,* vol. 11, pp. 218-227, 2017.
[62] H. Z. El-Hajjar, *Identification of Transformer Mechanical Faults Using Frequency Response Analysis*: The University of Manchester (United Kingdom), 2008.
[63] J. C. G. Arispe and E. E. Mombello, "Detection of Failures Within Transformers by FRA Using Multiresolution Decomposition," *IEEE Transactions on Power Delivery,* vol. 29, pp. 1127-1137, 2014.

[64] M. Bigdeli, M. Vakilian, and E. Rahimpour, "Transformer winding faults classification based on transfer function analysis by support vector machine," *IET Electric Power Applications,* vol. 6, pp. 268-276, 2012.
[65] D. K. Xu, C. Z. Fu, and Y. M. Li, "Application of artificial neural network to the detection of the transformer winding deformation," in *1999 Eleventh International Symposium on High Voltage Engineering*, 1999, pp. 220-223 vol.5.
[66] M. H. Samimi, S. Tenbohlen, A. A. S. Akmal, and H. Mohseni, "Effect of Different Connection Schemes, Terminating Resistors and Measurement Impedances on the Sensitivity of the FRA Method," *IEEE Transactions on Power Delivery,* vol. 32, pp. 1713-1720, 2017.
[67] E. C. Levy, "Complex-curve fitting," *IRE Transactions on Automatic Control,* vol. AC-4, pp. 37-43, 1959.
[68] J. Bak-Jenson, B. Bak-Jenson, S. D. Mikkelsen, and C. G. Jensen, "Parametric identification in potential transformer modelling," *IEEE Transactions on Power Delivery,* vol. 7, pp. 70-76, 1992.
[69] S. M. Islam, K. M. Coates, and G. Ledwich, "Identification of high frequency transformer equivalent circuit using Matlab from frequency domain data," in *IAS '97. Conference Record of the 1997 IEEE Industry Applications Conference Thirty-Second IAS Annual Meeting*, 1997, pp. 357-364 vol.1.
[70] M. Stace and S. M. Islam, "Condition monitoring of power transformers in the Australian State of New South Wales using transfer function measurements," in *Proceedings of 5th International Conference on Properties and Applications of Dielectric Materials*, 1997, pp. 248-251 vol.1.
[71] S. Birlasekaran, X. Yu, F. Fetherstone, R. Abell, and R. Middleton, "Diagnosis and identification of transformer faults from frequency response data," in *2000 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.00CH37077)*, 2000, pp. 2251-2256 vol.3.
[72] T. McKelvey, H. Akcay, and L. Ljung, "Subspace-based multivariable system identification from frequency response data," *IEEE Transactions on Automatic Control,* vol. 41, pp. 960-979, 1996.
[73] H. Akcay, S. M. Islam, and B. Ninness, "Subspace-based identification of power transformer models from frequency response data," *IEEE Transactions on Instrumentation and Measurement,* vol. 48, pp. 700-704, 1999.
[74] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Transactions on Power Delivery,* vol. 14, pp. 1052-1061, 1999.
[75] B. Gustavsen, "Improving the pole relocating properties of vector fitting," *IEEE Transactions on Power Delivery,* vol. 21, pp. 1587-1592, 2006.
[76] L. D. Tommasi, B. Gustavsen, and T. Dhaene, "Robust transfer function identification via an enhanced magnitude vector fitting algorithm," *IET Control Theory & Applications,* vol. 4, pp. 1169-1178, 2010.
[77] P. Karimifard, G. Gharehpetian, and S. Tenbohlen, "Determination of axial displacement extent based on transformer winding transfer function estimation using vector-fitting method," *European Transactions on Electrical Power,* vol. 18, pp. 423-436, 2008.
[78] M. Bigdeli, M. Vakilian, and E. Rahimpour, "A new method for detection and evaluation of winding mechanical faults in transformer through transfer function measurements," *Advances in Electrical and Computer Engineering,* vol. 11, pp. 23-30, 2011.
[79] K. Cornick, "Distribution of very fast transient overvoltages in transformer windings," *CIGRE report,* 1992.
[80] M. Popov, L. v. d. Sluis, R. P. P. Smeets, and J. L. Roldan, "Analysis of Very Fast Transients in Layer-Type Transformer Windings," *IEEE Transactions on Power Delivery,* vol. 22, pp. 238-247, 2007.
[81] M. Popov, L. V. D. Sluis, R. P. P. Smeets, J. Lopez-Roldan, and V. V. Terzija, "Modelling, simulation and measurement of fast transients in transformer windings with

consideration of frequency-dependent losses," *IET Electric Power Applications,* vol. 1, pp. 29-35, 2007.
[82] T. Ji, W. Tang, and Q. Wu, "Detection of power transformer winding deformation and variation of measurement connections using a hybrid winding model," *Electric Power Systems Research,* vol. 87, pp. 39-46, 2012.
[83] K. G. N. B. Abeywickrama, Y. V. Serdyuk, and S. M. Gubanski, "Exploring possibilities for characterization of power transformer insulation by frequency response analysis (FRA)," *IEEE Transactions on Power Delivery,* vol. 21, pp. 1375-1382, 2006.
[84] N. Abeywickrama, Y. V. Serdyuk, and S. M. Gubanski, "High-Frequency Modeling of Power Transformers for Use in Frequency Response Analysis (FRA)," *IEEE Transactions on Power Delivery,* vol. 23, pp. 2042-2049, 2008.
[85] A. Shintemirov, W. H. Tang, and Q. H. Wu, "A Hybrid Winding Model of Disc-Type Power Transformers for Frequency Response Analysis," *IEEE Transactions on Power Delivery,* vol. 24, pp. 730-739, 2009.
[86] F. d. Leon and A. Semlyen, "Complete transformer model for electromagnetic transients," *IEEE Transactions on Power Delivery,* vol. 9, pp. 231-239, 1994.
[87] D. Filipović-Grčić, B. Filipović-Grčić, and I. Uglešić, "High-Frequency Model of the Power Transformer Based on Frequency-Response Measurements," *IEEE Transactions on Power Delivery,* vol. 30, pp. 34-42, 2015.
[88] J. Li, "High Frequency Power Transformer Modelling for Frequency Response Analysis (FRA) Diagnosis," *PhD Thesis, The University of Manchester, Uk,* 2008.
[89] S. M. Islam and G. Ledwich, "Locating transformer faults through sensitivity analysis of high frequency modeling using transfer function approach," in *Conference Record of the 1996 IEEE International Symposium on Electrical Insulation*, 1996, pp. 38-41 vol.1.
[90] D. M. Sofian, Z.D. Wang, and J. Li, "Interpretation of Transformer FRA Responses— Part II: Influence of Transformer Structure," *IEEE Transactions on Power Delivery,* vol. 25, pp. 2582-2589, 2010.
[91] G. Stein, "A study of the initial surge distribution in concentric transformer windings," *IEEE Transactions on power Apparatus and systems,* vol. 83, pp. 877-893, 1964.
[92] Z.J.Wang, "The calculation of cross-capacitance between transformer windings in transisent analysis," *Journal of Chinese Society of Electrical Engineering, pp. 60-65,* 1990.
[93] Calantaluof, "Inductance Calculation Handbook," *Soviet Union,* 1986.
[94] N. Hashemnia, A. Abu-Siada, M. A. S. Masoum, and S. M. Islam, "Characterization of transformer FRA signature under various winding faults," in *2012 IEEE International Conference on Condition Monitoring and Diagnosis*, 2012, pp. 446-449.
[95] A. Abu-Siada, N. Hashemnia, S. Islam, and M. A. S. Masoum, "Understanding power transformer frequency response analysis signatures," *IEEE Electrical Insulation Magazine,* vol. 29, pp. 48-56, 2013.
[96] J. Pleite, E. Olias, A. Barrado, A. Lazaro, and J. Vazquez, "Transformer modeling for FRA techniques," in *IEEE/PES Transmission and Distribution Conference and Exhibition*, 2002, pp. 317-321 vol.1.
[97] J. Pleite, E. Olias, A. Barrado, A. Lazaro, and R. Vazquez, "Frequency response modeling for device analysis," in *IEEE 2002 28th Annual Conference of the Industrial Electronics Society. IECON 02*, 2002, pp. 1457-1462 vol.2.
[98] S. D. Mitchell and J. S. Welsh, "Modeling Power Transformers to Support the Interpretation of Frequency-Response Analysis," *IEEE Transactions on Power Delivery,* vol. 26, pp. 2705-2717, 2011.
[99] S. D. Mitchell and J. S. Welsh, "Initial Parameter Estimates and Constraints to Support Gray Box Modeling of Power Transformers," *IEEE Transactions on Power Delivery,* vol. 28, pp. 2411-2418, 2013.
[100] R. Aghmasheh, V. Rashtchi, and E. Rahimpour, "Gray Box Modeling of Power Transformer Windings for Transient Studies," *IEEE Transactions on Power Delivery,* vol. 32, pp. 2350-2359, 2017.

[101] Z. W. Zhang, W. H. Tang, T. Y. Ji, and Q. H. Wu, "Finite-Element Modeling for Analysis of Radial Deformations Within Transformer Windings," *IEEE Transactions on Power Delivery,* vol. 29, pp. 2297-2305, 2014.
[102] M. Tahir, S. Tenbohlen, and S. Miyazaki, "Optimization of FRA by an Improved Numerical Winding Model: Disk Space Variation," in *VDE High Voltage Technology 2018; ETG-Symposium*, 2018, pp. 1-6.
[103] N. Hashemnia, A. Abu-Siada, and S. Islam, "Improved power transformer winding fault detection using FRA diagnostics – part 2: radial deformation simulation," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 22, pp. 564-570, 2015.
[104] N. Hashemnia, A. Abu-Siada, and S. Islam, "Improved power transformer winding fault detection using FRA diagnostics – part 1: axial displacement simulation," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 22, pp. 556-563, 2015.
[105] G. M. V. Zambrano, A. C. Ferreira, and L. P. Caloba, "Power transformer equivalent circuit identification by artificial neural network using frequency response analysis," in *2006 IEEE Power Engineering Society General Meeting*, 2006, p. 6 pp.
[106] Y. Luo, J. Ye, J. Gao, G. Chen, G. Wang, L. Liu*, et al.*, "Recognition technology of winding deformation based on principal components of transfer function characteristics and artificial neural network," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 24, pp. 3922-3932, 2017.
[107] J. Liu, Z. Zhao, C. Tang, C. Yao, C. Li, and S. Islam, "Classifying Transformer Winding Deformation Fault Types and Degrees Using FRA Based on Support Vector Machine," *IEEE Access,* vol. 7, pp. 112494-112504, 2019.
[108] A. R. Abbasi, M. R. Mahmoudi, and Z. Avazzadeh, "Diagnosis and clustering of power transformer winding fault types by cross-correlation and clustering analysis of FRA results," *IET Generation, Transmission & Distribution,* vol. 12, pp. 4301-4309, 2018.
[109] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their Applications,* vol. 13, pp. 18-28, 1998.
[110] H. Chih-Wei and L. Chih-Jen, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks,* vol. 13, pp. 415-425, 2002.
[111] M. H. Hassoun, *Fundamentals of artificial neural networks*: MIT press, 1995.
[112] M. Rahmatian, M. Naderi, G. Gharehpetian, and A. Ghanizadeh, "Simultaneous detection of radial deformation location and axial displacement extent in a transformer winding using cross-correlation technique and ann," *Int'l. J. Eng. Innov. Res.,* vol. 1, pp. 478-483, 2012.
[113] K. R. Gandhi and K. P. Badgujar, "Artificial neural network based identification of deviation in frequency response of power transformer windings," in *2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD)*, 2014, pp. 1-8.
[114] V. Rashtchi, E. Rahimpour, and E. M. Rezapour, "Using a genetic algorithm for parameter identification of transformer RLCM model," *Electrical Engineering,* vol. 88, pp. 417-422, 2006.
[115] V. Rashtchi, H. Shayeghi, M. Mahdavi, A. Kimiyaghalam, and E. Rahimpour, "Using an improved PSO algorithm for parameter identification of transformer detailed model," *International Journal of Electrical Power and Energy Systems Engineering,* vol. 1, pp. 138-144, 2008.
[116] V. Rashtchi, E. Rahimpour, and H. Shahrouzi, "Model reduction of transformer detailed R-C-L-M model using the imperialist competitive algorithm," *IET Electric Power Applications,* vol. 6, pp. 233-242, 2012.
[117] P. Mukherjee and L. Satish, "Construction of Equivalent Circuit of a Single and Isolated Transformer Winding From FRA Data Using the ABC Algorithm," *IEEE Transactions on Power Delivery,* vol. 27, pp. 963-970, 2012.

[118] V. Rashtchi, E. Rahimpour, and H. Fotoohabadi, "Parameter identification of transformer detailed model based on chaos optimisation algorithm," *IET Electric Power Applications,* vol. 5, pp. 238-246, 2011.
[119] A. De and N. Chatterjee, "Impulse fault diagnosis in power transformers using self-organising map and learning vector quantisation," *IEE Proceedings - Generation, Transmission and Distribution,* vol. 148, pp. 397-405, 2001.
[120] D. M. Sofian, "Transformer FRA Interpretation for Detection of Winding Movement," *PhD Thesis, The University of Manchester, Uk,* 2007.

Appendices

Appendix A: Application of Dynamic Time Warping (DTW)

Considering the shifting/scaling of frequency response's characteristics caused by sizing including rated voltage and power, a technique called Dynamic Time Warping (DTW) can be helpful to pre-process the FRA data for better identifying the same characteristics at different frequencies. The DTW technique, which is able to scale or shift part of the features on FRA traces, is used in combination with Hierarchical Clustering to enhance the classification ability by identifying the features located at different frequencies.

1. Pre-processing by DTW

In time series analysis, DTW is one of the algorithms for measuring similarity between two temporal sequences, which may vary in speed. It can be used to identify same characteristics before and after scaling. DTW horizontally scales two traces to align them to each other, which means it can be used for the 'feature to feature' comparison after data being processed. Among all the possible paths, it chooses the one which has smallest moving distance. Horizontal points will be inserted to compensate the part where the height is not aligned. Generally the greater the difference between the two traces, the more insertion points there are. Indeed, any data that can be turned into a linear sequence can be analyzed with DTW.

1) Methodology

An example, using two data sets A and B, will be given to explain the DTW technique. Data set A consists of data points 1, 3, 4, 9, 8, 2, 1, 5, 7, 3 and data set B consists of data points 1, 6, 2, 3, 0, 9, 4, 3, 6, 3, as plotted in Figure 1. It can be seen that the peaks and troughs on the two sets of data are not aligned.

A distance matrix is helpful to find the possible moving path to match the features on the two datasets, as shown in Table 1. An element X_{ij} in this distance matrix contains the distances between the corresponding two points, A_i and B_j, from the two datasets, which represents a possible alignment in the possible moving path, and the value of X_{ij} is defined in Equation 1:

$$X_{ij} = |A_i - B_j| + \min(D_{(i-1,j-1)}, D_{(i,j-1)}, D_{(i-1,j)}) \quad \text{Equation 1}$$

where $D_{(m,n)}$ is the absolute difference between A_m and B_n, X_{ij} is the smallest distances among $D_{(i-1,j-1)}$, $D_{(i,j-1)}$, and $D_{(i-1,j)}$. Among all the possible moving paths, the path with the smallest accumulated moved distance between the two traces should be used. The distance matrix should be calculated starting from X_{11}. Though not all possible moving steps are shown in the distance matrix, the one with the smallest accumulated distance is surely included. After the distance matrix is calculated, start from the right bottom corner, the path with the least accumulated distance is selected. For the two datasets A and B, the selected path is noted in grey shade, as shown in Table 2. The mapping path for A is 1, 6, 2, 3, 0, 9, 9, 4, 3, 6, 6, 3, and for B is 1, 3, 3, 3, 4, 9, 8, 2, 1, 5, 7, 3. Figure 2 shows the traces of A and B after mapping.

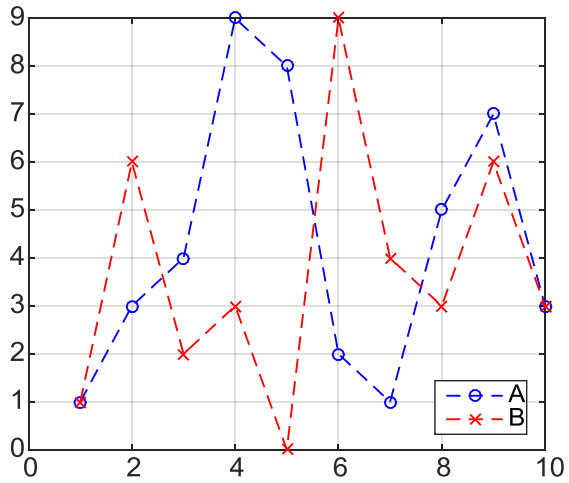


Figure 1 Data Sets A and B to be Processed by DTW Technique

Table 1 Distance Matrix of Data Sets A and B to be Processed

	1	6	2	3	0	9	4	3	6	3
1										
3										
4										
9										
8										
2				X_{ij}						
1										
5										
7										
3										

Table 2 Distance Matrix of Data sets A and B Processed

	1	6	2	3	0	9	4	3	6	3
1	0	5	6	8	9	17	20	22	27	29
3	2	3	4	4	7	13	14	14	17	17
4	5	4	5	5	8	12	12	13	15	16
9	13	7	11	11	14	8	13	18	16	21
8	20	9	13	16	19	9	12	17	18	21
2	21	13	9	10	12	16	11	12	16	17
1	21	18	10	11	11	19	14	13	17	18
5	25	19	13	12	16	15	15	16	14	16
7	31	20	18	16	19	17	18	19	15	18
3	33	23	19	16	19	23	18	18	18	15

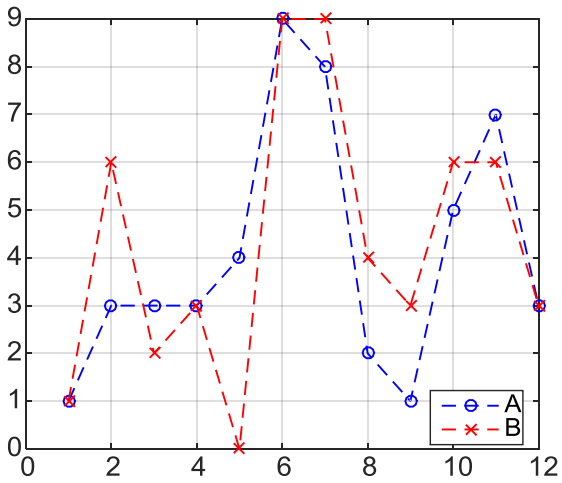


Figure 2 Data Sets A and B Processed by DTW Technique

2) Application of DTW on Two Datasets

In the first part of this section, sensitivity study on the DTW method is carried out using artificial frequency responses. Two influencing factors are investigated, i.e. the location of the features on the frequency response and the overall magnitude of the frequency response. In the second part of this section, the application of the DTW on the measured frequency responses from the same or different winding types is demonstrated.

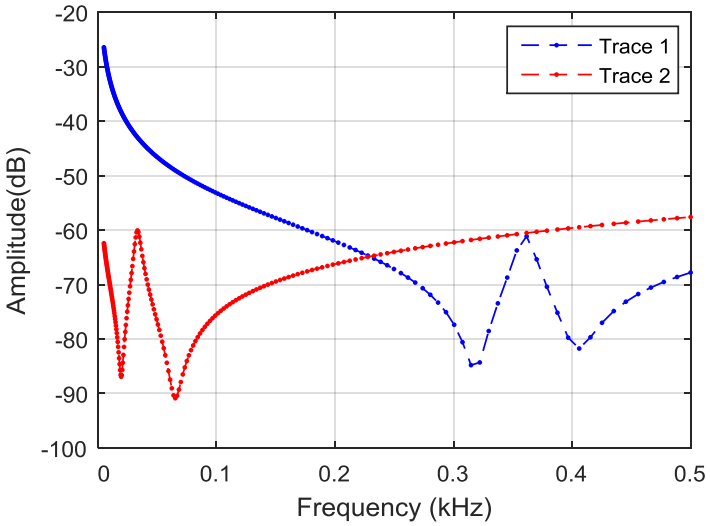
a. Sensitivity Study on Artificial FRA Trace

Two artificial FRA traces are constructed for the sensitivity study of DTW, as shown in Figure 3. For each trace, 200 data points are sampled linearly from 1 Hz to 0.5 kHz. Both trace 1 and trace 2 own two troughs and 1 peak, located at different frequencies. The

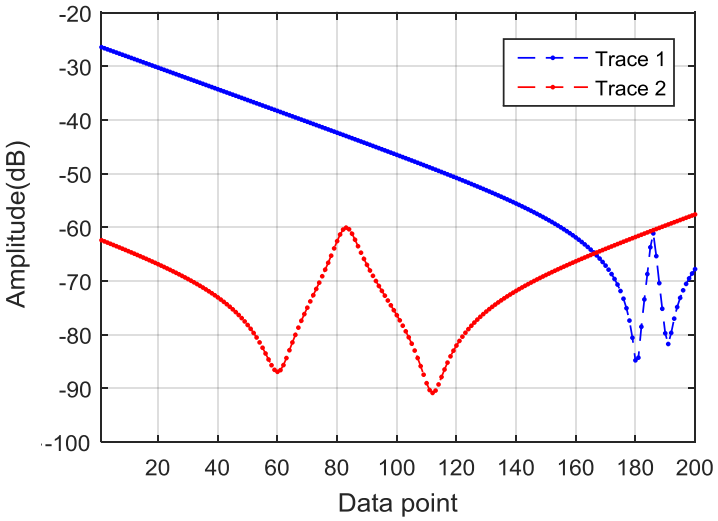
horizontal axis in Figure 3(b) is data point, instead of frequency, for the convenience of comparison after DTW is applied. After DTW technique is applied, the two traces are plotted in Figure 3(c). It can been seen that even if the original peaks and troughs on two traces are located far away, they can be matched together after the DTW is applied.

However, if the parameters of Trace 2 are changed so that the peak and troughs move to lower frequency range, i.e. the first trough from 0.0392 kHz to 0.024 kHz, the second tough from 0.0667 kHz to 0.0493 kHz, the first peak from 0.131 kHz to 0.0349 kHz, as shown in Figure 4(a)-(b), the matching result changes, as shown in Figure 4(c). It can be seen that the second trough on Trace 2 is matched with the first trough of Trace 1. Thus it can be concluded that the locations of the features on the original traces to be processed have an influence on the matching result. This is understandable since that the DTW seeks the smallest accumulated path. Therefore, it should be noticed that when applying DTW technique, it is important to select the relevant frequency range to be processed. In order to verify this conclusion, a further study is carried out. The parameters of Trace 1 and Trace 2 in Figure 4 remain the same, and the frequency range is extended from 0.1 Hz to 5 kHz. Because of the extension of frequency range, the features are comparatively closer located in terms of data points, and thus a more desired result is obtained, as shown in Figure 5.

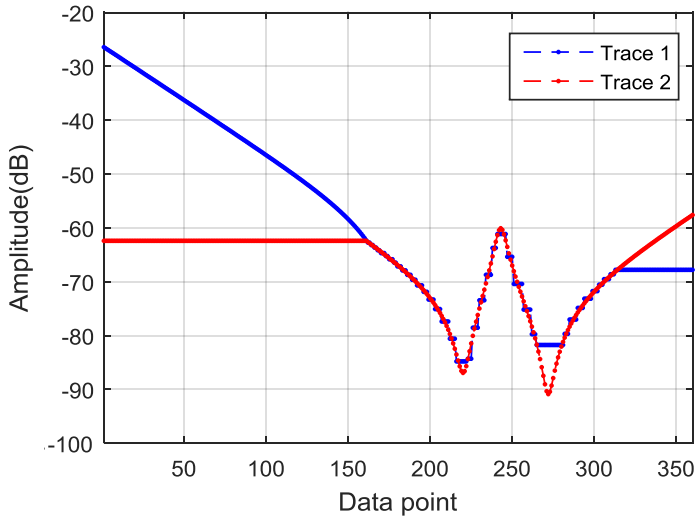
The influence of overall magnitude is investigated in Figure 6. The parameters of Trace 1 and Trace 2 in Figure 3 remain unchanged, except that Trace 2 is shifted upwards for 30dB. It can be seen that the peaks and troughs from the two traces are not aligned. Instead, part of the slope of Trace 1 is aligned with the peak and troughs of Trace 2. This is because that magnitude is also regarded as a feature of the original data. The measured frequency responses from different winding types do have different overall magnitudes. For example, the Single Helical winding type has the highest magnitude, roughly from 0 Hz to 100 kHz. Therefore, the magnitude of the FRA trace is also expected to play a role in the clustering process in combination of DTW technique..



(a) Horizontal Axis using Frequency

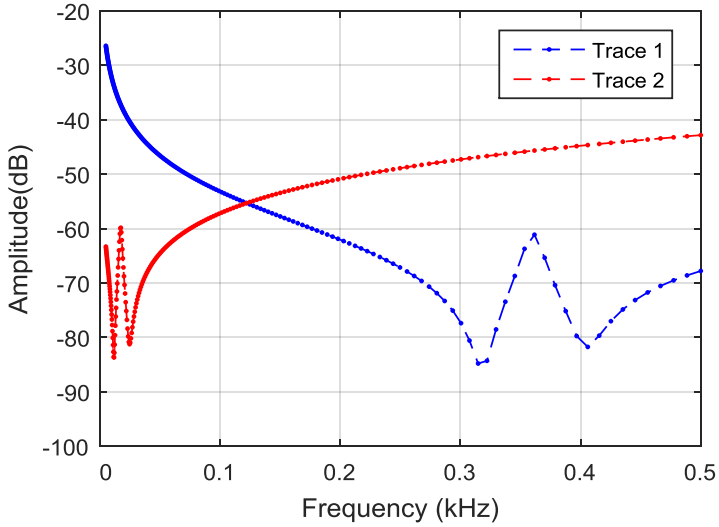


(b) Horizontal Axis using Data Point

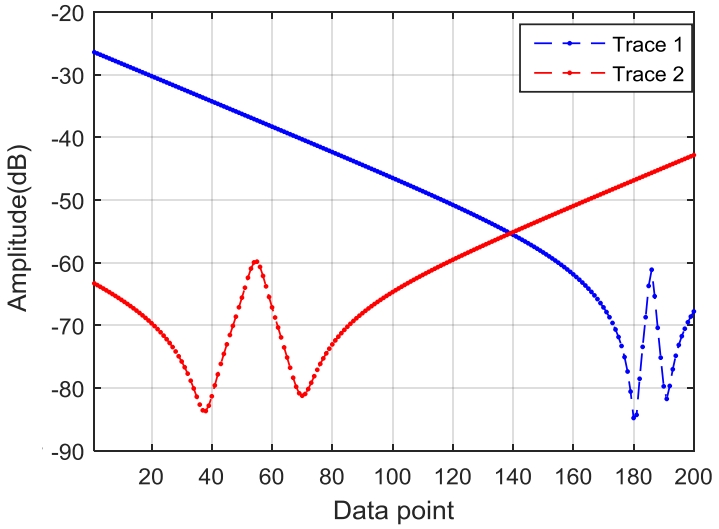


(c) Artificial traces processed by DTW

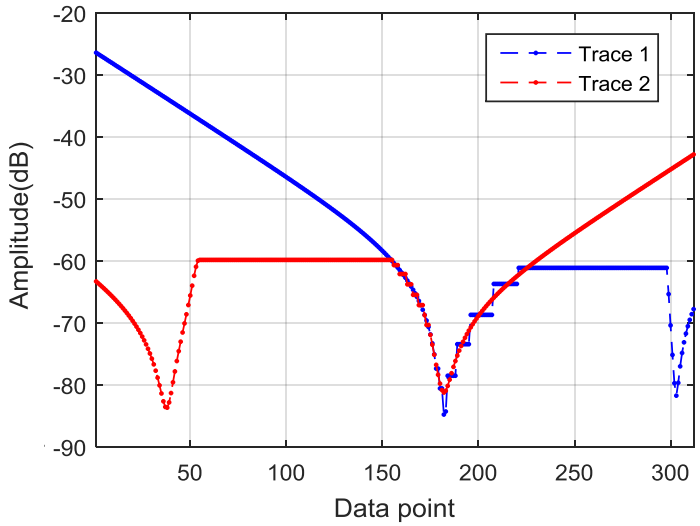
Figure 3 Artificial FRA Traces for Sensitivity Study on DTW Method



(a) Horizontal Axis using Frequency

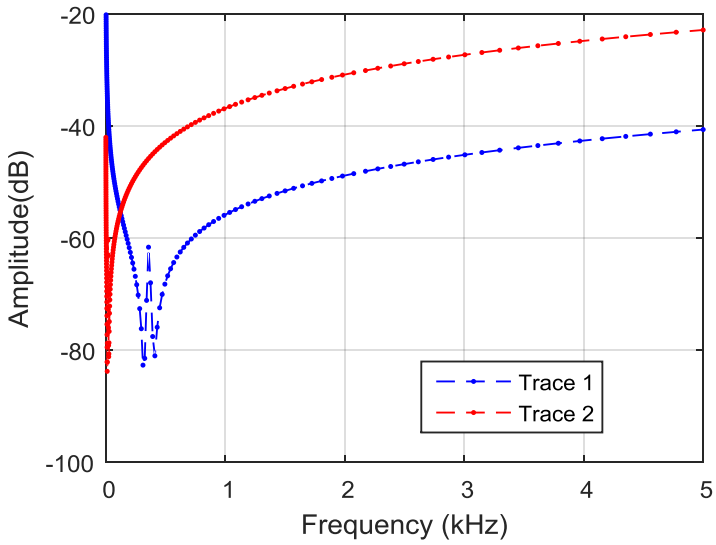


(b) Horizontal Axis using Data Point

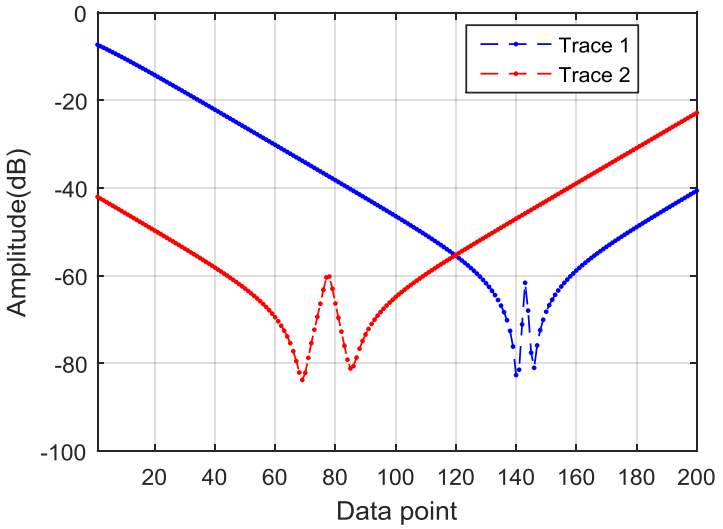


(c) Artificial traces processed by DTW

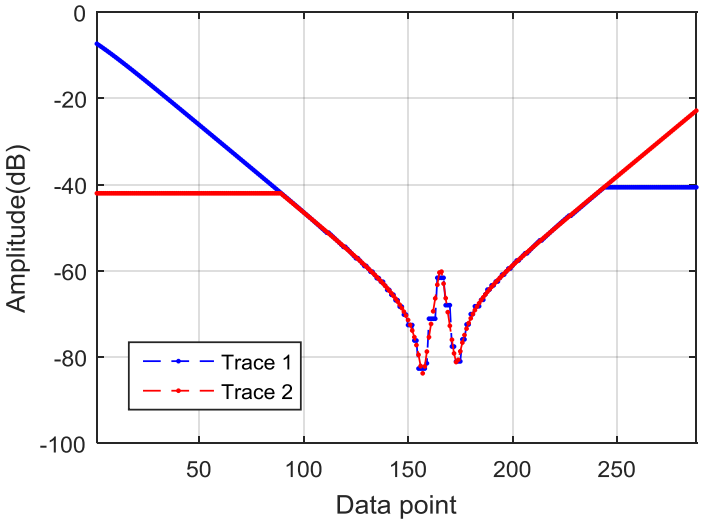
Figure 4 Influence of Feature Location on DTW Technique 1



(a) Horizontal Axis using Frequency

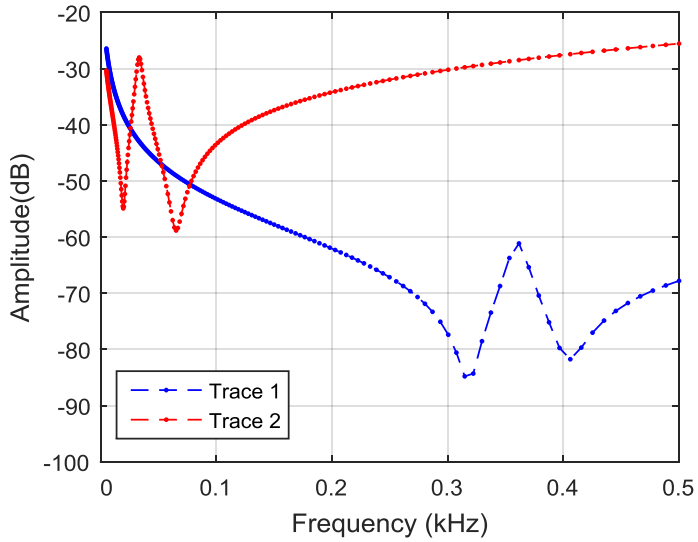


(b) Horizontal Axis using Data Point

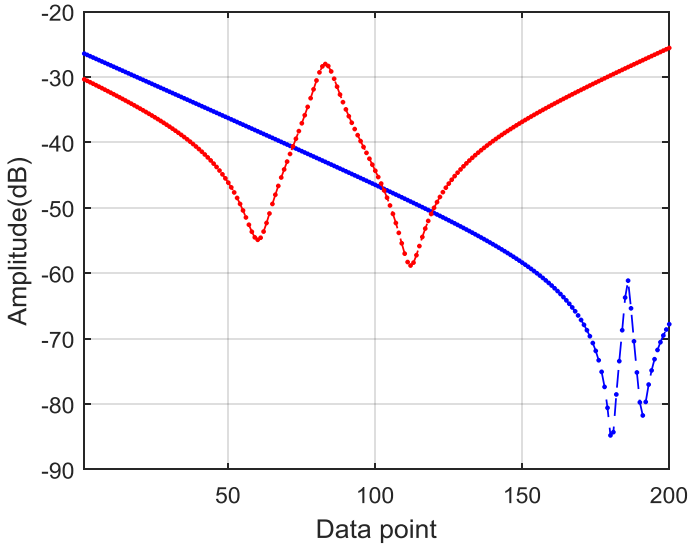


(c) Artificial Traces Processed by DTW

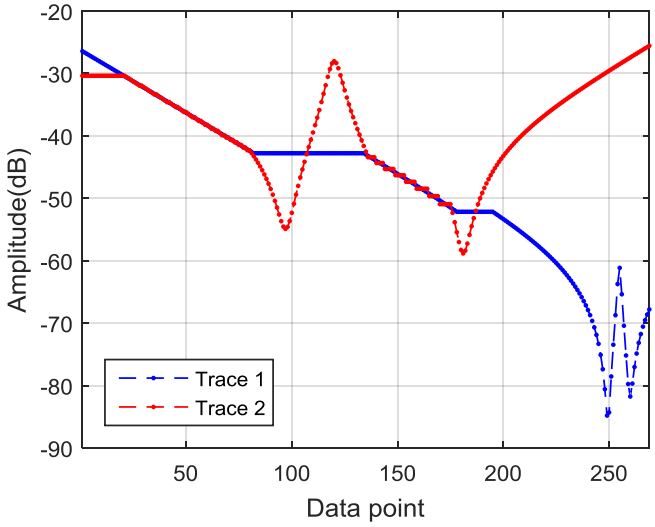
Figure 5 Influence of Feature Location on DTW Technique 2



(a) Horizontal Axis using Frequency



(b) Horizontal Axis using Data Point



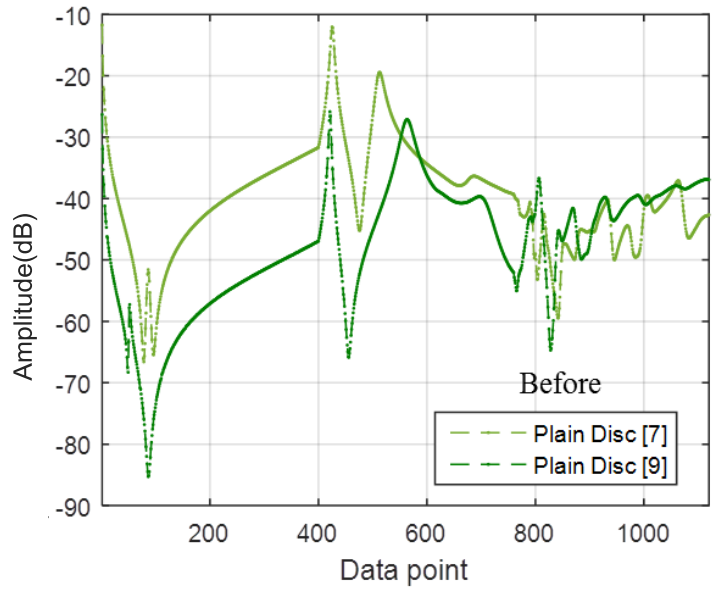
(c) Artificial Traces Processed by DTW

Figure 6 Influence of Magnitude on DTW Technique

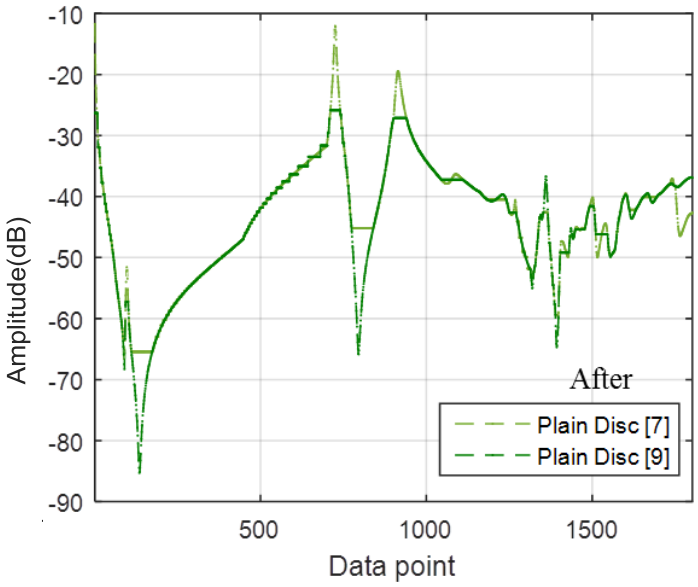
b. Application on Measured FRA Trace

In Figure 7, comparison of two frequency responses from Plain Disc winding, from 5 Hz to 200 kHz, before and after DTW applied is illustrated. The horizontal axis is data point for both Figure 7(a) and Figure 7(b). For the frequency region controlled by the core properties below 2 kHz, it can be seen that after application of DTW, the two troughs and one peak from both traces are shifted and scaled to match each other. For the frequency region 2 kHz to 20 kHz controlled by winding properties, the flatting area on the trace 7 can be clearly seen, which is used to match the trough and two peaks of trace 9. For the most concerned frequency region 20 kHz to 200 kHz, the peaks and troughs are aligned to each other with several flatting areas on both traces, which may be hard to find visually due to their intensity. The quantity of data points increases from 1120 to 1802, and the distance between the two traces changes from 406 to 137 after DTW is applied. Therefore, DTW is able to enhance the similarity between two traces.

Another example is shown in Figure 8 when DTW is applied on two frequency responses from two difference winding types, i.e. Plain Disc winding and Interleaved Disc Winding. For the two traces, the peaks and troughs from 5 Hz to 20 kHz are matched well with small amount of flatting data points, since their magnitude difference is small. However, the difference in the winding controlled frequency region, 20 kHz to 200 kHz, remains huge after DTW processing. The quantity of data points for both race rises from 1120 to 1565. The distance between the two traces is reduced from 383 to 233. Though the similarity between them is enhanced, there still exists a significant difference between the distance of frequency responses from same winding type (137) and the distance of frequency responses from difference winding type (233), after application of DTW.

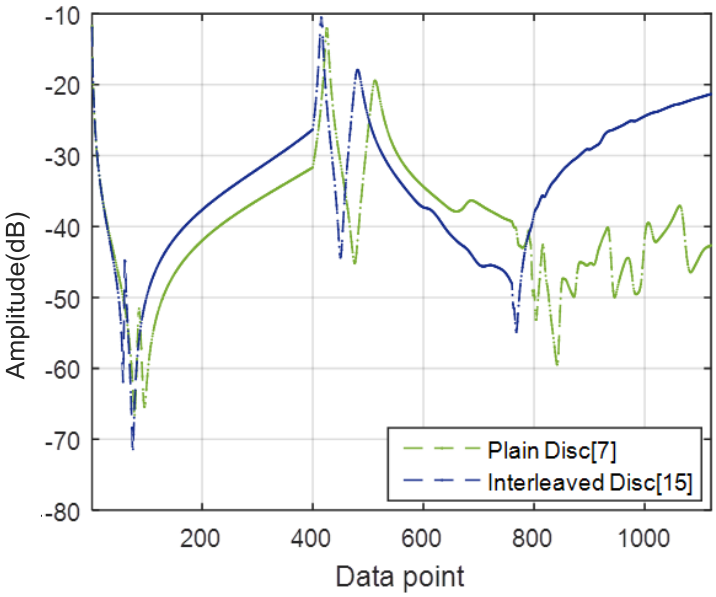


(a)

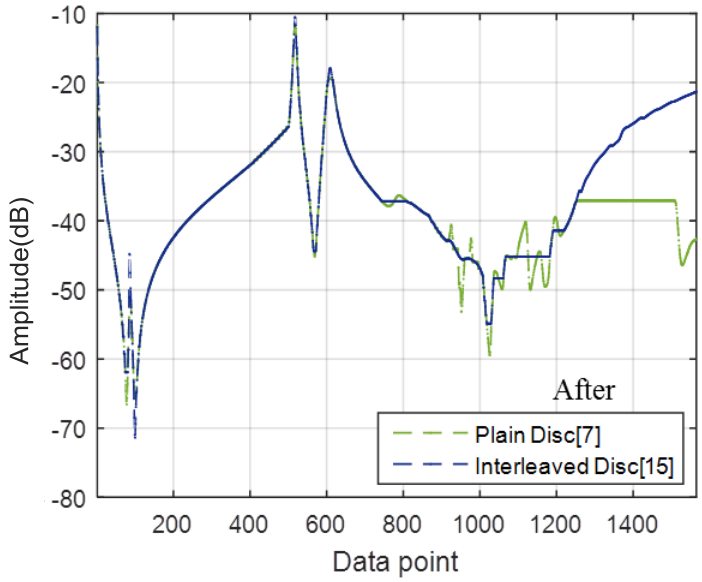


(b)

Figure 7 DTW Applied on Two Windings of Same Type (Plain Disc Windings)



(a)



(b)

Figure 8 DTW Applied on Two Different Windings (Plain Disc and Interleaved Disc Windings)

2. Clustering DTW processed Data

The Euclidean Distance should be calculated after the traces being processed by DTW. In order to make the Euclidean Distance comparable, all the Euclidean Distance are divided by the number of data points of frequency response after processed by DTW, and then multiplied by the original data points 1120.

When DTW is applied on every two of the 28 frequency responses in Chapter 6 to calculate their Euclidean Distance, a better clustering result than the former attempts can be produced as shown in Figure9.

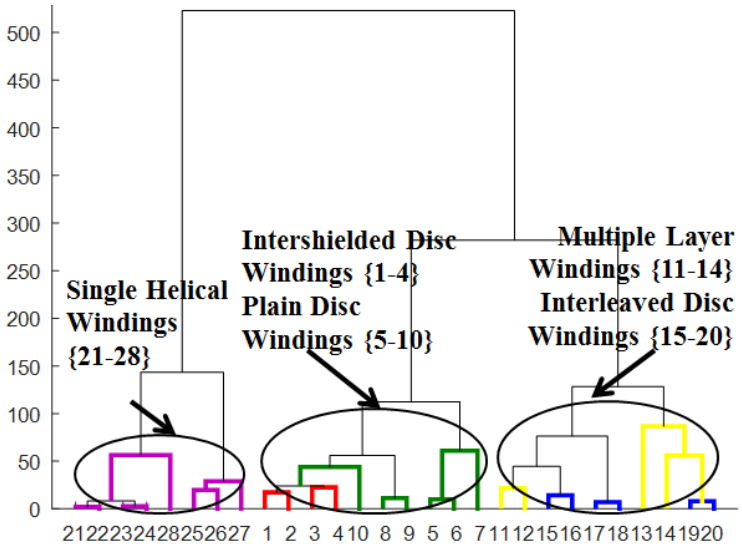


Figure9 5 Hz to 200 kHz Clustering Dendrogram using Euclidean Distance Weighted Linkage

The overall distance is reduced as compared with Figure 6 2, with the final distance between the last two clusters decreasing from about 900 to about 550. It can be seen that all traces are divided into three distinctive groups; one group contains all the 8 Single Helical windings and the others are separated into two groups, one with low capacitance (Plain Disc winding and Intershielded Disc winding) and the other with high capacitance (Multiple Layer winding and Interleaved Disc winding).

3. Summary

With DTW technique applied to pre-process the FRA data, the same FRA features at different frequencies can be scaled or shifted. This pre-processing technique significantly increases the chance to correctly group the same winding types. Three clustering groups can be identified as Single Helical winding, winding with low series capacitance (Plain Disc winding and Intershielded Disc winding) and windings with high series capacitance (Multiple Layer winding and Interleaved Disc winding).

Appendix B: List of Publications

1) **X. Mao**, Z.D. Wang, Z.J. Wang, P. Jarman: 'Accurate Estimating Algorithm of Transfer Function for Transformer FRA Diagnosis', *IEEE Power & Energy Society General Meeting*, Portland, OR, 2018, pp. 1-5.

2) **X. Mao**, Z.D. Wang, P. Jarman and A. Fieldsend-Roxborough: 'Winding Type Recognition through Supervised Machine Learning using Frequency Response Analysis (FRA) Data', *2019 2nd International Conference on Electrical Materials and Power Equipment,* Guangzhou, China, 2019, pp. 588-591.

3) **X. Mao**, S. Ji, Z.D. Wang, P. Jarman, A. Fieldsend-Roxborough, G.Wilson: 'Applying Unsupervised Machine Learning Method on FRA Data to Classify Winding Types', *International Symposium on High Voltage Engineering,* Budapest, Hungary, 2019, pp. 969-981

4) **X. Mao**, Z.D. Wang, P. Crossley, P. Jarman, A. Fieldsend-Roxborough, G.Wilson:'Winding Type Recognition through Applying Support Vector Machine on Transformer Frequency Response Analysis Data', *IET High Voltage* (accepted)

5) **X. Mao**, Z.D. Wang, P. Crossley, P. Jarman, A. Fieldsend-Roxborough, G.Wilson:'Extreme Points Identification Algorithm of Transfer Function Estimation for Transformer FRA Diagnosis', *IEEE Power Delivery* (in preparation)