ESCUELA POLITÉCNICA NACIONAL

DOCTORADO EN INGENIERÍA ELÉCTRICA
MENCIÓN EN SISTEMAS ELÉCTRICOS DE POTENCIA

DETERMINATION OF THE RESIDUAL LIFE OF POWER TRANSFORMER WINDINGS THROUGH THE ANALYSIS OF CUMULATIVE FATIGUE DUE TO INTERNAL ELECTROMAGNETIC FORCES

FAUSTO RAMIRO VALENCIA ARCOS

fausto.valencia@epn.edu.ec

DIRECTOR: DR. ING. HUGO NEPTALÍ ARCOS MARTÍNEZ

hugo.arcos@epn.edu.ec

Quito, January 2024

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por Fausto Ramiro Valencia Arcos bajo mi supervisión.

Dr. Ing. Hugo Neptalí Arcos Martínez
DIRECTOR DE PROYECTO

DECLARACIÓN

Yo, Fausto Ramiro Valencia Arcos, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

Fausto Ramiro Valencia Arcos

DEDICATORIA

A mi familia y a mis amigos, quienes tuvieron más fe que la que yo tuve en que este proyecto habría de llegar a concretarse.

AGRADECIMIENTOS

Agradezco al Dr. Hugo Arcos, cuya contribución a esta investigación permitió que la misma mejorara día a día hasta llegar a su consecución final.

Un agradecimiento especial al Dr. Franklin Quilumba quien fue un apoyo tanto profesional como personal durante los momentos más difíciles de la investigación.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SOURCE CODES

RESUMEN

Los transformadores de potencia se encuentran entre los elementos más importantes de un sistema de potencia. Dependiendo de la topología del sistema, una falla en un transformador de potencia podría causar la pérdida de energía en muchos usuarios y hasta en ciudades enteras. Por esta razón, un gran esfuerzo se ha realizado para preservarestos equipos in perfectas condiciones y así, evitar eventos de falla.

Conforme el transformador envejece la probabilidad de falla aumenta. Por lo tanto, es necesario monitorear su vida remanente de manera continua de tal manera que quienes administran al sistema estén concientes de un posible problema y tomen las acciones necesarias para mitigar las consecuencias de una falla total del equipo. Por esta razón, muchas investigaciones han sido llevadas a cabo para estimar la vida útil de un transformador de potencia.

La vida útil de los transformadores de potencia es usualmente analizada desde el punto de vista del aisalmiento. Es una práctica generalizada el tomar la vida del papel aislante como la vida del transformador de potencia. Sin embargo, se han visto casos en los cuales la falla del transformador aparece en el conductor y no en el aislamiento. Debido a la gravedad de este hecho, la resistencia mecánica del conductor ante fuerzas electromecánicas ha llegado a ser parte del diseño del transformador.

El análisis de los efectos de las fuerzas en los conductores de los devanados es generalmente realizada en relación con las fuerzas estáticas. Es decir, du-

rante el diseño, se calculan las fuerzas para el peor escenario y, de acuerdo a éstas, se especifica el conductor. Sin embargo, existen efectos continuos de las fuerzas, aún cuando dichas fuerzas son de menor valor que aquellas de diseño. Esos efectos se presentan durante toda la vida del transformador y podrían disminuir la vida útil de los conductores del devanado.

Con estos antecedentes, esta investigación propone un método para analizar las consecuencias de la fatiga mecánica en la vida útil de los transformadore de potencia. A diferencia de investigaciones previas que consideran sólo casos de falla, en esta tesis, se consideran todas las condiciones operativas. El típico método de elementos finitos fue utilizado para calcular la fuerza y la tensión mecáncia en los conductores. Además, considerando posibles aplicaciones en tiempo real, se diseñó un modelo de aprendizaje de máquina, basado en árboles aleatorios. Se obtuvo un resultado interesante, ya que los tiempos para determinar las fuerzas se redujeron considerablemente del tiempo utilizado por el método de elementos finitos.

El procedimiento fue aplicado a un transformador del sistema de prueba de confiabilidad del IEEE. Este sistema contiene datos de carga con variaciones a lo largo del año, del mes, de la semana y del día. Por lo tanto, fue posible simular las condiciones operativas del transformador de una forma horaria con un buen de detalle de la carta. Para estar más cercano a lo que sucede en la realidad, la carga total del sistema fue incrementada de una manera anual. La potencia alimentada por cada generador se la estableció de manera proporcional de acuerdo a su capacidad. Con la finalidad de realizar una comparación entre materiales, se simularon casos tanto para devanados de aluminio como para devanados de cobre.

Se monitoreó la condición del transformador hasta que éste alcanzó el final de

su fida útil, de acuerdo al análisis de fatiga.

En cuanto al desarrollo teórico, los últimos avances relacionados con el análisis de fatiga de transformadores de potencia han sido descritos. Además, se presenta la aplicación del método de elementos finitos para el análisis del campo magnético, así como el respectivo cálculo de las fuerzas electromagnéticas.

Palabras Clave: Análisis de Fatiga, Árboles Aleatorios, Fuerzas Electromagnéticas, Método de Elementos Finitos, Tensión Mecánica, Transformadores de Potencia

ABSTRACT

Power transformers are among the most important elements of a power system. Depending on the topology of the system, a failure in a power transformer may cause the lack of energy for many customers and even for entire cities. For that reason, many efforts have been made to preserve these devices in perfect conditions so that failure events are avoided.

As the transformer ages, the probability of failure increases. It is necessary to monitor its lifespan continuously so that the managers are conscious of a possible problem and take the required actions to mitigate the consequences of a total failure of the equipment. For this reason, many researches have been done in order to estimate the lifespan of a power transformer.

The lifespan of power transformers is usually analysed from the point of view of the insulation. It is common knowledge that the life of the insulation paper is the life of the power transformer. However, there have been found cases where the fault of the transformer appears in the conductor and not in the insulation. This fact has been so serious that the strength of the conductor in regards to electromagnetic forces has become part of the transformer design.

The analysis of the effects of forces in the winding conductors is generally performed in regards to static forces. That means that during the design, the forces are calculated for the worst scenario and the conductor is specified accordingly. However, there are continuous effects of the forces, even when these forces are lower than those of the design. These effects work during the whole life and could diminish the life of the winding conductors.

Given those antecedents, this research proposes a method to analyse the consequences of fatigue in the lifespan of power transformers. Different from previous investigations that consider only fault cases, in this thesis, the whole operating conditions of the transformer are taken into account. The usual finite element method was used to calculate the force and the stress in the conductors. Besides, considering the possibility of real time applications, a machine learning model was designed, based on random forests. This was an interesting result because the time to find the forces was considerably reduced from the direct use of the finite element method.

The procedure was applied to a 400 MVA transformer of the IEEE Reliability Test System. This test system has data of load that varies along the year, month, week, and day. Hence, it was possible to simulate the transformer operating conditions on an hourly basis with a good detail of the load curve. To be closer to reality, the total load of the system was increased each year. The power served by each generator was set proportionally to its full capacity. For comparison purposes, windings made by copper and aluminium were simulated.

The condition of the transformer was monitored until it reached the ultimate point of life, according to the fatigue analysis.

In regards to the theoretical framework, the latest developments related to the fatigue analysis of power transformers have been described. Also, the application of the finite element method for the analysis of magnetic fields is presented, together with the calculation of forces.

Keywords: Fatigue Analysis, Random Forests, Electromagnetic Forces, Finite Element Method, Mechanical Stress, Power Transformers

1 INTRODUCTION

Power transformers are among the most important elements of a power system. A fault in a power transformer could cause even an entire city to lose electrical energy. To make things worse, the repair or replacement actions could last a long time.

For that reason, power transformers must have good maintenance. In the optimal case, utilities must have an asset management program to be very well acquainted with the state of these devices.

1.1 ASSET MANAGEMENT FUNDAMENTALS

One of the best practices to preserve the condition of power transformers is asset management. Some definitions of asset management can be found in standard PAS-55 [1] and CIGRE [2]; they have in common the life cycle of the asset and the reduction of risks and costs. The objective of maintenance of power transformers is threefold: reducing the risks of a failure, reducing the costs as a consequence of a lack of the equipment or an non-optimal point of operation, and finally, procuring to extend the life cycle of the transformer.

Asset management has three types of activities: maintenance plan, condition assessment, and life evaluation [3]; see Figure 1.1.

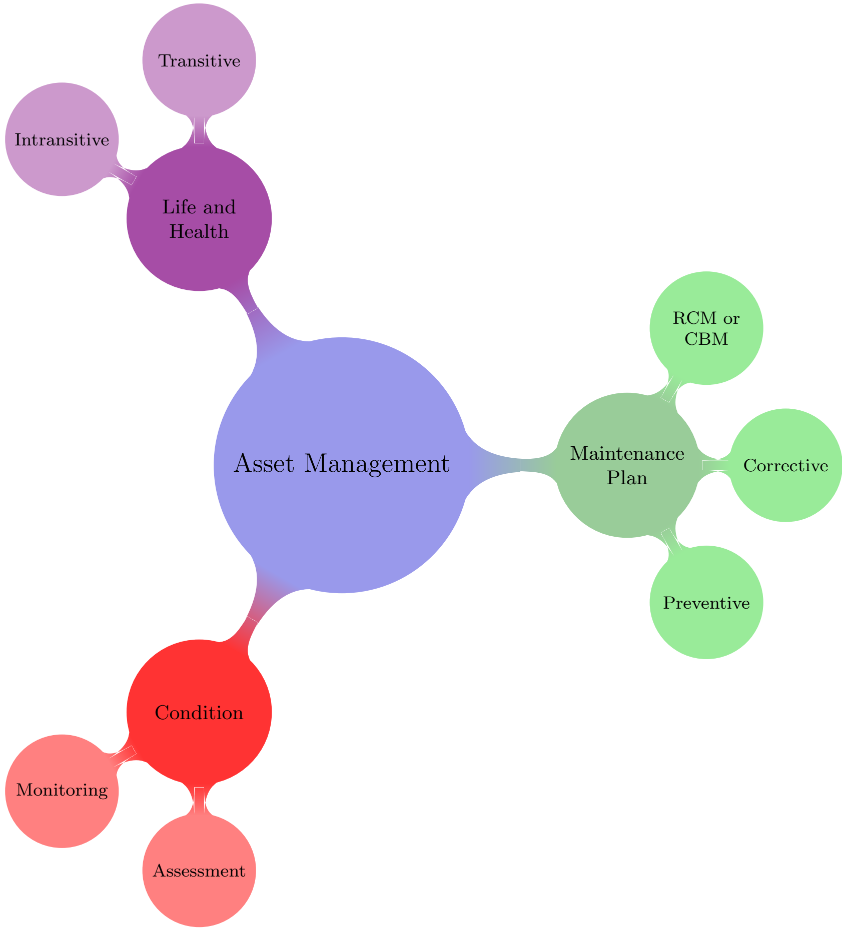


Figure 1.1: Asset Management Areas: (1) Life and Health, (2) Condition Evaluation, (3) Maintenance

Table 1.1: Condition monitoring related to Condition Assessment

Condition Monitoring	Condition Assessment
Hot spot temperature	Thermal analysis
Wall and winding vibration	Vibration analysis
Dissolved gas in oil	Dissolved gas analysis
Partial discharges	Partial discharges analysis
Winding movements and deformations	Frequency response analysis

1.1.1 Condition Monitoring and Assessment

Condition monitoring allows the utilities to know the health state of the transformer. It works along with condition assessment, as it is seen in Table 1.1 [3].

Nowadays, there are also some variables that are monitored online, such as hot spot temperature and dissolved gas analysis in oil [4]. Moreover, all this monitoring could be part of a multiagent system, which functions as a collector of all the monitoring data [5].

Hot spot temperature gives a picture of the loading characteristics and of the possibility of some internal short circuits. Vibration analysis is performed to detect changes in the internal structure of windings [6], whereas frequency response analysis does the same evaluation but including the core and bushings [7], [8]. The dissolved gas analysis checks for the existence of faults related with high temperature of the oil such as arcing, corona discharges or hot spots [9]. Finally, partial discharge analysis look for incipient faults in the insulation system, and it is usually performed through acoustic methods [10].

1.1.2 Maintenance Planning

There are three basic types of maintenance planning: corrective, preventive, and reliability centered maintenance (RCM) [3].

Corrective maintenance was applied in the early days of power transformers. Nowadays, it is only used for not important elements, i.e. when there are no crucial consequences in case of a fault. The actions to be taken in this maintenance are general inspection, electrical tests, and liquid insulating tests [11]. The first and the second actions are performed when the fault occurs in an accessory of the transformer. On the other hand, all three actions are made in case the whole transformer is in a fault condition.

Preventive maintenance can be divided in time based maintenance (TBM) or condition based maintenance (CBM) [3]. TBM means that the transformer is checked once within a constant period of time, for example, each year. The disadvantage is that the period could be too short, i.e., too many resources are spent in maintenance, or it could be too long, i.e., there is a higher risk of transformer failure. On the other hand, CBM is based on the results of the monitoring system. If it is seen that an important change has happened in the monitoring variables, then a maintenance action is planned and performed.

It could be the case that an intervention in any equipment increases the probability of failure. This has been seen, for example, in overhauling events. In other words, the maintenance activities worsen rather than improve the condition of the equipment. Hence, a maintenance where the importance of the equipment and the risk of failure are considered simultaneously, is advantageous. RCM is based on this consideration. So it is more focused on more important elements, and even recommends only corrective actions for those elements whose failure does have paramount consequences. RCM has the aim of reducing maintenance costs; however, one of its main disadvantages is that the personnel needs to have a deep knowledge about how the system works and the consequences of each failure [12].

1.1.3 Life Assessment

The asset management evaluation gives four types of life condition:

- ❐ Normal Condition
- ❐ Accelerated Aging
- ❐ Uneconomical Condition
- ❐ End of Life Status

The normal condition means that the transformer is operating as was expected during the design. There is no overload or an excessive number of fault events. The accelerated aging occurs when there are overload cases, for example, if working in parallel with other transformers, and the last is out of service for some time. The uneconomical condition means that the transformer can still work, but the power system must be operated far from the optimal point. Finally, if the transformer is no longer able to operate, it has reached the end of life status.

The lifetime of any device, including power transformers, could be analyzed from three points of view: physical lifetime, technical lifetime, and economical lifetime. Physical lifetime means that the transformer can no longer work as expected [13], maybe the winding ratio is not as original, or the insulation cannot withstand the power system voltages. Technical lifetime means that new technologies have been developed so that keeping the old technologies may be expensive or impractical [14]; this lifetime is not frequent in power transformers but a case of a new magnetic material for the iron could be an example of this situation. The economical lifetime ends when the transformer has zero asset value [15]; it is related to the amortization value that must be considered each year. This research is more concerned with physical lifetime.

Physical lifetime analysis can be divided into intransitive and transitive aging [16]. Intransitive aging refers to the power transformer deterioration when it is operating in normal conditions. On the other hand, transitive aging considers unusual operating conditions such as overloading, harmonics, higher temperatures, etc.

1.1.3.1 Intransitive Aging

Intransitive aging is mainly focused on the analysis of the insulating paper. Insulating paper is made of cellulose, which is formed by chains of polymers, Figure 1.2. It is generally accepted that the number of chains in a unit of polymer is a good reference of the age or the state of the insulating paper. This number is known as the degree of polymerization (DP). So, a DP of 1200 is considered to belong to a new power transformer [17], and a DP of 200 means that the transformer is no longer guaranteed to be able to operate in secure conditions[17], [18]. Ideally, a sample of the paper would be taken and then measure the DP. However, this action is intrusive and it is impossible to perform in an operating transformer.

It has been observed that the quantity of furans in the insulating oil are directly related with the DP. Furans are chemical components that are formed by the breaking of the polymer chains. There are five types of furans as shown in

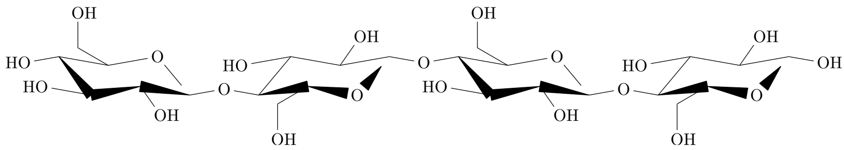
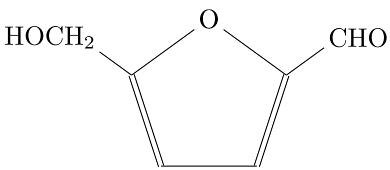
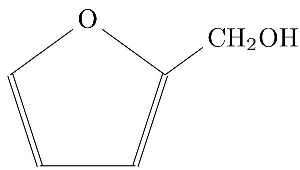


Figure 1.2: Chain of polymers in the insulating paper cellulose.

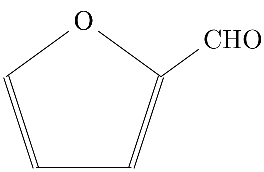
Figure 1.3. From all the furans, 2FAL (2-furaldehyde) is the most abundant [19] and has been found to be the most related to DP [20]. Thus, in practice, the content of furans in oil is measured, and then the DP is determined by (1.1) and (1.2) for the upgraded and non upgraded insulating paper respectively [21]. The amount of furans is measured in ppb per weight (μg/kg).



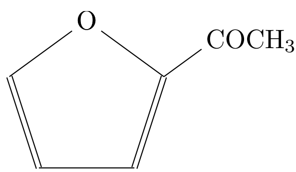
5-Hydroxy methyl-2-furfural
(5-HMF)



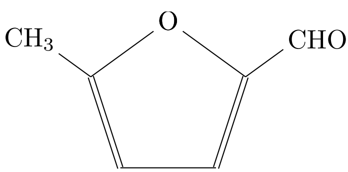
2-Furfurol
(2-FOL)



2-Furfural
(2-FAL)



2-Acethylfuran
(2-ACF)



5-Methyl-2-furfural
(5-MEF)

Figure 1.3: Furans dissolved in insulating oil. 2-FAL is the most used in the analysis of the degree of polymerization.

$$DP = -\frac{\log_{10}(\text{amount of furans}) - 4.51}{-0.0035} \tag{1.1}$$

$$DP = -\frac{\log_{10}(\text{2FAL} \times 0.88) - 4.0355}{-0.002908} \tag{1.2}$$

A disadvantage of determining the DP through the measure of furans is that the oil should be free of any previous treatment. During drying and filtering of the oil, some furans could be lost. Furthermore, care must be taken about the temperature during the tests and recollection of oil samples because the furans could react chemically and disappear.

It is worth noting that some advancements have been done in regards to the analysis of the DP. For example, fuzzy logic has been used to include the effects of thermal degradation and moisture [22], or using moisture content, interfacial tension, and furan content, together with a neuro fuzzy logic model [23].

1.1.3.2 Transitive Aging

The most used method for transitive aging is the analysis of temperature. The power transformer is designed to work for some pre-established load and electrical current. This current circulating through the winding produces heat and a correspondent value of temperature. A too high temperature could mean the presence of overload.

Nowadays, it is set that the hottest spot temperature for a rated load should be 110° [24]. This value is taken as the base for per unit calculations. In Chapter 4 the variation of temperature with load is used to find the lifespan of the power transformer in regard to the insulation.

1.2 RESEARCH JUSTIFICATION

As it has been exposed, the asset management of power transformers is mainly focused on the analysis of the insulating system. Almost nothing has been said about the deterioration of the winding conductor, although it is well known that during the power transformer operation lifetime, the winding conductor suffers the effects of electromagnetic forces.

Faults in windings, Figure 1.4, are about the 19.2 % of all the faults in power transformers [25]. From those faults, many are due to electromagnetic forces in the windings which cause deformations that, together with the deterioration of the insulation, could cause the failure of the whole transformer [26]. This problem has become so serious that many studies have been developed, mainly focused on the defformation of the windings [27]–[29]

Usually, the ability to withstand the electromagnetic forces caused by the highest fault current is evaluated. This evaluation is even recommended in the design review process. But, this evaluation is taken from a static perspective, although the electromagnetic forces can also continuously affect the conductor, both under normal and fault transient operating conditions. This research analyzes and describes the dynamic effects of the electromagnetic forces and compares the results with those of a dynamic load, i.e. it performs a fatigue analysis of the electromagnetic forces.

In practice, the fatigue analysis could be included in asset management as a complementary procedure. Thus, besides all the monitoring of the insulating materials, the evaluation of the winding conductors' state would also be consid-

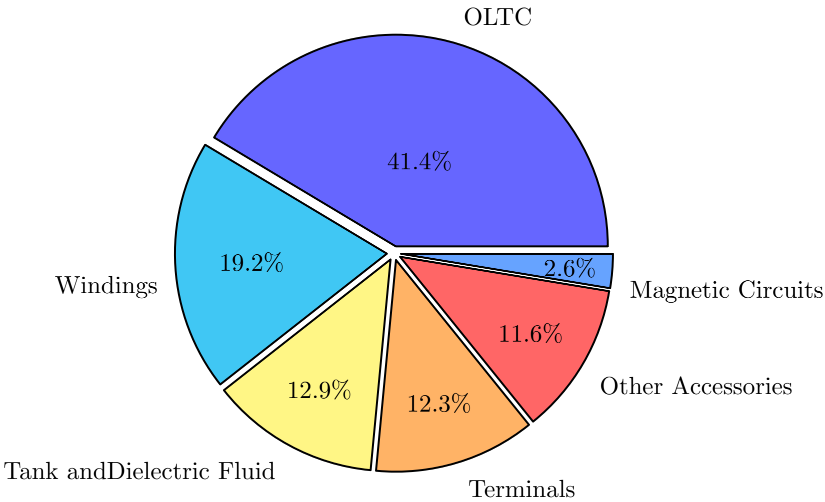


Figure 1.4: Statistics for the fault modes in power transformers.

ered. In that way, regarding the results related to transformer aging, there could be some additional, considering the conductor deterioration:

- ❒ Normal condition: The conductor deterioration is similar to that of the insulating system.
- ❒ Accelerated aging: The conductor deterioration is faster than that of the insulating system, maybe due to the presence of faults too close to the transformer.
- ❒ Uneconomical condition: The transformer requires reducing the exposure to high current faults, perhaps the utilities must reduce the power system fault currents, which might be an uneconomical solution.
- ❒ End of life status: The fatigue analysis could predict the possible end of the lifetime of the power transformer, given that the operating conditions do not change significantly.

In regard to the theory of power transformers, the contribution of the research is to give a deeper knowledge about the dynamic effects of the forces in winding conductors. The fatigue analysis allows studying the effects that the mechanical stresses continuously make on those conductors. Hence, the electromagnetic forces become important during the operation of power transformers and not only during the design process.

Finally, as a byproduct, a procedure for the real-time determination of the mechanical stresses is developed. Thus, it is expected that the methodology that is exposed in this research may be applied to actual power systems and may be included in the current operating practices.

1.3 RELATED WORK

Not too much research has been done regarding the study of the effects of fatigue in winding conductors. Generally, the investigation has been limited to the static effects of electromagnetic forces. In a direct fashion, two works have been found and are described in the following sections.

1.3.1 Inrush Current and Cold Load Events

When a system has been out of service for several minutes and is reenergized, the current could be much higher than the one previous to the event. This phenomenon is known as cold pick up [30]. The reasons for the higher values of current are load inrush currents, loss of load diversity, magnetizing currents, etc. Note that the inrush current could become between 10 to 20 times the rated current [31].

Experimental setups to find the effects of cold pick up load when energizing a distribution transformer have been performed. Thus, it has been found that the continuous events could cause a transformer to reach its life end, previous to the time it was expected [32].

Considering that fatigue failures could be from 25% to 75% of the mechanical faults in the engineering industry [33], [34], Beniwal et al. researched about the effect of fatigue in distribution transformers when subjected to energization events [32].

The hypothesis in their work was that since fatigue is responsible for mechanical equipment deterioration, the same may happen for winding conductors. Hence, they investigated how aluminum and copper winding conductors are affected by fatigue when subjected to energization events similar to those of cold load events. The main concern was the fact that many non-well-conditioned systems could have a great amount of cold load, which worsens the effects of transformer energization.

The authors performed tests to find the number of cycles that the conductor sample could withstand without failure. They found the SN curve for both copper and aluminum. Copper had a better strength regarding mechanical stress. A lgood contribution of this research was the construction of a homemade artifact to apply and measure the mechanical stress.

The authors set a number of energization events per day. With the relation between the number of cycles the conductor can withstand and the number of energization events per day, the number of days that the transformer is supposed to live is calculated. A better approach would have been the inclusion of a failure probability to get a closer approximation of the transformer operating condition.

As a main result, it was seen in this research that the distribution transformers had less life than designed.

1.3.2 Fault Currents Fatigue Analysis

Araujo et al. gave a different approach [35]. They studied the fatigue effects of fault currents and compare them to the generally applied Von Mises criterion, which verifies that the transformer can withstand electromagnetic forces resulting from the worst fault. The particularity of the Von Mises criterion is that the analysis is performed from a static perspective.

In Araujo's research, the hypothesis is that even stresses lower than those of the original design, applied continuously, such as in the transient period of an electrical fault, can destroy the transformer winding.

For the static analysis, the authors use the Von Mises criterion, which employs the three components of the mechanical stress, contrary to the recommended criterion that utilizes only the component in the direction of the stress, i.e. perpendicular to the cross-section of the conductor [36]. On the other hand, for the fatigue analysis, the well known Miner's rule is applied. This rule assumes that the steady deterioration of any material relates linearly with the number of cycles the stress acts on that material.

For the transient period of the fault, the authors made simulations considering various resistance/reactance relations. Moreover, three fault duration times are used: 0.05, 0.25 and 3 seconds. The first time is taken as an immediate fault clearance, 0.25 as a backup clearance, and 3 seconds are taken as the maximum time that a fault can be present in the system.

In summary, the authors found that depending on the time constant of the transient, the winding conductor can deteriorate with much lower mechanical stresses than that of the original design. Besides, they also analyzed the cases where the conductor is already damaged because of previous fault currents. If the previous damage is high e.g. 90% of loss of life for a predetermined stress, a minimum fault could end the transformer life.

A limitation of the study is the fact that only fault currents are considered in the analysis. Thus, the normal operating conditions of the transformer are neglected.

1.4 OBJECTIVES

According to what has been exposed before, the main objective of this investigation is to determine how the mechanical stress, produced by fault and inrush currents, affect the life of the power transformer. The main difference with the previous research is threefold:

- ❒ The effects of both fault and inrush currents are considered.
- ❒ The transformer life is analyzed from a global perspective, as part of a power system. Reliability indexes are considered so that the simulations are closer to a real condition of operation.
- ❒ A nonlinear fatigue model is used so that a comparison with the linear model could be performed, as well as the limits of using nonlinear models in this type of problem.

A second objective is to compare the lifespan of the conductor and the insulating system. For this analysis, the procedure recommended by IEEE Std C57.91-2011 [24] is followed. The question, in this case, is under which conditions the conductor deteriorates faster than the insulating material.

A third objective is a comparison between aluminum and copper conductors. The question is how much the life of the conductor is influenced by the material. Note that, in the transformer design for aluminum and copper, the current density, the cross-sections and so, the mechanical stresses are different.

Finally, the Miner's rule results are compared with nonlinear models, which use only SN curves so that non-complexity is added to the final model. Hence, it should be clear if the Miner's rule is an optimistic or pessimistic model when evaluating the life of the winding conductors.

2 THEORETICAL FRAMEWORK

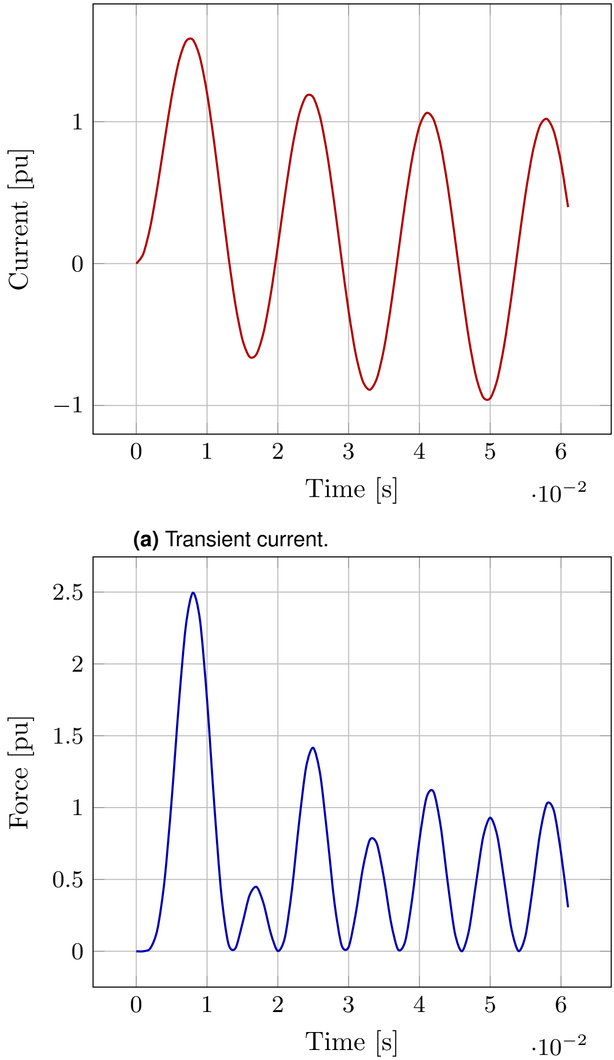
2.1 ELECTROMAGNETIC FORCES

2.1.1 Problem Statement

If an electrical current I is circulating through a conductor immersed in a magnetic field B, there is a force F acting on the conductor. If the conductor has length l, and if the magnetic field makes an angle α with the current, in scalar form, the magnitude of the force is given by (2.1), which is known as Lorentz Force [37]. It can be seen that, as B also depends on the current, F varies as I squared. Moreover, this relationship doubles the frequency of the force, i.e. the force pulses appear twice each power system period. For example, compare the periods in Figure 2.1 [38].

$$F = B \cdot I \cdot l \sin \alpha \tag{2.1}$$

For a point in space the force per unit of volume $\mathbf{f}$ can be expressed by (2.2), where $\mathbf{J}$ is the current density and $\mathbf{B}$ is the vector of magnetic induction.



(a) Transient current.

(b) Transient force.

Figure 2.1: Transient behavior of current and force before a fault on the system.

$$\mathbf{f} = \mathbf{J} \times \mathbf{B} \tag{2.2}$$

It is seen that before calculating $\mathbf{f}$, $\mathbf{B}$ must be determined. Due to the fact that the magnetic field is irrotational, a magnetic vector potential $\mathbf{A}$ can be defined by (2.3).

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{2.3}$$

Thus, the first step in finding the force is the determination of $\mathbf{A}$. From the quasi static Maxwell's equations, it can be demonstrated that $\mathbf{A}$ can be obtained from the Poisson's equation for magnetics (2.4), where μ is the permeability of the medium [39].

$$\nabla^2 \mathbf{A} = -\mu \mathbf{J} \tag{2.4}$$

The Poisson equation is an elliptic partial differential equation [40]. It cannot be solved analytically except for some cases where one of the coordinates is constant, such as a sphere, a cylinder or a capacitor with infinite parallel plates [41]. Fortunately, the advancement of numerical methods has permitted to obtain solutions to those partial differential equations. From these, the Finite Element Method (FEM) is the most spread among the electrical engineering.

2.1.2 Finite Element Method

The characteristic of FEM that has made it popular in the engineering applications is its capacity to solve problems with complex geometry. In the following paragraphs, a brief description of the FEM theory is presented.

In its basic form, an elliptic partial differential equation can be expressed as a boundary problem. It has the form (2.5), where $\mathcal{L}$ is a differential operator, f is the excitation function and ϕ is the unknown field [42].

$$\mathcal{L}\phi = f \tag{2.5}$$

The boundary problem for two-dimensional geometries is shown in (2.6). The coefficients α_x, α_y and β_x characterize the medium and Ω represents the set of points in the medium [42].

$$\begin{aligned} -\frac{\partial}{\partial x}\left(\alpha_x \frac{\partial \phi}{\partial x}\right) - \frac{\partial}{\partial y}\left(\alpha_y \frac{\partial \phi}{\partial y}\right) + \beta\phi = f \\ (x, y) \in \Omega \end{aligned} \tag{2.6}$$

The problem is complemented by the boundary conditions (2.7). On the boundary Γ_1 the value p is set for the field ϕ, this conditions is known as Dirichlet boundary condition. On Γ_2, the behavior of the angle of the going out field ϕ with respect to Γ_2 on each point is established. The presence of the unit vectors $\hat{x}$, $\hat{y}$ indicates that the expression between brackets is the scalar product between the vector of the material characteristics and the divergence of the field. The unit vector $\hat{n}$ represents the direction of the normal vector to the boundary. This boundary condition on Γ_2 is known as the Neumann boundary condition. It is worth to note that the boundary conditions cause the existence of a unique solution for the problem; without the specification of at least one boundary condition, the problem would have infinite solutions.

$$\begin{aligned} \phi = p \quad \text{on} \quad \Gamma_1 \\ \left(\alpha_x \frac{\partial \phi}{\partial x}\hat{x} + \alpha_y \frac{\partial \phi}{\partial y}\hat{y}\right) \cdot \hat{n} + \gamma\phi = q \quad \text{on} \quad \Gamma_2 \end{aligned} \tag{2.7}$$

It can be demonstrated that the boundary problem is equivalent to the variational formulation given by (2.8), where F is given by (2.9) [42]. Note that the Neumann boundary condition is included in the functional, whereas the Dirichlet conditions are imposed externally.

$$\begin{aligned} \delta F(\phi) &= 0 \\ \phi &= p \end{aligned} \tag{2.8}$$

$$\begin{aligned} F(\phi) = \frac{1}{2} \iint_{\Omega} \left[\alpha_x \left(\frac{\partial \phi}{\partial x} \right)^2 + \alpha_y \left(\frac{\partial \phi}{\partial y} \right)^2 + \beta \phi^2 \right] d\Omega + \\ \int_{\Gamma_2} \left(\frac{\gamma}{2} \phi^2 - q\phi \right) d\Gamma - \iint_{\Omega} f \phi \, d\Omega \end{aligned} \tag{2.9}$$

FEM is performed in the following steps [43]:

- Discretization of the domain: For two-dimensional problems, the elements are usually triangles.
- Define the interpolation functions: The linear interpolation is the simplest function to be implemented.
- Incorporation of the boundary condition.

The field ϕ within an element e is (2.10), when approximated by a linear function interpolation. The coefficients a, b and c depend on the geometry of the discretization, i.e. on the position of the vertices and on the area of each element.

$$\phi^e(x, y) = a^e + b^e x^e + c^e y^e \tag{2.10}$$

Although not an obligatory requisite, it is recommended that the parameters α_x,

α_y, β and f are constant inside each element. In that way, the implementation of the method is easier.

After finding the extreme of the variational F, i.e. after applying (2.8) to (2.9), and with some algebraic treatment, the matrix K and the vector b can be constructed by (2.11) and (2.12), where i and j denote the matrix row and the column respectively, $\triangle^e$ is the area of the element e, and δ_{ij} is the Kronecker delta.

$$K_{ij}^e = \frac{1}{4\triangle^e}\left(\alpha_x^e b_i^e b_j^e + \alpha_y^e c_i^e c_j^e\right) + \frac{\triangle^e}{12}\beta^e\left(1+\delta_{ij}\right) \tag{2.11}$$

$$b_i^e = \frac{\triangle^e}{3} f^e \tag{2.12}$$

Finally, the matrix equation (2.13) is obtained. A detailed exposition of the derivation of the equation is presented by Jin Jian Ming [42].

$$K\phi = b \tag{2.13}$$

2.1.3 FEM Applied to Magnetostatics

In two-dimensional problems, the current is taken as directed along the z axis. In agreement with (2.3), the magnetic vector potential has the same direction to the current density, therefore, $\mathbf{A}$ also has only a z component. Thus, the Poisson equation is expressed as (2.14). Developing this equation in rectangular coordinates x and y, (2.15) is obtained.

$$\nabla \times \left(\frac{1}{\mu_r} \nabla \times A_z \hat{z} \right) = -\mu_0 J_z \hat{z} \tag{2.14}$$

$$-\frac{\partial}{\partial x} \left(\frac{1}{\mu_r} \frac{\partial A_z}{\partial x} \right) - \frac{\partial}{\partial y} \left(\frac{1}{\mu_r} \frac{\partial A_z}{\partial y} \right) = \mu_0 J_z \tag{2.15}$$

Therefore, the magnetostatics equation has the form of a boundary problem and can be solved by FEM if the varaibles are set according to (2.16).

$$\begin{aligned} \alpha_x = \alpha_y &= \frac{1}{\mu_r} \\ \phi &= A_z \\ f &= -\mu_0 J_z \end{aligned} \tag{2.16}$$

2.1.4 Iterations in FEM

The discretization should be enough to get results approximate to the actual phenomenon. If the discretization has too little points, the approximation of the method will give results far from the actual values. On the other hand, if there are too many points, the internal treatment of the matrices will be computationally expensive, and sometimes they will even make the problem impossible to be solved.

FEM looks for a good level of discretization. For that purpose, the method begins with an approximate discretization, according to the geometry of the problem [44]. For example, there will be more points in regions with corners or where the geometry abruptly changes.

After the initial discretization, FEM runs a simulation and finds the first results, i.e. the first values of the magnetic vector potential A_z for the case of the power transformer. Then, there are two options to verify the goodness of the discretization:

1. To compare the values of A_z in the vertices of each triangle. If they are too different, a denser discretization is created in that region. Then a new simulation is performed, and the process is repeated until some tolerance is achieved [45].
2. To compare the values of A_z for two simulations, with the second discretization denser than the first one. If the results are too different, a third discretization, denser than the second, is performed, and a new simulation is run. The comparison and the creation of new discretization continues until a tolerance is achieved [44].

Therefore, independent of the method followed for evaluating the discretization, an iterative process is followed. On each iteration, a matrix equation is solved.

In the case of power transformers, the nonlinear characteristic of the core permeability worsens the discretization procedure. The boundary problem includes the permeability, and some values are assumed at the beginning. Then the results for the magnetic induction are obtained, but these results may not correspond to the permeability assumed, according to the saturation curve. As a consequence, a new permeability is obtained, closer to the magnetic field found before, and a new simulation is run. The process continues until the magnetic induction and the permeability agree to what is expected from the saturation curve.

In summary, there exist iterations for both the discretization and the saturation curve. Due to these iterations, FEM takes a lot of time for each electrical current state. When the objective is to evaluate the forces or the mechanical stresses for the worst condition, this amount of spent time is not an issue. However, if the required analysis is to be used in a continuous basis, like the fatigue analysis of this research, FEM is not suitable, and a new method must be developed.

2.2 MECHANICAL STRESS

By FEM the magnetic induction and the electromagnetic force are calculated. The power transformer must be designed to withstand the highest forces. The possible required actions include variations in the internal dimensions and the use of stronger supports in the winding structure. It is expected that the transformer winding will not suffer any deformation after the completion of the design.

A side consequence of the transformer design, and the final winding support structure, is that there is no vertical movement of the winding disks, at least under the design operating conditions. There is only one degree of freedom in the expansion of the disks, which corresponds to the radial forces.

For that reason, only the mechanical stresses caused by radial forces were considered for fatigue analysis. Note that even if vertical forces were unbalanced (i.e. not nullified by the structure supports), the direct effect would be a vertical movement of the disk, and not a fatigue damage. On the other hand, for radial forces, there is no structure that contains the movement of the disk. The result is a constant expansion and compression of the conductor, with all the fatigue effects.

With the previous considerations, the winding disk was modeled as a ring as shown in Figure 2.2. When the radius of the disk is much greater than the radius of the conductor, the mechanical stress σ can be calculated by (2.17) and (2.18) [46], where f_r is the radial force, R_{ring} is the radius of the equivalent ring, P is the cross-sectional force, and Sc is the cross-section of the conductor.

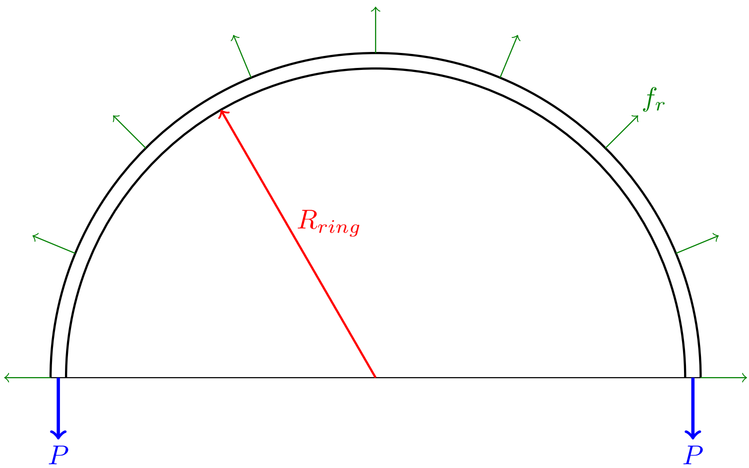


Figure 2.2: Ring model for the determination of the cross sectional force in a winding disk.

$$P = f_r \cdot R_{ring} \tag{2.17}$$

$$\sigma = \frac{P}{Sc} \tag{2.18}$$

2.3 RANDOM FORESTS

2.3.1 Machine Learning Basics

Random forests belong to the set of Machine Learning tools. Machine learning is understood as the process where a computer system can learn from past data.

The learning may refer to discovering patterns or relationships in the existing data [47].

The problems treated by machine learning are of two types: classification and regression problems. Classification problems are related to categorical variables, those that belong to a finite set and cannot be ordered. On the other hand, regression problems handle numerical variables, which are ordered and belong to the set of real numbers $\mathbb{R}$ [48].

Random forests are of the supervised learning type. As that, it needs the existence of input variables or features that are part of input vectors or instances, and output vectors or targets [49]. From those input/output data, the learning process proceeds until a model is found.

The input/output data form the learning set $\mathfrak{L}$. The elements of this set come in pairs of vectors ($\mathbf{x_i}$, $\mathbf{y_i}$). The vector $\mathbf{x_i}$ is an instance of the input data and the vector $\mathbf{y_i}$ is an instance of the output data.

If $\mathbf{X}$ is the matrix containing all the instances $\mathbf{x_i}$, and $\mathbf{Y}$ is the matrix containing all the targets $\mathbf{y_i}$, the learning process can be defined as finding a function $\varphi_{\mathfrak{L}}$ modeling the relation between $\mathbf{X}$ and $\mathbf{Y}$ from the learning set $\mathfrak{L}$.

The generalization error err of the function $\varphi_{\mathfrak{L}}$ is defined by (2.19). In other words, the error is the expected value $\mathbb{E}$ of some function L that measures the difference between the output and the function φ of the model. Note also that the error depends on the learning set. A different learning set would yield a different error.

$$\mathrm{err}(\varphi_{\mathfrak{L}}) = \mathbb{E}\left\{L(\mathbf{y}, \varphi_{\mathfrak{L}}(\mathbf{x}))\right\} \tag{2.19}$$

For regression problems, the most common function L is the squared error defined by (2.20) [50], [51].

$$L = [\mathbf{y} - \varphi_{\mathfrak{L}}(\mathbf{x})]^2 \tag{2.20}$$

The learning set $\mathfrak{L}$ does not contain all the possible cases of the problem or phenomenon to be modeled. Actually, the elements are probabilistically chosen. Thus, if the probability distribution function were known, (2.19) would give the error for the whole set of instances in the problem [52]. But, as it is usually the case in regression problems, the set of instances is infinite, and so the error as formulated in (2.19) it cannot be exactly calculated. It is worth noting that the interest is settled in knowing the generalized error, i.e. the error involving the data not only in $\mathfrak{L}$ but also in those possible cases outside $\mathfrak{L}$.

The simplest method to determine the generalized error is to assume that the error comes from the learning process [52]. This is not a good assumption; this is a too optimistic error, since it is minimized during the training process.

Another method, which is mostly used in machine learning design, is to divide the learning set into training and testing sets. Only the training set is used as a reference for the learning process [53]. Then the generalization error of the model is calculated by making predictions in the testing set. Sometimes the testing error is calculated during training, so that the process finishes when the testing error begins to rise. This technique is known as early stopping. A drawback of this method of calculating the generalized error is the reduction in the number of samples to train the model because 10% to 30% of the whole learning set is taken as testing data [52].

Independently of the model or the machine learning technique employed, there is a minimum error value that cannot be reduced. Moreover, this error could be determined only if the true function is known, which is never the case. The model that has the minimum error value is known as the bias model φ_B [54].

In spite of the impossibility of obtaining φ_B, an approximate model can be found by varying some variables at hand, depending on the machine learning tool. These variables that the designer can vary are known as hyperparameters [55].

The process of finding the best combination of hyperparameters is a deterministic procedure. Some authors suggest using a grid to form the combinations [56], [57], while others propose the use of random variables with or without using the Monte Carlo method [58], [59]. In any case, the designer of the machine learning model must be aware that the model should not be too simple, neither too complex.

If the model is too simple, it is wasting information kept in the learning set, and the model is not close to the bayes model. When this is the case, the design has fallen in underfitting [52].

If the model is too complex, the function is representing the relation between the input and output data, together with the noise in the data. As a consequence, the model fits too well the output signal, but it poorly models data outside the learning set, i.e. the model is overfitting [52].

During the design, the complexity of the model can be monitored by examining the behavior of the training and testing error. When the model is too simple, both errors are high. As long as the complexity increases, both errors decrease until they reach a point where the training error continues decreasing, but the testing error begins to increase. This is the optimal point of the design, where the complexity of the model has reached a balance, and is the point that marks the early stopping, beyond which it is not recommended to continue the training process [60], [61]. If the model becomes is too complex, the training error is very low and the testing error is high.

2.3.2 Decision Trees

The machine learning methods can be divided in parametric and non-parametric. Parametric methods are those in which the designer makes previous assumptions about the model. For example, if a linear regression model has been chosen, the designer decides the degree of the polynomial, the number of terms, and the like. Hence, linear regression is parametric [62].

On the other hand, in non-parametric models, no previous assumption is made. The size of the model or the inner parameters are decided by the algorithm itself. The decision trees method is non-parametric. Except when used as stopping criteria, there is no limitation regarding the number of leafs or elements inside each branch [63].

Another characteristic of decision trees is that they are easy to be interpreted. In other machine learning models, it may happen that they become black boxes which nobody knows what is inside. This happens, for example, in artificial neural networks. On the other hand, as decision trees are formed by nodes of decisions, once they are created, it is easy to follow the logic inside the model.

Some definitions related with decision trees are

Tree Graph whose nodes are connected by only one path.

Parent Node that is connected downwards to a child node.

Root Node that has children but not a parent.

Leaf Node that has a parent but no children.

Internal Node Node that has a parent and children.

Thus, a decision tree is one with the following characteristics:

- ❒ It has a root node.
- ❒ The root node contains all the input data.
- ❒ The intermediate nodes contain subsets of the input data according to rules given on each previous node.
- ❒ The leaf node sets the input data in a place where the output data can be obtained.

Figure 2.3 gives an example of a decision tree. It models the function $f(x) = x^2$. Although this is a basic function, it helps in clarifying some principles that have been defined. The tree is constructed so that, at the end, three subsets are formed, thereby there are three leaves. If, for example, the prediction of $x = 0.3$ is desired, the path followed by the model would be (1) to the left in nd_0 and (2) to the right in nd_1, therefore $f(x) = 3.64 \cdot 10^{-2}$. The actual value is $f(x) = 9 \cdot 10^{-2}$. Following the prediction procedure of the example, it can be visualized why decision trees models are considered more interpretable than other machine learning types, black box models, such as deep neural networks or support vector machines.

There are some things that could be seen in Figure 2.3 that characterize a decision tree. First, the output value for regression problems is the average of the set of outputs according to the input values in the leaf. This is a limitation for decision trees models, because it means a discrete output. Second, the deeper

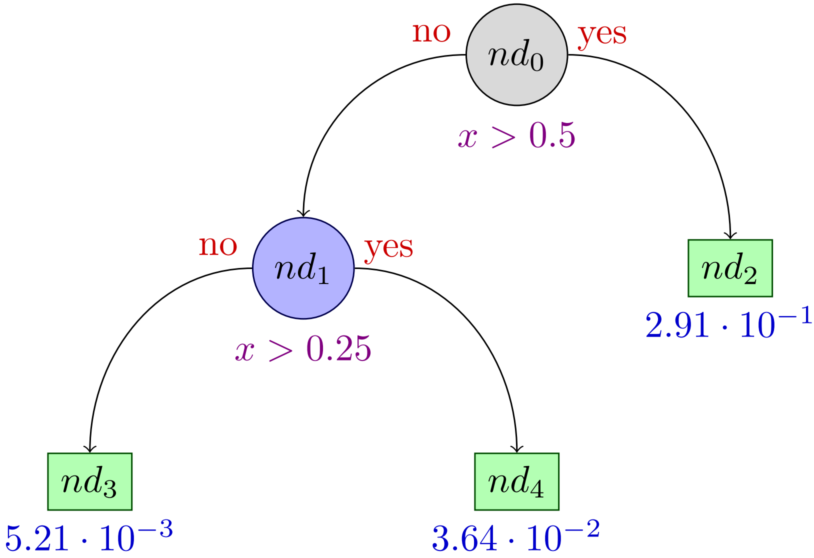


Figure 2.3: Example of a decision tree for the function $f(x) = x^2$. nd_0 is the root node, nd_1 is an intermediate node, and nd_2, nd_3 and nd_4 are the leaves.

the layers, the more accurate is the model. This is a consequence of the previous characteristic because if there are more output subsets, the discretization will be closer to the actual value. Finally, not all the branches in the tree have the same deepness. In the example, the leave nd_2 is directly connected to the root node nd_0, whereas the other two leaves nd_3 and nd_4 are connected to an intermediate node nd_1.

The training process of a decision tree consists in finding the optimal structure in regard to the number of nodes and the deepness of the tree [52]. In other words, the training process looks for the smallest possible tree that represents the problem. This objective is directly related to the split criteria executed on each node.

Formally speaking, when the loss is defined as the squared error (2.20), the value on each leave is defined as the mean value of the outputs corresponding to the inputs that fall into that leave (2.21) [52]. Note in the equation that the learning set has been reduced so that only the learning points of each node belong to it.

$$y_{nd} = \frac{1}{N_{nd}} \sum_{x,y \in \mathfrak{L}_{nd}} y \tag{2.21}$$

If the number of branches and leaves of the decision tree is unlimited, there is a point where each leave will represent an individual point of the training set. This model would have a zero training error because each element of the training set would be represented on at least one leaf. However, the noise in the signal would also be represented, and the model would not be prepared to predict values outside the learning set. In other words, the model has the problem of overfitting.

In order to avoid overfitting, the training must stop, i.e. none of the nodes is any longer divided, at some adequate point. In practice, a node is not split and becomes a leave under the following circumstances [52]:

1. When all the output values y_{nd} in the node are the same. In this condition, no split criterion could be deduced because the impurities cannot be differentiated. A particular case of this condition is when the number of elements in the leave is one.

2. When all the input values x_{nd} in the node are the same. In this condition, no split criterion could be applied.

3. If the designer has set a minimum number of samples in the nodes and the node has reached that number.

4. If the designer sets a maximum number of nodes that could be followed in the decision path. In other words, if a maximum depth has been pre-established.

5. If the designer has set a minimum decrease in impurity for the node.

6. If the designer has set a minimum number of samples in the resulting leaves after the split.

The first two conditions result naturally from the training process and cannot be controlled. The other conditions can be controlled and are useful to obtain an optimal model, avoiding too simple or too complex models.

In regression problems where the accuracy is measured by the squared error, the impurity of a node imp_{nd} is evaluated by the variance of y_{nd} (2.22), where $\hat{y}$ is the mean value of the outputs in the node [52].

$$\mathrm{imp}_{nd} = \frac{1}{N_{nd}} \sum_{x,y \in \mathfrak{L}_{nd}} (y_{nd} - \hat{y}_{nd})^2 \tag{2.22}$$

When the phenomenon to be modeled has multiple outputs, there may be two options: (1) Build one decision tree for each output or (2) Create a single model for all outputs. The first option is easier to understand, but has the main disadvantage of losing information about correlations that may exist between the outputs [64]. Furthermore, the time of design and execution of the model would increase. As long as possible, the second option must be followed. For multiple output models, the measure of impurity is the mean value of the impurities for each output, and the optimal split is determined in base of that value.

2.3.3 Random Forests as Ensemble Systems

It is commonly known in machine learning theory that the error of a model can be expressed in three terms: the noise error, the bias and the variance. For models created from different learning sets for the same problem, the expected value of the error is (2.23) [52], [65]. The definition of the three terms is shown in (2.24).

$$\mathbb{E}_{\mathfrak{L}}\{\mathrm{err}\left[\varphi_{\mathfrak{L}}(x)\right]\} = \mathrm{noise}(x) + \mathrm{bias}^2(x) + \mathrm{var}(x) \tag{2.23}$$

$$\begin{aligned} \mathrm{noise}(x) &= \mathrm{err}\left[\varphi_B(x)\right] \\ \mathrm{bias}^2(x) &= \left[\varphi_B(x) - \mathbb{E}_{\mathfrak{L}}\{\varphi_{\mathfrak{L}}(x)\}\right]^2 \\ \mathrm{var}(x) &= \mathbb{E}_{\mathfrak{L}}\left\{\left[\mathbb{E}_{\mathfrak{L}}\left\{\varphi_{\mathfrak{L}}(x)\right\} - \varphi_{\mathfrak{L}}(x)\right]^2\right\} \end{aligned} \tag{2.24}$$

The noise is the error of the bayesian model, i.e. it is the minimum error which cannot be eliminated or reduced [54]. The bias is the error of the machine learning model in regard to the bayesian model. The variance measures the difference between models created from different learning sets.

In general, simple models have high bias and low variance. For their simple nature, the models have only a slight variation when the learning set is changed [65]. However, they are far from the bayesian model.

On the other hand, complex models try to be closer to the learning set. For that reason, when the learning set is changed, so does the model. Thus, complex models have low bias but high variance.

Therefore, the designer looks for a model which is balanced in variance and bias. Too high variance means overfitting, and too high bias means underfitting.

The ideal situation would be to have a model with low bias and that does not increase the variance while becoming more complex. For decision trees alone, that model is not possible. However, if different decision trees are combined to work together, such as in ensemble systems, a model closer to ideal could be obtained. That is the main characteristic of Random Forests (RF) [52].

RF are ensemble systems whose units are formed by decision trees. Each decision tree is built according to random generation of hyperparameters. If ξ is a particular set of hyperparameters for a single decision tree (e.g. maximum depth, minimum decrease of impurity, etc.), the expected value of the error of the model will now depend on the learning set $\mathfrak{L}$ as well as on ξ, as shown in (2.25) and in (2.26).

$$\mathbb{E}_{\mathfrak{L},\xi}\{\mathrm{err}\left[\varphi_{\mathfrak{L},\xi}(x)\right]\} = \mathrm{noise}(x) + \mathrm{bias}^2(x) + \mathrm{var}(x) \tag{2.25}$$

$$\begin{aligned} \mathrm{noise}(x) &= \mathrm{err}\left[\varphi_B(x)\right] \\ \mathrm{bias}^2(x) &= \left[\varphi_B(x) - \mathbb{E}_{\mathfrak{L},\xi}\{\varphi_{\mathfrak{L},\xi}(x)\}\right]^2 \\ \mathrm{var}(x) &= \mathbb{E}_{\mathfrak{L},\xi}\left\{\left[\mathbb{E}_{\mathfrak{L},\xi}\left\{\varphi_{\mathfrak{L},\xi}(x)\right\} - \varphi_{\mathfrak{L},\xi}(x)\right]^2\right\} \end{aligned} \tag{2.26}$$

The output value of a random forest is the mean value of the outputs from the decision trees. For instance, if there are M models created from different hyper-parameters ξ, the output value is given by (2.27) [52].

$$\psi_{\mathfrak{L},\xi_1,\ldots,\xi_M}(x) = \frac{1}{M}\sum_{m=1}^{M}\varphi_{\mathfrak{L},\xi_m}(x) \tag{2.27}$$

In regard to the noise, it can be seen in (2.26) that the noise of the RF is the same as that of the individual decision trees. This characteristic agrees with the definition of the bayesian model, which contains the minimum error for the phenomenon.

In the bias, the term $\mathbb{E}_{\mathfrak{L},\xi}\{\varphi_{\mathfrak{L},\xi}(x)\}$, when considering different combinations of ξ_m, has the same form of (2.24). Thereby, the bias does not change when combining the decision trees in a random forest.

For the analysis of the variance, first of all the Pearson's correlation coefficient $\rho(x)$ is defined by (2.28) [66] when determined for two models built from hyperparameters ξ_1 and ξ_2, where $\mu_{\mathfrak{L},\xi}(x)$ and $\sigma^2_{\mathfrak{L},\xi}(x)$ is the mean value and the variance respectively of the distribution of the random forests models according to the learning set $\mathfrak{L}$ and the variation of hyperparameters ξ. A $\rho(x)$ close to one means that the models are very related, i.e. there is not too much randomness in the creation of the hyperparameters. Otherwise, if $\rho(x)$ is close to zero, the models are not related at all, i.e. the randomness in the creation of hyperparameters is high [52].

$$\rho(x) = \frac{\mathbb{E}_{\mathfrak{L},\xi_1,\xi_2}\{\varphi_{\mathfrak{L},\xi_1}(x) \cdot \varphi_{\mathfrak{L},\xi_2}(x)\} - \mu^2_{\mathfrak{L},\xi}(x)}{\sigma^2_{\mathfrak{L},\xi}(x)} \tag{2.28}$$

Thus, it can be demonstrated that the variance of $\psi_{\mathfrak{L},\xi_1,\ldots,\xi_M}(x)$ can be expressed as (2.29). If the randomness is low, $\rho(x)$ is close to one and only the first term of the equation remains. Thereby, the variance of the whole ensemble random forest model is about the same as the individual decision trees, and there is no advantage in using random forests. On the other hand, if the randomness is high, $\rho(x)$ is close to zero, and only the second term remains. If the quantity of decision trees used M is high, the variance is very reduced in regard to the variance of single decision trees.

$$\mathrm{var}(x) = \rho(x) \cdot \sigma^2_{\mathfrak{L},\xi}(x) + \frac{1-\rho(x)}{M} \cdot \sigma^2_{\mathfrak{L},\xi}(x) \tag{2.29}$$

In conclusion, if the designer is careful about the correlation between the individual decision trees models, the utilization of random forests achieves the initial objective of getting a model with low variation in the bias and reduction in the variance. This is the main advantage of using random forests. As a complementary gain, the use of random forests allow the reduction in the discreteness of the outputs, due to the use of the mean value in the final output.

2.4 FATIGUE CUMMULATIVE DAMAGE

2.4.1 Strain - Stress Curves

A body subjected to the action of a force F is also subjected to the action of a mechanical stress S, (2.30), where A_0 is the cross-sectional area before the force is applied [67]. S is also known as the nominal engineering stress [68], which is different from the true stress that accounts for the variation of the area. This expression is not complete since in reality the stress is a tensor [69]. However, in this work only a scalar analysis is performed because only the axial component is considered.

$$S = F/A_0 \tag{2.30}$$

On the other hand, the tensile stress causes an extension of the body. If l_0 is the original length and l is the length after applying the stress, the strain is defined by (2.31) [70].

$$\varepsilon = \frac{l - l_0}{l_0} \tag{2.31}$$

In Figure 2.4, the different zones of the stress-strain relationship are shown [71]. Until the point of proportional limit (pl), the relation is linear and can be expressed as the Hooke's law (2.32), where E is the Young's modulus or modulus of elasticity. The point el represents the elastic limit, where the body can return to its original length if the stress is withdrawn. Beyond the elastic limit, the behavior of the material is plastic and cannot return as original even when the action of the stress is gone. S_y is the yield strength, usually taken as the reference to trace the line to get the yield point. The yield strength line is parallel to the proportional limit of the stress-strain curve, i.e. the slope is the same Young modulus. The point S_u deserves especial attention because it is important in the determination of the SN curve. It marks the maximum stress the material can reach in the stress-strain curve.

At this point, it is worth to make a distinction between strength and stress. Strength is an intrinsic property of a material and does not depend on the external conditions, for instance, there is the yielding strength, the ultimate strength, etc. In this research, the strength is symbolized by the letter S. On the other hand, the stress is part of the system where the material element is installed. To be clearer, the strength is always part of the element, it exists even before the element is installed in the equipment. In contrast, the stress only appears when the element is installed or is already working. The stress is symbolized by the Greek letter σ.

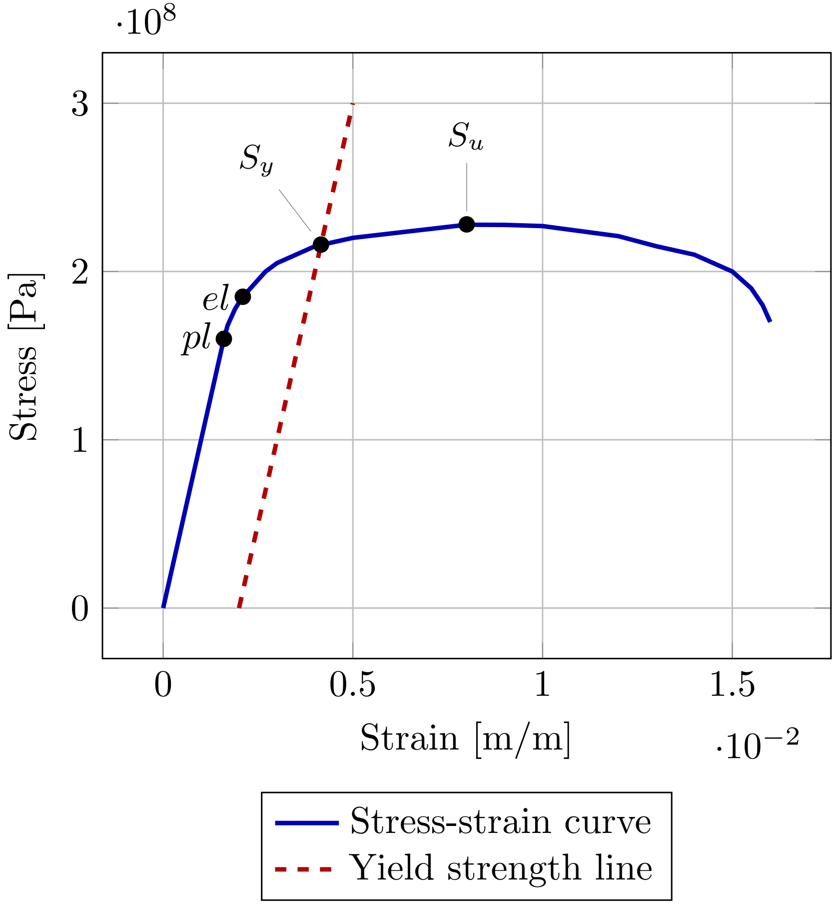


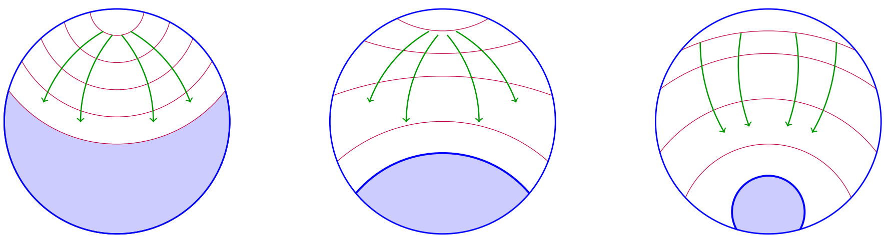
Figure 2.4: Zones of the stress - strain curve. The proportional limit is located until the point pl. The elastic limit is marked by the point el. S_y is the yield strength, from which the yield line is drawn to get the offset point. S_u is the ultimate strength.

$$\sigma = E\varepsilon \tag{2.32}$$

2.4.2 Fatigue Fundamentals

The phenomenon where a body is destroyed by the continuous action of a mechanical stress is known as fatigue [72]. It is important to note that the action must be continuous. This characteristic implies a dynamical effect. A big difference between a dynamic and a static failure is that the latter is usually visible and sometimes could be prevented with a good inspection. On the other hand, the dynamic failure is invisible until the material has already been broken.

The fatigue damage is produced in three stages, as Figure 2.5 sketches [73]. Firstly, one or more microcracks are initiated. These microcracks cannot be seen at plain sight, and often appear in places with notches. Some beach marks are present as seen in Figure 2.5a. In Stage Two, the stress is concentrated in the microcracks, thus increasing their size and the extension of the beach marks, Figure 2.5b. Finally, at Stage Three, the beach marks occupy most of the body's cross section, and the fracture occurs. In Figure 2.5c it could be seen the filled area where the fracture is produced. When evaluating a fatigue failure, the inspectors can recognize the beach marks and the area with the fracture.



(a) Stage I **(b)** Stage II **(c)** Stage III

Figure 2.5: Fatigue stages.

The factors that may accelerate the crack propagation are [74]

1. Residual tensile stress
2. Corrosive environment
3. Elevated temperatures
4. Temperature cycling
5. High-frequency cycling

From them, the last three are clearly part of the operating conditions of a power transformer. Firstly, the temperature around the windings when the transformer is working at rated load is 110° [24]. Secondly, in regard to the power load, it is not constant and, as a consequence, the temperature is subjected to variations from hour to hour [75]. Finally, the mechanical forces vary at twice the power frequency (e.g. 120 Hz in a 60 Hz power system) and so does the mechanical stresses.

2.4.3 Basics of Stress-Life Analysis

In a stress based test to determine the fatigue characteristics, the sample is subjected to cycles of constant stress until it is broken [76]. The locus formed by the pair number of cycles - stress is known as SN curve. Once the SN curve is obtained for a determined material, it can be used to perform a fatigue analysis.

The determination of the SN curve is a statistical phenomenon. Multiple samples are subjected to the test and the curve with the 50 % of probability of occurrence is generally published for future use [77].

The SN tests begin with a stress a little lower than S_u. Then, the value of stress is reduced and the SN curve is sketched [76]. Some metals present a lower limit, known as endurance limit S_e, under which no fracture is presented no matter how many cycles of stress are applied [78]. Nonetheless, nonferrous metals, such as copper or aluminum, do not have S_e [68]. An example of an SN curve in a log-log scale can be seen in Figure 2.6.

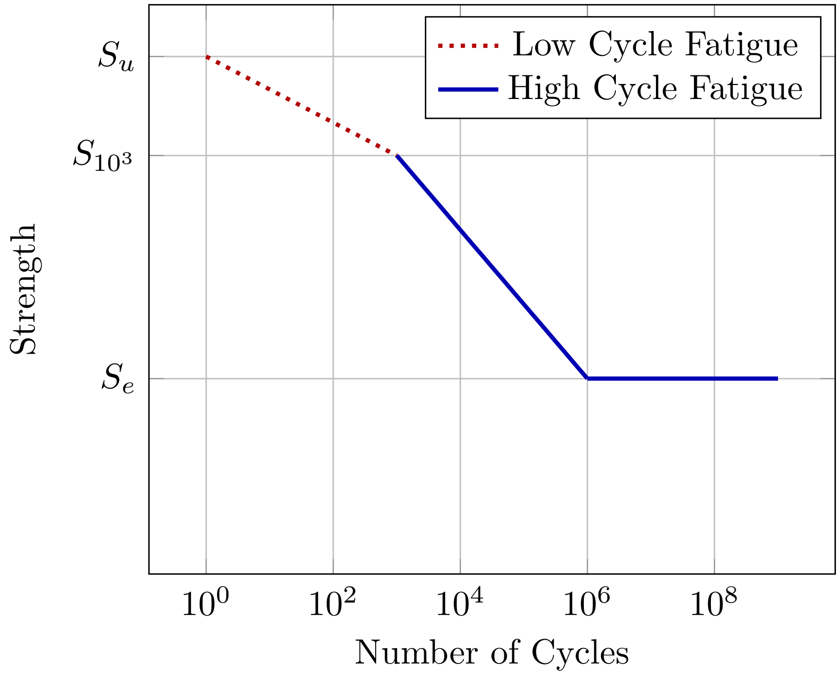


Figure 2.6: SN curve. The low cycle is located in $1 < N < 10^3$ whereas the high cycle in $N > 10^3$. For ferrous metals, the endurance limit S_e is located at $N = 10^6$ or $N = 10^7$.

Those events where the cumulative fatigue damage happens between $N = 1$ and $N = 10^3$ cycles are known as low cycle fatigue. They are characterized by a relative high mechanical stress and a little difference in regard to S_u [79].

For $N > 10^3$, the phenomenon is known as high cycle fatigue. In the SN curve, it is represented by the Basquin's formulation (2.33) [80], where S_f is the strength where the fracture occurs, and N_f is the corresponding number of cycles to failure. The coefficient a is the strength to fracture for one cycle, i.e., it coincides with S_u, and the exponent b is the slope when the function is graphed in a log-log scale. This expression has the advantage that in a log-log graph, it has the form of a straight line, as seen in (2.34). In Section 3 this formulation is used to find an approximate SN curve.

$$S_f = a \cdot N_f^b \tag{2.33}$$

$$\log_{10} S_f = \log_{10} a + b \cdot \log_{10} N \tag{2.34}$$

2.4.4 SN Curve Modifying Factors

The SN curve found from experimental tests is valid for laboratory conditions. If possible, for mechanical design, the curve must be found in the conditions of operation, including the shape of the piece of the equipment. However, that is not generally viable. Hence, the SN curve must be modified to have a closer representation of the actual phenomenon. The common way to modify the curve is to use (2.35) to the high frequency stress limit S_e', usually located at $N = 10^7$ cycles.

$$S_e = k_{\text{sf}} \cdot k_{\text{T}} \cdot k_{\text{sh}} \cdot k_{\text{rel}} \cdot k_{\text{misc}} \cdot S_e' \tag{2.35}$$

The meaning of the factors is as follows: k_{sf} is the finished surface factor, k_{T} represents the temperature, k_{sh} considers the shape and size of the element, k_{rel} is the reliability factor, k_{misc} is the miscellaneous factor. In the ongoing paragraphs,

Table 2.1: Variables for the determination of the surface factor [82].

Surface Finish	α	β
Ground	1.58	-0.085
Machine or cold-drawn	4.51	-0.265
Hot rolled	57.7	-0.718
As forged	272	-0.995

a detail of these factors is given.

2.4.4.1 Surface Condition

When a specimen is tested in laboratory, it has a perfectly polished surface, a characteristic that is known as mirror polished. When the same material of the specimen is used in a piece of a machine, it may not have the same finished surface because the piece may have imperfections [68].

According to Noll et al., the factor of the surface finish could be expressed as (2.36) [81]. The variables α and β, when S_u is in MPa are related to the manufacturing process according to Table 2.1 [82]. For the case of copper used in power transformers, it is generally hot rolled [83].

$$k_{\text{sf}} = \alpha \cdot S_u^{\beta} \tag{2.36}$$

2.4.4.2 Size and shape

The effect of size and shape is evident in bending and torsional stresses. In winding conductors, the stress is located perpendicular to the cross-section, thus, it could be taken as an axial stress [68]. For that reason, the size does not have a direct effect on the SN curve, and k_{sh} could be taken as 1.

2.4.4.3 Type of Loading

The SN curves are typically obtained by test performed in the R.R. Moore machine. These tests apply bending tests. To adapt the results to the different type of loads: bending, torsional or axial, the factor k_{L} is included. For stress acting in the axial direction, the factor is 0.85 [82]. This value is the average of various tests and statistical comparisons between the results.

2.4.4.4 Temperature

If the temperature is too low, a ductile material becomes a little brittle, i.e. it has a higher tendency to be broken, instead of stretched. On the other hand, if the temperature is too high, the yielding point of the metal suffers variations. These effects are included in the SN curve through the factor k_{T}. For 100° the $k_T =$1.02 [82], which could be employed considering a rated load.

2.4.4.5 Reliability

The test for stress-strain and SN curves are not deterministic. They are performed in several samples and the published results are those with a 50% of probability. From a security point of view, this probability is too low because it would mean that the analysis performed have a probability of 50% of being accurate. In a mechanical device, that could be catastrophic.

In order to have more accurate results, a higher probability of analysis can be

Table 2.2: Values of Z_a according to the reliability.

Reliability	Z_a
0.9	1.288
0.99	2.326
0.999	3.091
0.9999	3.719

used. Considering a normal distribution with an 8% [84] of deviation, the factor k_{rel} is found by (2.37). Z_a represents the Z distribution, which depends on the level of reliability desired, as shown in Table 2.2 for values greater than 90%.

$$k_{\text{rel}} = 1 - 0.08 \cdot Z_a \tag{2.37}$$

2.4.4.6 Miscellaneous Effects

The last factor k_{misc} accounts for those effects not included in the other factors [82]. These effects could be

- Static stress
- Electrolyte concentration
- Dissolved oxygen in electrolyte
- Local crevices
- Mechanical frequency

Static stress refers to the existence of a tension or compression that is constantly applied to the body. Note that it is not the mean value of the stress signal. The static stress exists by itself, independent of the external forces. That could be the case, for example, of a compressed spring.

The electrolyte concentration and the dissolved oxygen in the electrolyte are part of liquids or the medium surrounding the body. Finally, the local crevices are holes or gaps that could exist inside the material. None of those phenomena are generally present in the winding conductors.

The only direct effect that could affect the power transformer windings is the mechanical frequency of the stress. The frequency, combined with the temperature surrounding the conductor material, may accelerate the propagation of the internal cracks. A factor $k_{\text{misc}} = 0.5$ is considered as an average effect [73].

2.4.5 Cumulative Damage Evaluation Techniques

For the evaluation of the fatigue damage, the damage index D, for one cycle of a given stress, is defined by (2.38), where N_f is the number of cycles to failure, and n is the number of cycles the stress acts over the body.

$$D = \frac{n}{N_f} \tag{2.38}$$

2.4.5.1 Palmgren-Miner Rule

The most known and used method to perform a fatigue analysis for variable stresses is the Palmgren-Miner rule, also known as the Miner's rule [85]. Say, for example, that there are two levels of stress that have acted over a body, as shown in Figure 2.7. Besides, in the SN curve, the stress levels and the number of cycles to failure are related as shown in Figure 2.8. Thereby, D is calculated by (2.39).

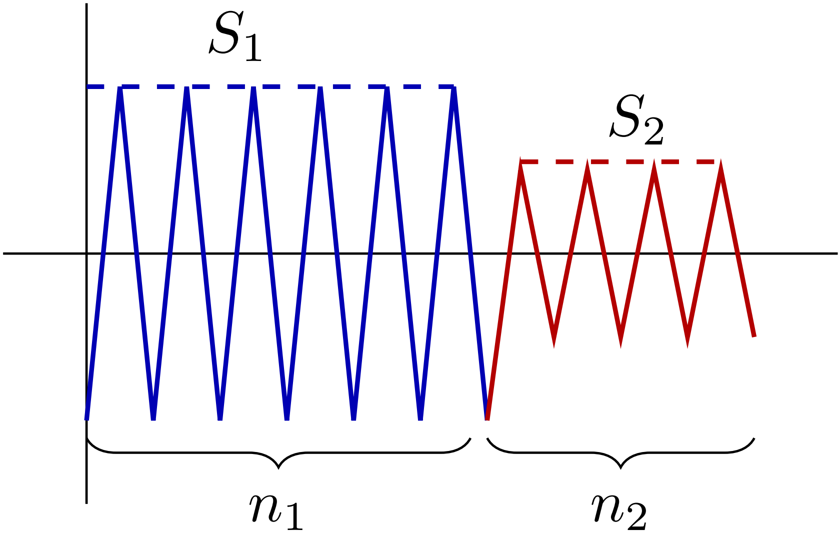


Figure 2.7: Variable stress. S_1 acts during n_1 cycles whereas S_2 during n_2 cycles

$$D = \frac{n_1}{N_1} + \frac{n_2}{N_2} \tag{2.39}$$

Generalizing for multiple stresses, the Miner's rule is expressed as (2.40).

$$D = \sum_i \frac{n_i}{N_i} \tag{2.40}$$

This formulation has two drawbacks [86]: (1) It does not make any distinction in the order of the stresses. (2) In the real phenomenon, the effects of high stresses are different from the effects of low stresses, in regard to the development of cracks and notches; Miner's rule cannot distinguish between those

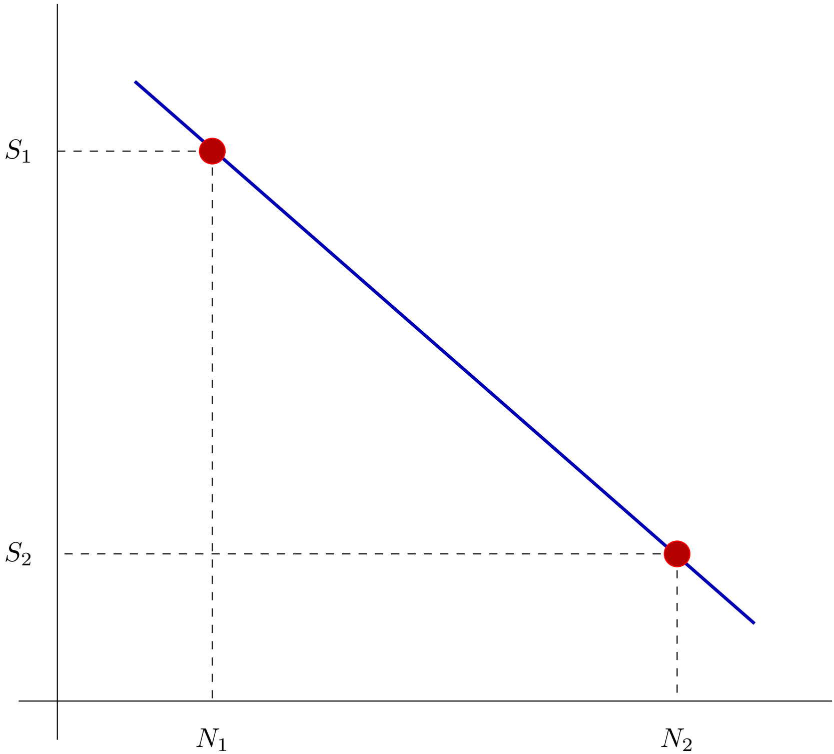


Figure 2.8: SN curve for the hypothetical case. If S_1 is applied constantly, there is failure after N_1 cycles. Similar situation happens if S_2 is applied during N_2 cycles.

stresses.

3 METHODOLOGY

3.1 SCOPE OF THE RESEARCH

It was seen in previous chapters that little research has been done in regard to the effects of the mechanical stress on power transformer windings. The studies have been focused on the effects of stresses due to the energization of power transformers, or in the variation of the stress when the parameters of an electrical fault are changed. The effects of the mechanical stresses in power transformers' lifespan has not been treated. Thus, this research started as exploratory, and as the theory has been developed, the research became descriptive, quantitatively detailing the stresses values and the reduction of the power transformer's life according to the electrical faults in the system.

The research is not performed using actual measurements. Rather, the evaluation is made through simulations of the behavior of electromagnetic forces corresponding to a set of currents circulating through the transformer windings. In addition to that, the currents are also a result of simulated conditions of a power system. Therefore, the research is experimental with controlled conditions.

The tasks that were performed during the development of the research are

- ❒ Evaluation of the reduction in power transformer lifespan when the transformer is working under rated load and when it is affected by power system faults and inrush currents.
- ❒ Comparison of the lifespan of the winding conductors and the lifespan of insulation, under the assumption that the transformer is working with the load condition, and with the reliability indexes given by the IEEE Reliability Test System [87].
- ❒ Comparison of the lifespan of aluminum and copper winding conductors from a fatigue analysis point of view.
- ❒ Comparison of the fatigue analysis results from Miner's rule and the Damage Transfer Concept methods.

In the evaluation of the lifespan considering rated load and mechanical stresses, a variation of the fault conditions, such as magnitude and fault impedance, were included. The mechanical stresses were obtained by using a combination of FEMM and Python, where the geometry of the transformer was given as data. The same tools were used for the comparison of aluminum and copper windings.

In regard to the simulations of faults according to reliability indexes, the load conditions were included in a power flow analysis so that the loading of the transformer varies depending on each week of the year, day, or hour. To automate this process, a Python script was developed. The fault currents were also part of this script so that the current magnitude could be used in the determination of the mechanical stresses. As this process simulates a real time operation, a faster method to calculate the mechanical stresses was developed through the use of random forests.

For the comparison of the Miner's rule and the Damage Transfer Concept, a script was developed in Python, besides the use of FEMM.

The order in which the study was developed is

1. The effect of the winding material was evaluated using the Miner's rule. The material with less lifespan regarding the fatigue effects is chosen for the next steps.

2. The Miner's rule and the DTC were compared for low ultimate strength values and without fault events because the DTC was unstable for the latter.

3. The operation of the power transformer was assessed along the years until damage for fatigue is achieved. The results of the power flow, the random forests and the Miner's rule were used for this step.

With respect to the limitations, this study did not contemplate the following issues:

- Variations of the insulation lifespan due to transformer overloads.

- Variations in the power transformer capacity. Only one transformer of the IEEE Reliability Test was analyzed.

- Variations in the power transformer design. Hence, the dimensions of both aluminum and copper windings are constant during the whole study.

- Only the SN curve obtained with the simplified method was used during the fatigue analysis. Non-experimental curve was employed.

A flowchart that summarizes the procedure of analysis is shown in Figure 3.1.

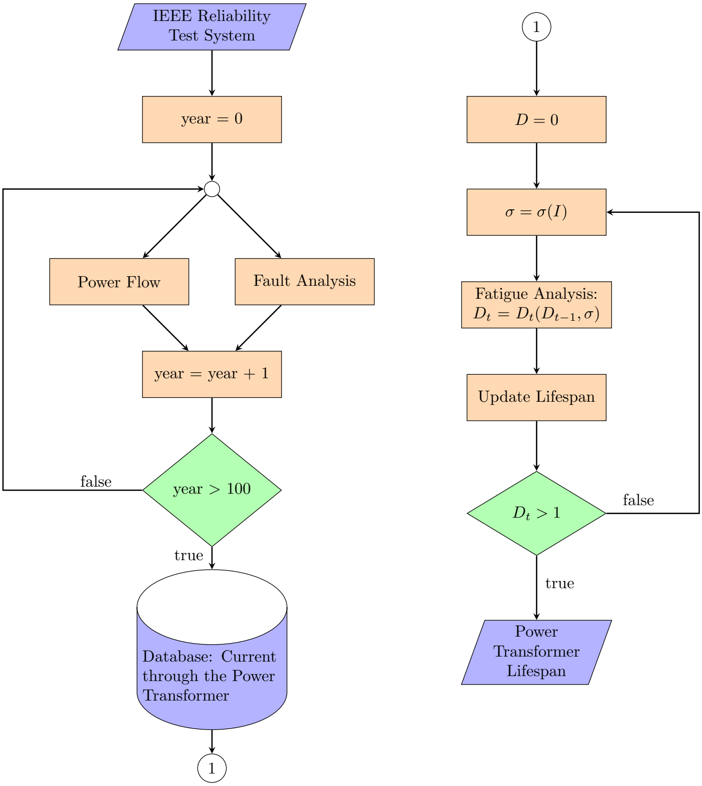


Figure 3.1: Procedure of Analysis

3.2 DESCRIPTION OF THE POWER SYSTEM AND POWER TRANSFORMER

3.2.1 IEEE Reliability Power Test System

In order to evaluate the effect of the mechanical stresses during the transformer lifetime, operating and fault conditions in a power system were simulated. The

IEEE reliability power test system was used for this purpose [87], see Figure 3.2.

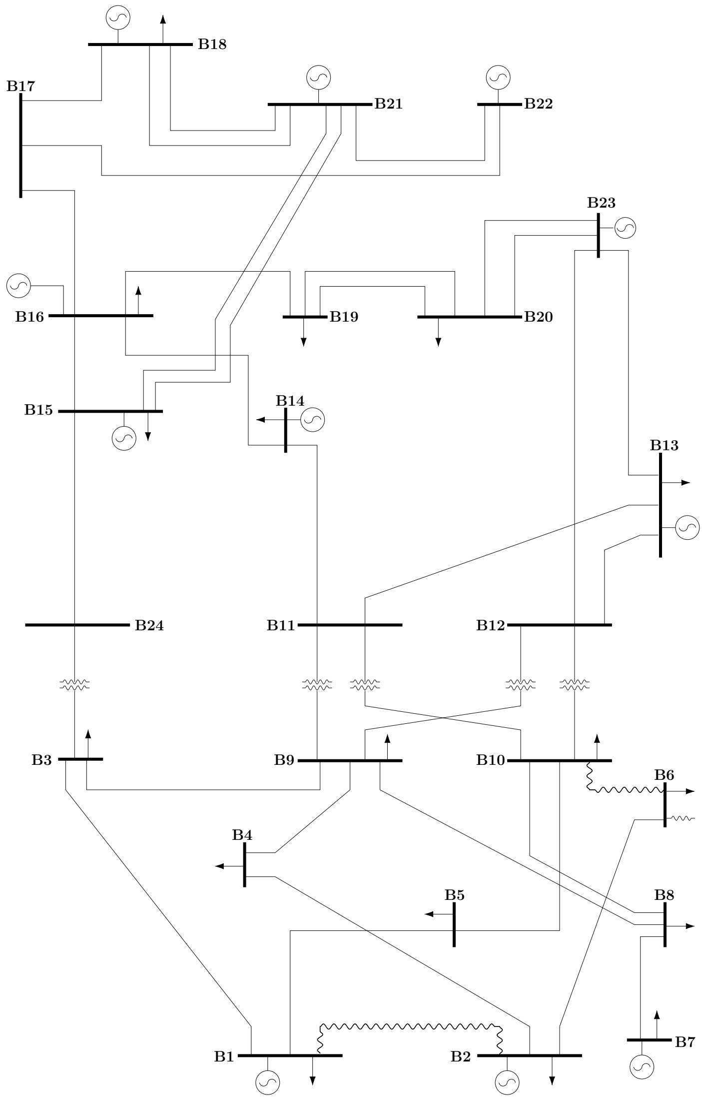


Figure 3.2: IEEE Reliability Power Test System

Among other information, the IEEE reliability power test system has the following data:

Table 3.1: Power Transformer Characteristics

Variable	Value	Unit
Rated Capacity	400	MVA
High Voltage	230	kV
Low Voltage	138	kV
Leakage Impedance	0.0023 + j0.0839	pu (100 MVA base)
Type	Three - Phase	

- ❒ Reliability indexes of the components of the system.
- ❒ Load characteristics in a season, weekly, daily, and hourly basis.
- ❒ Capacity of the generating units.

Therefore, the operating conditions of the power transformer could be simulated under normal and fault conditions.

3.2.2 Power Transformer

From the IEEE Power System, the transformer connected in buses 9 and 11 has been chosen for the analysis. The characteristics of the transformer are shown in Table 3.1.

For the copper windings, a current density of 3 A/mm^2 was set, whereas for the aluminum windings, the current density was 1.5 A/mm^2. These values come from typical power transformer designs [88]. The internal dimensions and characteristics are shown in Tables 3.2, 3.3, and 3.4. In FEMM the disks were considered as one body, so that the forces are calculated as an average of the set of conductors that form the disk. The disks are equally spaced along the winding, and they are symmetrically located in the vertical directions in relation to

Table 3.2: Low Voltage Winding Characteristics

Variable	Copper Winding	Aluminum Winding	Unit
Number of Disks	100	110	u
Number of Turns	300	300	u
Disk Height	17	24	mm
Disk Width	100.5	142.5	mm
Inside diameter	1108	1323	mm
Winding Height	2740	3785	mm

Table 3.3: High Voltage Winding Characteristics

Variable	Copper Winding	Aluminum Winding	Unit
Number of Disks	105	126	u
Number of Turns	500	500	u
Disk Height	13	19	mm
Disk Width	133	183	mm
Inside diameter	1429	1728	mm
Winding Height	2468	3719	mm

the core window. The designs were made so that both windings have about the same height.

Table 3.4: Core Characteristics

Variable	Copper Winding	Aluminum Winding	Unit
HV-HV separation	158	158	u
Limb-Limb Distance	1853	2252	u
Limb Diameter	958	1173	mm
Inner Window Height	3034	4079	mm

3.3 ELECTRICAL CURRENT UNDER NORMAL AND FAULT CONDITIONS

One of the goals is to simulate the power transformer operation working conditions during its lifetime. Hence, periods of normal load and fault events must exist. For the normal conditions, the IEEE test system is analyzed with power flow simulations. A fault analysis is performed to get the conditions under fault events.

3.3.1 Power Flow Analysis

The IEEE Reliability Test System gives the load variations for each of the 52 weeks, for each day of the week, and for each hour of the day. The weeks are organized according to the four seasons of one year. The load is considered constant during each hour. In addition, workday loads are different from weekend loads.

The Test System also gives the characteristics of the generation for each bus. In some buses, the generation system is divided in units so that individual generator faults can be studied.

For a year with 365 days, there are 8760 hours. The same number of simulations is needed to get the power transformer operation under normal conditions.

In order to simplify the analysis, the generation system on each bus is taken as a whole unit. Bus 1 was considered as the slack bus. The generator in Bus 1 does not have the highest capacity, but it has been chosen as the slack bus for facility of programming. Buses with generation were classified as PV buses and

Table 3.5: Load Determination for the First Week, First Day and First Hour

Variable	Value	Unit
Total Base Load	2850	MW
Week Factor	86.2	%
Day Factor (Monday)	93	%
Hour Factor (00:00 - 01:00)	67	%
Final Load	1530	MW

Table 3.6: Load Sharing for each of the Generation Units

Bus	Capacity (MW)	Percentage (%)	Load Shared MW
1	192	5.64	86.27
2	192	5.64	86.27
7	300	8.81	134.80
13	591	17.36	265.56
15	215	6.31	96.61
16	155	4.55	69.65
18	400	11.75	179.74
21	400	11.75	179.74
22	300	8.81	134.80
23	660	19.38	296.56

the rest were PQ buses.

The generation on each PV bus was determined in proportion to the individual capacity and according to the total demand. In order to clarify how the generation was shared, the procedure for the first week, first day and first hour is detailed. In Table 3.5 the Load is calculated for the given week, day, and hour. The final load results from the multiplication of the base load by the three factors.

The final load is shared among the generation plants according to its maximum capacity, as shown in Table 3.6.

In the procedure, a PV bus remains as that even though the reactive power is surpassed. This means that the reactive limits are not checked. The same situation happens with the lines, whose limits are not evaluated on each simulation. The reason for this criterion is that the variations in the power transformer loading will not greatly affect the fatigue analysis, more so if the magnitude of the fault currents is compared with that of normal operating conditions.

For the power flow analysis, a script was developed in Python. The Gauss-Seidel method was used for the solution of the power flow equations. The tolerance for the iterations difference was 10^{-5} and a maximum of 200 iterations were considered. The magnitude of the power flow through the transformer connected between buses 9 and 11 for each hour was the only datum stored.

3.3.2 Fault Currents

Only three-phase fault currents were calculated, mainly due to two reasons:

- The lack of data in the IEEE Test System for the negative and zero sequence models.
- The conservative approach of the analysis since a three phase fault current causes a higher effect on the derating of the winding conductor.

The fault current was determined by using the impedance bus matrix Z_{bus}. The original Z_{bus}, i.e. for a normal operating system, was calculated through the inversion of the admittance bus matrix Y_{bus} that was used in the power flow analysis, but modified by the inclusion of the subtransient impedance of the generators.

Only faults on transmission lines or their correspondent terminal buses were considered in the procedure. The currents that result from faults on buses are calculated directly from the original Z_{bus}. The current $I''_{9,11}$ through the transformer when Bus k is faulted, was calculated by (3.1) [89]. V_f is the prefault voltage, considered as 1 pu. for simplification of the analysis. Z_t is the leakage impedance of the transformer. Z_{9k}, Z_{11k}, and Z_{kk} are the 9th, 11th and kth row elements of the kth column, under the assumption that the fault occurred on Bus k.

$$I''_{9,11} = -\frac{V_f}{Z_t}\left(\frac{Z_{9k} - Z_{11k}}{Z_{kk}}\right) \tag{3.1}$$

Equation (3.1) gives the steady RMS value of the fault current. However, for fatigue analysis, the transient period is also required. For the simulation of the transient period, the resistance and reactance characteristics of Z_{kk} were assumed.

When the fault occurs inside a line located between buses i and j, a new bus must be created in the place where the fault is assumed. This activity requires three steps:

1. Eliminate the line from Z_{bus} by adding an impedance $-Z_b$ between Buses i and j, where Z_b is the impedance of the line to be eliminated.

2. Add an impedance $\alpha \cdot Zb$ from Bus i to a new Bus k, where α is the fraction of the length of the line, representing the location of the fault measured from Bus i.

3. Add an impedance $(1 - \alpha) \cdot Zb$ between Bus j and the new Bus k.

After the new bus has been created, the procedure is the same as before and the current through the transformer can also be calculated by (3.1).

3.3.3 Generation of Faults in the System

Faults are generated only on transmission lines. The frequency of faults depends on the outage rate (faults per year) given by the IEEE Reliability Tests System [87].

For the generation of faults, a Monte Carlo simulation was implemented. Two values were required: (1) The time when each line is expected to have a fault, and (2) the location of the fault.

For the time of fault occurrence T_i, the exponential function of probability was assumed, so that the time could be obtained by (3.2) [90], where U_i is a random number, and λ_i is the outage rate. T_i is given in years.

$$T_i = -\frac{1}{\lambda_i} \ln U_i \tag{3.2}$$

A second random number is used to determine the fault location. If the line goes from Bus i to Bus j, a random number of 0 means a fault at Bus i, and a random number of 1 means a fault at Bus j. Random numbers between 0 and 1 mean faults located between both buses.

3.4 DETERMINATION OF MECHANICAL STRESSES

There are two different types of applications when determining mechanical stresses. In the first type, the mechanical stresses effects are analyzed individually, as in an off-line situation; in this case, the force produced by a single value of current is calculated. In the second type, the same effects are analyzed as if the forces are acting continually in the transformer windings, similar to an on-line analysis, so the cummulative fatigue analysis is performed. Both types of analysis have the electrical current circulating through the windings as the input data, and the mechanical stress as the output data.

The electrical current has two distinct stages, the transient and the steady stage. The transient stage represents the time from the first appearance of the fault to its clearing through the protection system. The currents during this stage are modeled by (3.3). I_{LV} is the RMS current once the fault is in steady state, ω is the angular frequency of the system, t is the time, ϕ is the angle representing the fault starting point, θ is the angle between phases (120° in a three-phase balanced system), $\lambda = \omega r / x_l$, where r and x_l are the equivalent resistance and inductive reactance seen by the fault. The high voltage current is determined with the ratio of the respective windings because the fault current is by far higher than the excitation current, and so, the latter can be neglected.

$$\begin{aligned}
i_a(t) &= I_{LV} \cdot \sqrt{2} \cdot [\sin(\omega t - \phi) + \sin(\phi) \cdot \exp(-\lambda t)] \\
i_b(t) &= I_{LV} \cdot \sqrt{2} \cdot [\sin(\omega t - \theta - \phi) + \sin(\theta + \phi) \cdot \exp(-\lambda t)] \\
i_c(t) &= I_{LV} \cdot \sqrt{2} \cdot [\sin(\omega t + \theta - \phi) + \sin(-\theta + \phi) \cdot \exp(-\lambda t)]
\end{aligned} \tag{3.3}$$

For the duration of the transient stage, the operation of the protection was considered uniformly for all the faults with a backup time of 0.5 s as recommended by Hewitson et al. [91].

For all the fatigue analyses, only the peaks of each period are used in the calculation of the mechanical stresses because in that way the stresses can be considered as those pulses represented in SN curves.

3.4.1 Mechanical Stresses Using FEM

The FEM analysis of mechanical stresses and fatigue was applied when comparing the effects of fatigue in copper and aluminum.

For this case, the software FEMM (Finite element method magnetics Version 4.2) was used [92]. FEMM allows the simulation of magnetic fields when design data is available, such as number of turns per disk, electrical current through the windings, internal geometry of the transformer, etc. As a result of the simulations, the magnetic induction is determined on each point of the medium. Finally, FEMM has an algorithm to determine the average force per volume on each region formed during the construction of the internal geometry. In this research, each disk was constructed as an independent region.

FEMM does not simulate time variable fields. Nonetheless, it can be used for power frequency problems because the displacement current can be neglected in those cases. Therefore, for each electrical current value, a different simulation had to be performed. In order to automate the process, the library Pyfemm was used. This library allows the interaction between FEMM and Python with all its practical features.

An example of the Phase A transient current along with the resultant mechanical stress is shown in Figure 3.3 and Figure 3.4. Note how the mechanical stress has a frequency that doubles that of the electrical current.

3.4.2 Mechanical Stresses Using Random Forests

The random forests type of analysis is intended to simulate the operation of a power transformer in actual conditions. Hence, the operation includes variations

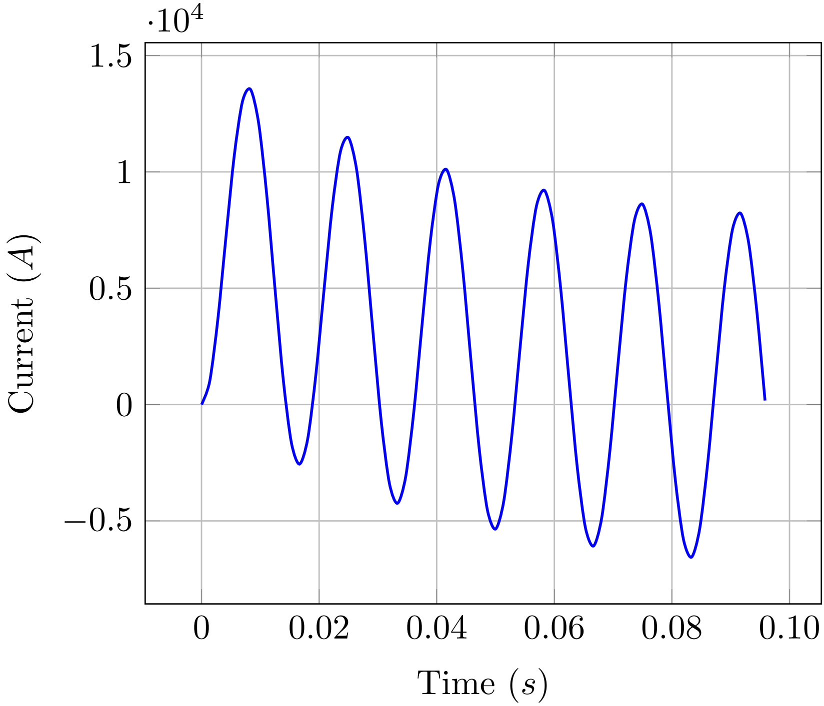


Figure 3.3: Transient current in Phase A.

of load and the existence of events, such as three-phase faults on the lines.

The simulations for random forests analysis are performed from the first energization to the end of life of the transformer. The number of simulations required is huge and hence the speed of solution becomes important.

The finite element method that is generally used for this type of problems has the disadvantage of needing a considerable computational time to get one solution. The necessity of handling matrices, and the need of iterations makes it not suitable when the speed of solution is essential.

For that reason, a new faster method was to be found. In this research, numer-

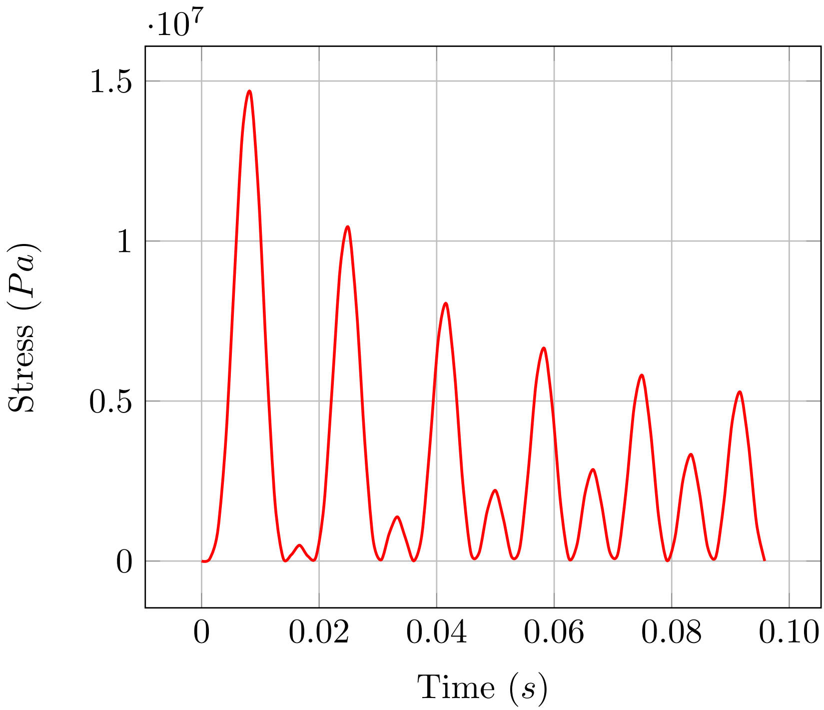


Figure 3.4: Mechanical stress correspondent to the transient current in Phase A.

ical experiments were made to find a faster method among machine learning tools.

It has been demonstrated that when comparing four machine learning techniques, namely linear regression, random forests, support vector machines, and deep learning, random forests presented the most accurate results [93], as can be seen in Figure 3.5.

Random Forests is a supervised learning tool, so it needs a set of input and output data from which to learn. For the mechanical stress problem, the input data is formed by the electrical currents and the output data by the mechanical stress. This set is known as the learning set.

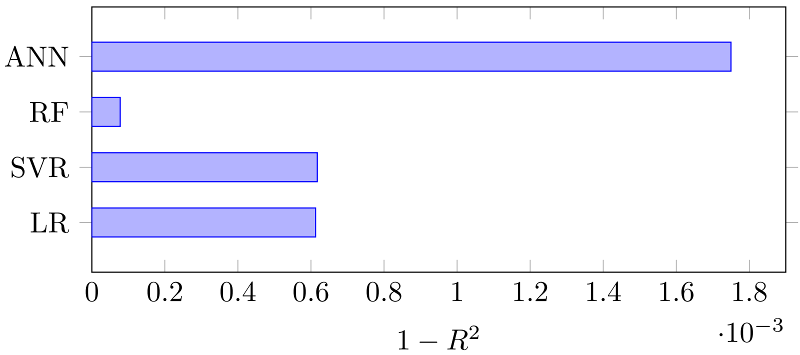


Figure 3.5: Comparison of the error for four machine learning techniques: Artificial neural network, Random forests, Support vector regression, Linear Regresion. Random forests has the least error.

In this research, the learning set was obtained from FEM simulations. The number of samples, i.e. the number of FEM simulations, were 8634. The simulations involved fault impedances with variations in the resistive and reactive component. The impedances varied from $1 + j15\Omega$ to $5 + j80\Omega$. They represent currents from fault in the transformer terminals to working loads at half the rated value. In addition to that, the simulations also alternatively considered cases where the fault initiated in Phases A, B, and C.

The learning set was divided into training and validation data set, with 7869 and 795 samples respectively. The validation data set was used to check for the existence of overfitting, which could appear as a result of using too many trees inside the Random Forest. The results of this checking are shown in Figure 3.6. The minimum error value was obtained at 100 trees. Note that after that value, the validation test error is somehow stabilized and has a small increment which could be an indication of a possible overfitting.

In order to avoid that the highest mechanical stresses have a stronger influence

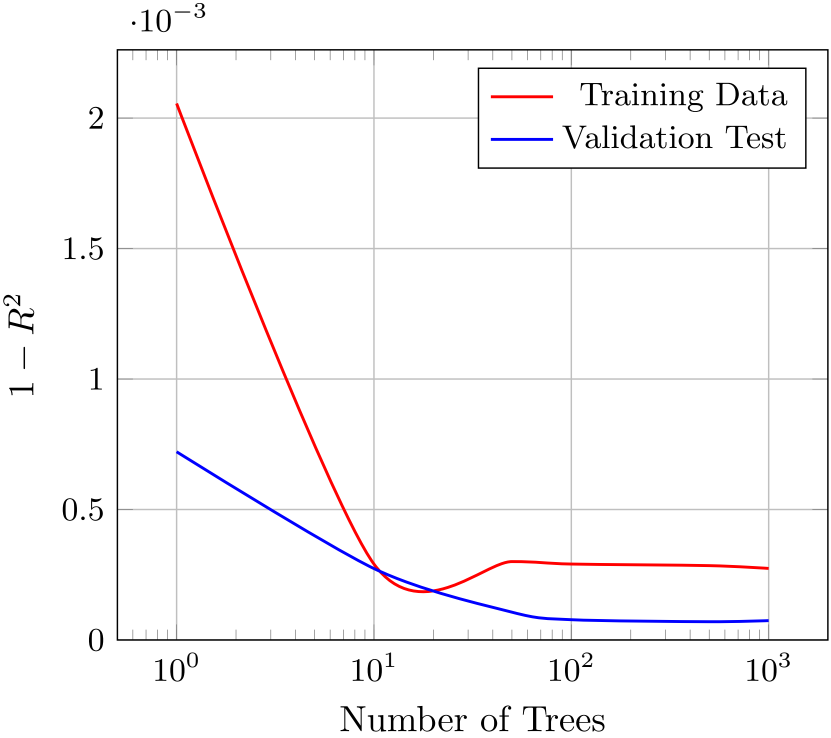


Figure 3.6: Error behavior in regards to the number of trees. After 100 trees, the model does not present any improvement.

during training the model, the data were previously normalized, so that the values are in the range 0 to 1. This standardization is performed by (3.4)

$$
\begin{aligned}
x_{\text{std}} &= \frac{x - x_{\min}}{x_{\max} - x_{\min}} \\
y_{\text{std}} &= \frac{y - y_{\min}}{y_{\max} - y_{\min}}
\end{aligned}
\tag{3.4}
$$

The machine learning process was made using the scikit-learning library of Python [94]. Its algorithm optimizes the model by minimizing the error represented by the coefficient of determination R^2, defined by (3.5), where y_{true} is

Table 3.7: Modifying factors for Aluminum and Copper

Factor	Value
k_{sf}	1.162
k_{T}	1.02
k_{sh}	1
k_{rel}	0.689
k_{L}	0.85
k_{misc}	0.5

the true value of the output variable, y_{pred} is the output value predicted by the model, and $\bar{y}_{true}$ is the mean value of the output variable.

$$\begin{aligned} R^2 &= 1 - \frac{u}{v} \\ u &= \sum \left(y_{true} - y_{pred}\right)^2 \\ v &= \sum \left(y_{true} - \bar{y}_{true}\right)^2 \end{aligned} \tag{3.5}$$

3.4.3 SN Curves and Fatigue Analysis

The SN curves are determined using the following criteria [73]:

- At $N = 10^3$ cycles, set a stress $S = 0.9 \cdot S_u$.
- At $N = 10^7$ cycles, set a stress $S = 0.3 \cdot S_u$.
- Apply the modifying factors to the stress at $N = 10^7$.

In that way, two points are given so that the expression for a straight line (2.34) can be obtained. The modifying factors are common for both aluminum and copper and are shown in Table 3.7. For the surface factor k_{sf}, an ultimate stress of 90 MPa has been considered, which is the minimum available from factories [38].

The SN curves are the same, independent of the material of the winding conductors. Hence, the comparison is limited to the geometrical and field characteristics of the transformer.

Therefore, there are two points to get the line: $[10^3, 81]$ and $[10^7, 27]$. In that way, two equations with two unknowns are obtained, as shown in (3.6).

$$\begin{aligned} \log_{10}(81) &= \log_{10} a + b \cdot \log_{10}(10^3) \\ \log_{10}(27) &= \log_{10} a + b \cdot \log_{10}(10^7) \end{aligned} \tag{3.6}$$

From the solution of both equations, the SN curve is given by (3.7).

$$S_f = 153.55 \cdot N^{-0.1078} \tag{3.7}$$

The modifying factors are applied to the stress value at $N = 10^7$. So, the second point was $[10^7, 9.37]$. Therefore, after solving the system of equations, the modified SN curve is given by (3.8).

$$S_f = 367.5 \cdot N^{-0.234177} \tag{3.8}$$

Both, the theoretical SN curve as well as the modified curve are shown in Figure 3.7.

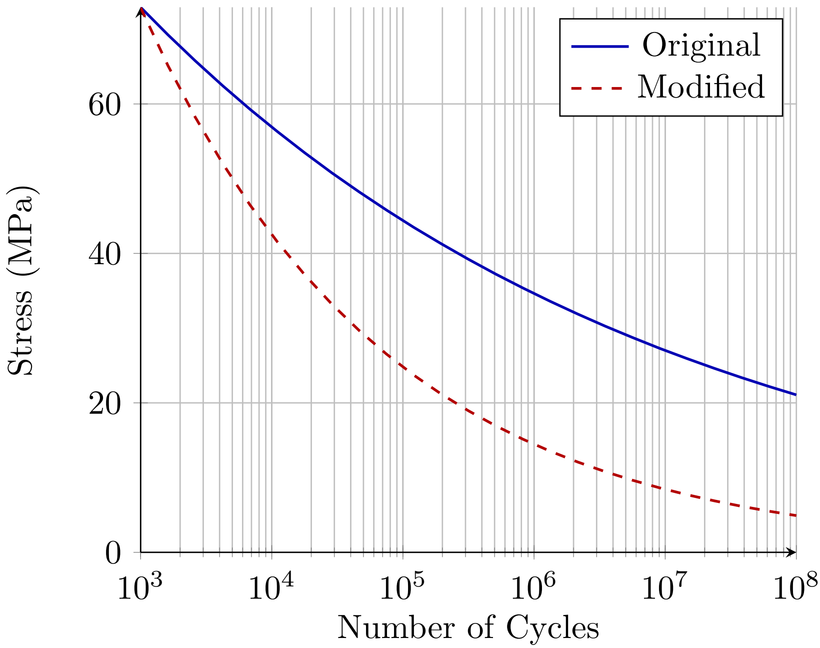


Figure 3.7: SN original and modified curves for the winding conductor.

4 RESULTS AND DISCUSSION

In this chapter, the fatigue analysis results for the 400 MVA power transformer are presented. This analysis was performed in three kinds of problems:

- ❒ To compare the effect of the mechanical stress in copper and aluminum windings.
- ❒ To compare the fatigue analysis results when using Miner's rule and the Damage Transfer Concept.
- ❒ To evaluate the behavior of the fatigue strength of a power transformer as part of a power system.

4.1 DESIGN AND DEVELOPMENT OF THE RANDOM FORESTS MODEL

The aim of the random forests model is to obtain the mechanical stresses in the middle disks of each low voltage and high voltage windings as a function of the electrical currents circulating through the windings. The number of forests in the final design was 100. So, a training error of 0.0002798 was obtained. The validation error was $9.5008 \cdot 10^{-5}$.

The mean absolute percentage error (MAPE) for three cases of impedance: $1 + j15\ \Omega$, $1 + j47\ \Omega$, $1 + j80\ \Omega$, is shown in Figures 4.1, 4.2, and 4.3. The lowest impedance has the lowest error. In the low impedance case, the error is about the same for the high and low voltage windings. In the high impedance case, the low voltage windings have less error than the high voltage ones.

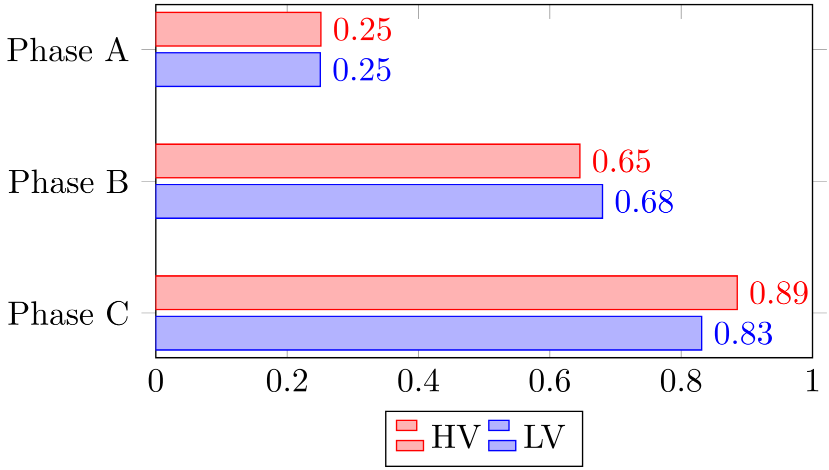


Figure 4.1: Mean absolute percentage error for a transient fault with impedance $1 + j15$ Ω.

Figure 4.4 shows the location of the errors in regard to the value of the mechanical stress. The highest errors are located in low values of mechanical stress. Nonetheless, most of the errors are lower mainly because the stress is usually high.

Figure 4.5 shows the distribution of the error. It is clear that 50% of the errors are located between near 0 to 2.8% approximately. Moreover, 99% of the errors are contained between 0 to near 8%.

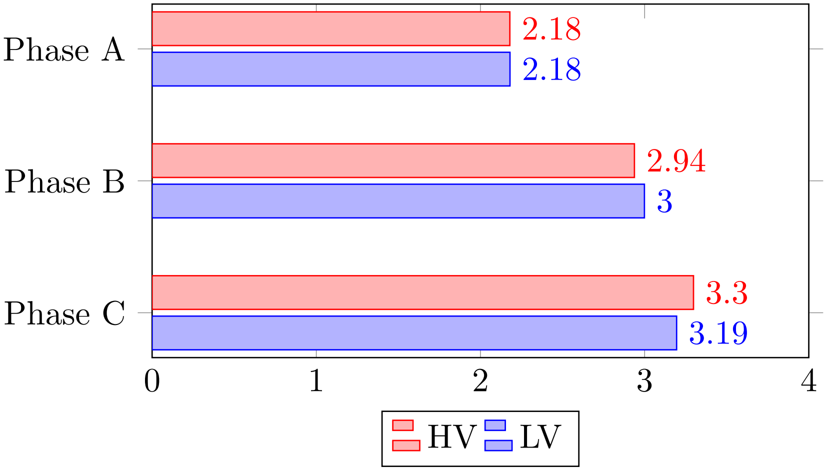


Figure 4.2: Mean absolute percentage error for a transient fault with impedance $1 + j47$ Ω.

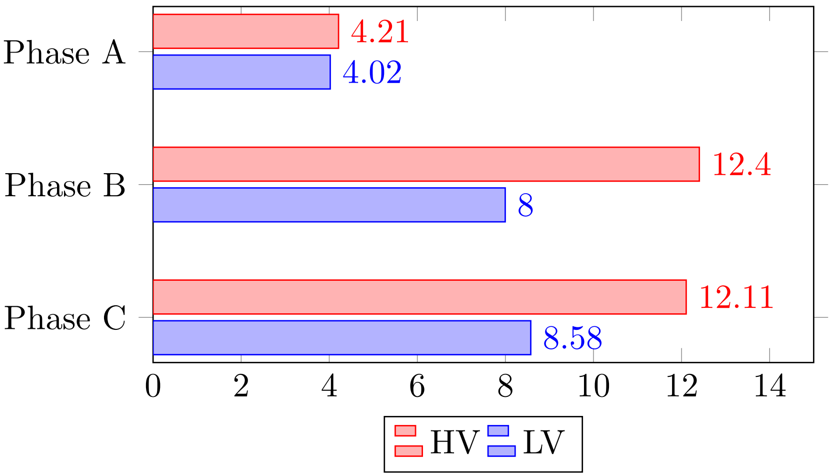


Figure 4.3: Mean absolute percentage error for a transient fault with impedance $1 + j80$ Ω.

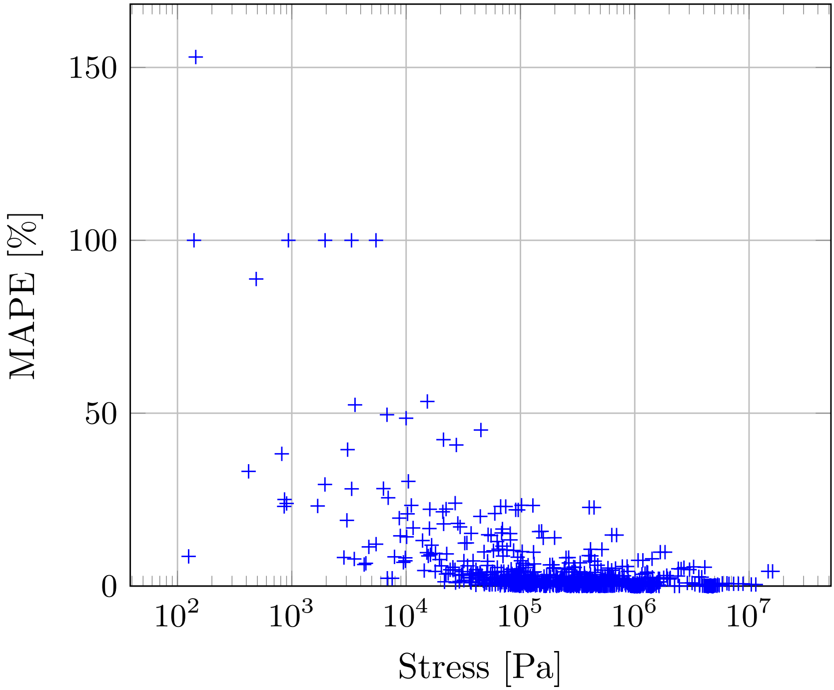


Figure 4.4: Behavior of the error according to the value of the mechanical stress. The highest error is located in values of low stress, i.e. the less serious for the fatigue cummulative effects in the winding conductors.

4.2 INRUSH CURRENT

The magnetization curve used to find the inrush current is shown in Figure 4.6. This curve was obtained using FEM simulations, by injecting the excitation current to the high voltage side of the power transformer, and registering the magnetic induction of the limb of Phase A. Phase A is chosen because the excitation current was at its peak at the time of simulation.

The inrush current when the transformer is energized from the high voltage side is shown in Figure 4.7. The voltage on Phase A is crossing zero at the instant of

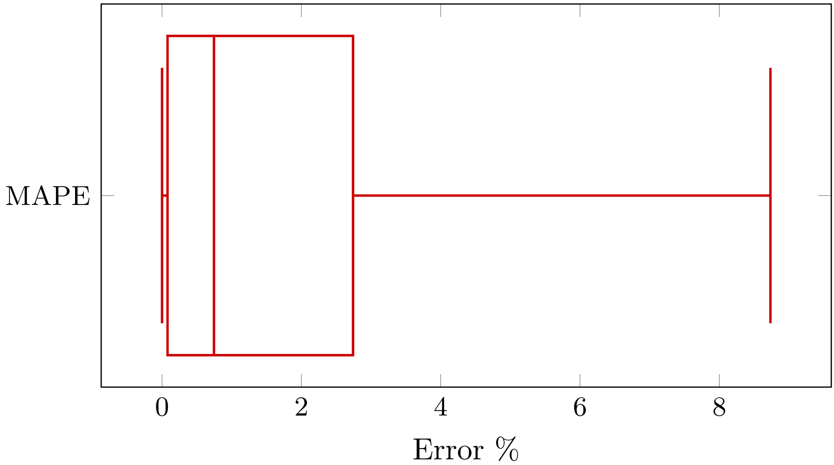


Figure 4.5: Distribution of the errors for the whole set of validation values. Most of the erros are less than 3%

energization, i.e. it is the worst case for Phase A from the magnetization point of view. Non-remnant magnetic field has been considered in the simulation. Only the peaks of the inrush current affect the fatigue behavior of the conductor.

4.3 SHORT CIRCUIT CHARACTERISTICS OF THE 400 MVA POWER TRANSFORMER

The winding conductor is expected to withstand the mechanical stress due to the maximum fault current. Table 4.1 shows the results of the procedure to calculate the maximum stress. The values are calculated assuming that Phase A is under the maximum short circuit current, when the fault is limited only by the transformer impedance and the source is an infinite bus. These values were entered in FEMM and the maximum stress was determined.

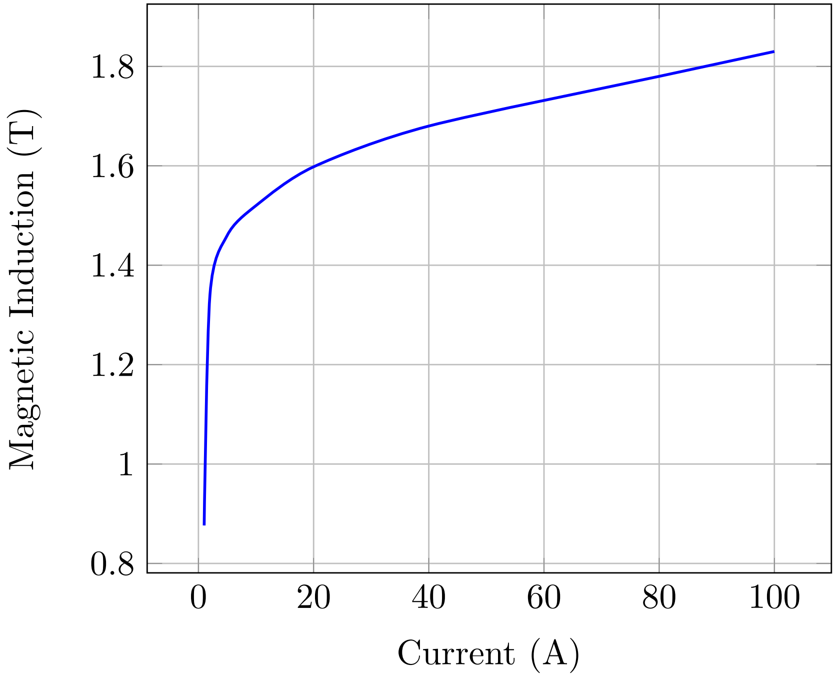


Figure 4.6: Magnetization curve for the 400 MVA power transformer, seen from the high voltage side.

Ideally, a material with a strength slightly higher than the obtained maximum stress should be used, say e.g. a material with an ultimate strength of 20 MPa. Usually, the minimum value of strength utilized in copper is 90 MPa [38].

4.4 COMPARISON OF THE FATIGUE STRENGTH OF ALUMINUM AND COPPER WINDING CONDUCTORS

Faults on the low voltage side of the power transformer were simulated. Table 4.2 shows the fault impedances considered. The transient period lasts 0.5

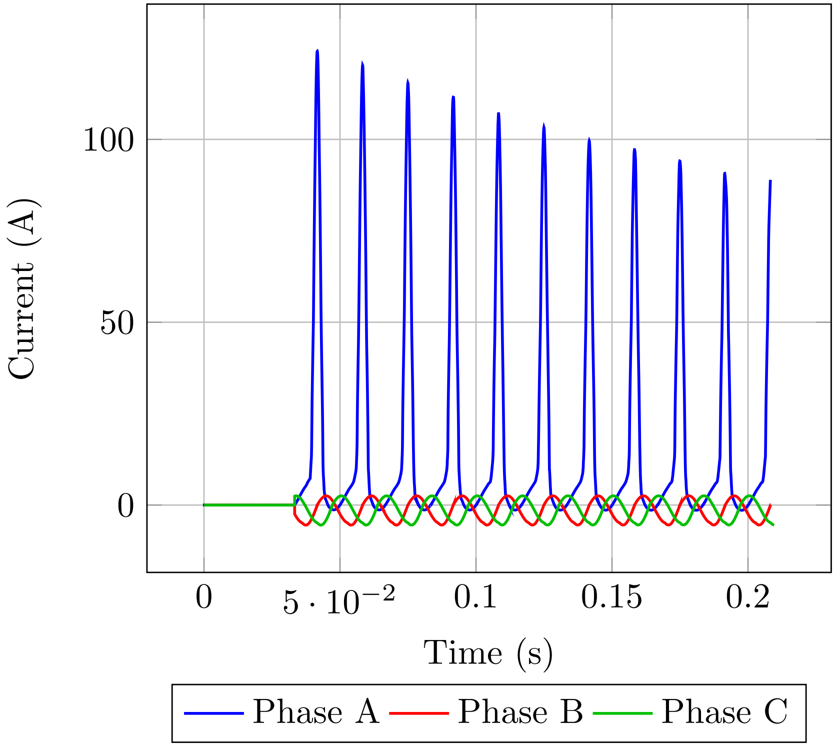


Figure 4.7: Inrush current of the 400 MVA power transformer, energized when the voltage on Phase A is crossing zero.

s, i.e. a time for backup protection settings [91].

In the two-dimensional models used in the simulations, each disk has two values of mechanical stress, one for the right side and one for the left side. One disk per phase was modeled, thus there are six disks in total, three for the low voltage winding and three for the high voltage side. Therefore, considering the nine different fault currents (one for each impedance), there are a total of 108 simulations.

For the comparison of aluminum versus copper, it was calculated how many

Table 4.1: Determination of the maximum stress in the 400 MVA Power Transformer

Item	Value	Unit
Leakage Impedance	0.335726	pu (own base)
Leakage Impedance - LV side	15.9839	Ω
Short Circuit Current at LV side (RMS value)	8.63369	kA
Short Circuit Current at LV side (Peak value)	12.2098	kA
Short Circuit Currents at Phases B and C	-6.10490	kA
Maximum mechanical stress Cu	15.4222	MPa
Maximum mechanical stress Al	4.3226	MPa

Table 4.2: Fault impedances for the simulation of short circuit events at the power transformer terminals

Case	Resistance (Ω)	Reactance (Ω)
1	1	15
2	1	30
3	1	47
4	1	80
5	2	15
6	2	30
7	3	15
8	3	47
9	5	60

Table 4.3: Years to fault for a constant exposition to a fault with an impedance $Z = 1 + 15j$ Ω. Each cycle lasts 0.5 s.

Disk	Aluminum (years)	Copper (days)
Phase A LV left	28.99	50.53
Phase A LV right	30.75	45.23
Phase B LV left	18.49	36.05
Phase B LV right	19.26	36.07
Phase C LV left	31.00	45.33
Phase C LV right	28.83	52.61
Phase A HV left	18.32	32.84
Phase A HV right	12.78	22.82
Phase B HV left	8.42	19.59
Phase B HV right	8.01	18.33
Phase C HV left	12.84	24.41
Phase C HV right	10.12	33.12

mechanical stress cycles of the fault transient (0.5 s) would cause the failure of the conductor, i.e. how many cycles reach the value of $D = 1$ in the fatigue analysis.

As an example, Table 4.3 shows the results for faults with the impedance of Case 1. All the results for copper windings are given in days, whereas for aluminum windings, the results are given in years.

For the low voltage disk, the mechanical stress effect on the left side presents the results shown in Table 4.4. The aluminum conductor windings show some cases that may be considered as of infinite life, mainly when the reactance is greater than 40 Ω.

The probability distribution of the ratio between the lifespan of aluminum and copper conductors is shown in Figure 4.8. The median of the ratios is 110, and the first and third quartiles are 90 and 144, which represents 50% of the cases. Finally, the lower and upper limits, 99% of the cases, are 8 and 226 respectively. Therefore, it can be concluded that the aluminum conductors last longer than

Table 4.4: Lifespan years of the low voltage phase A disk for a constant exposition to a fault in regards to the fault impedance

Disk	Aluminum (years)	Copper (years)
Case 1	14.49	0.1384
Case 2	7080	52.67
Case 3	369416	2457
Case 4	32129950	217870
Case 5	13.15	0.1344
Case 6	6830	52.15
Case 7	12.19	0.1317
Case 8	353753	2440
Case 9	2824150	19632

copper conductors when the ultimate stress is the same for both materials.

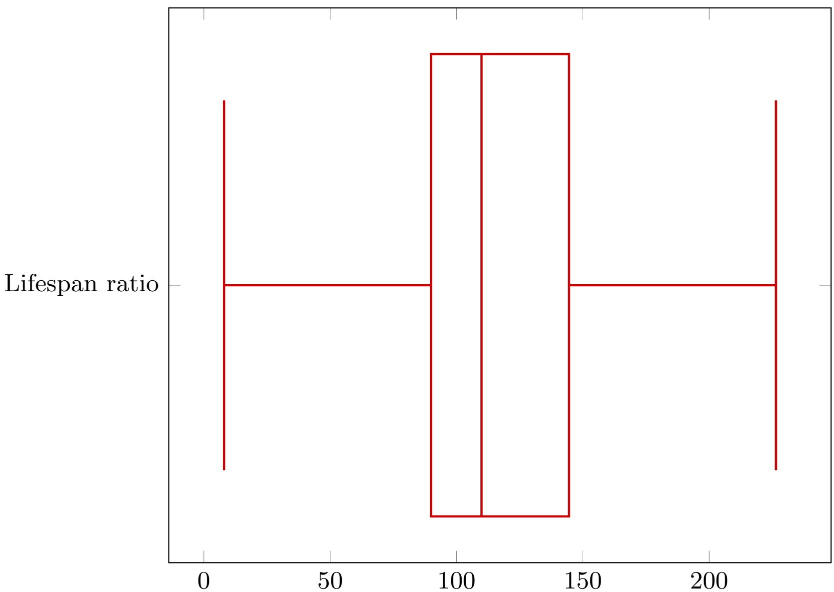


Figure 4.8: Distribution of the ratios between Al and Cu lifespans.

4.5 ANALYSIS OF THE LIFESPAN OF THE 400 MVA POWER TRANSFORMER WINDING CONDUCTORS

Algorithm 4.1 presents the general steps to find the number of years previous to transformer fault. There are three different conditions when updating the index D. In the first condition, the algorithm checks if the year has just started. At the beginning of each year a maintenance has been considered, thus mechanical stresses due to inrush currents are the input data to update D. The second condition is the existence of a fault in one of the lines of the power system. If a fault is found, the transient mechanical stress is the input data. The third kind of updating is the calculation of D considering normal loading conditions in the power transformer.

In this research, the inrush current is the same for every maintenance event and so is the correspondent mechanical stress. This mechanical stress is calculated previous to the determination of D and is valid for the whole fatigue analysis. Due to this invariable characteristics, the mechanical stress coming from inrush currents is calculated directly with FEM. Note that the random forests model cannot be used for this type of event because the model was not trained for zero currents in one of the windings.

The fault events have been determined by Monte Carlo simulations. If in the hour under analysis a line presents a fault, the transient mechanical stress is calculated and then the index D is updated. This event is similar to what would happen in a real time situation. For that reason, the random forests model is used instead of FEM.

Finally, independently of the existence of faults or being at the beginning of the

Algorithm 4.1: Procedure to find the lifespan of winding conductors

```
Data: Mechanical stress from normal, fault, and energizing conditions
Result: Lifespan of winding conductors in years
while Index D < 1 do
    foreach hour in a year do
        if hour = 0 then
            Calculate D with inrush mechanical stress;
            /* Mechanical stress calculated with FEM.            */
        end
        foreach Line in the power system do
            if Fault at line then
                Calculate D with transient fault mechanical stress;
                /* Mechanical stress calculated with random forests
                   model.                                        */
            end
        end
        Calculate D with normal operating load conditions;
        /* Mechanical stress calculated with random forests
           model.                                                */
    end
    number of years ← number of years + 1;
end
return number of years
```

year, the next step is to find the index D under normal operating conditions. For the loading data, power flow analyses were performed. Similar to fault events, the analysis in normal conditions may represent a real time problem, so the random forests model must be utilized.

The fatigue analysis was performed on Phase A because this is the phase where the highest currents are presented both during energization and during fault events.

4.5.1 Lifespan analysis without considering fault events

In the following analyses, a growth rate of 3% per year has been assumed for both load and fault current. The percentage is based on the original load and the maximum capacity of the transformer, so that the latter is reached in a 40-year time frame.

First of all, a comparison between the DTC and Miner's rule results is shown in Figure 4.9, where faults have not been considered. The DTC model is more pessimistic, all the results are lower than those of Miner's rule mainly for low values of ultimate strength.

Figure 4.10 shows how the index D rises along with the lifespan spent for the fatigue analysis. Also, the proportion of lifespan spent regarding the load, calculated using the thermal model, is included. This last curve could be used as a basis to compare the behavior. Note that DTC gives an optimistic state of the conductor, since it gives a lower value for D the whole time. On the other hand, the Miner's rule gives lower values until the life has been spent on a factor of 0.6. After that, the Miner's index D is more pessimistic in the sense that it presents a higher value than that of the thermal analysis.

4.5.2 Lifespan analysis considering fault events

4.5.2.1 Lifespan according to the load characteristics

One of the objectives of this research is to compare the fatigue analysis results with those of the lifespan determined using the load characteristics in the

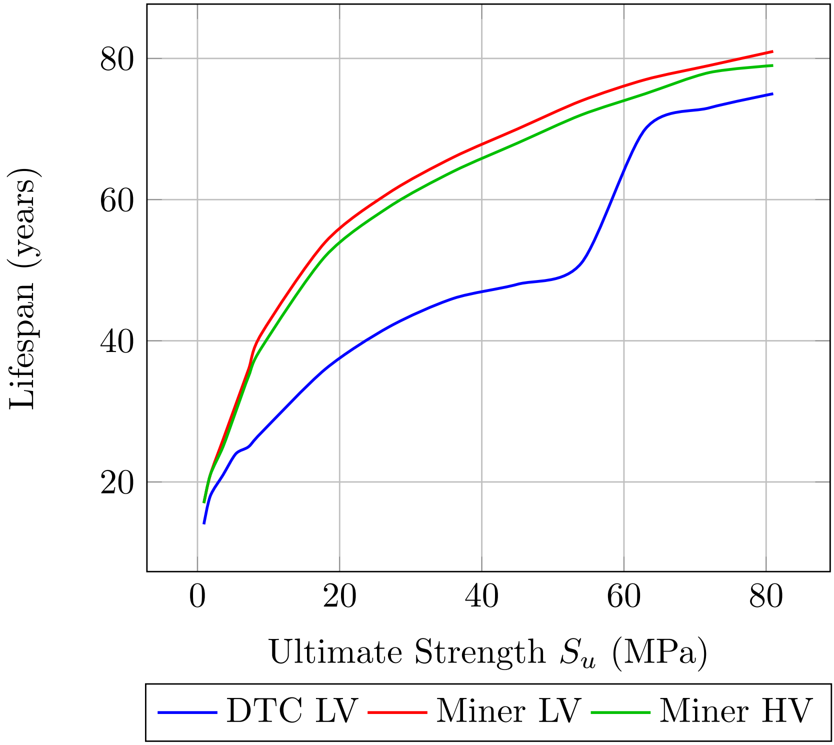


Figure 4.9: Comparison between the lifespan calculated with DTC and Miner's rule.

thermal model. The procedure followed was that indicated by IEEE Std C57.91-2011 [24], which follows a thermal model where the transformer suffers a higher aging when the top oil temperature surpasses the 110° C that is considered under rated operating conditions. According to the standard, if the top oil temperature is always 110° C, the lifespan of the transformer is 180000 hours.

For instance, Figure 4.11 shows the load characteristics of the transformer for the first day of the year, and Figure 4.12, the temperature characteristics for the same day. Note how the temperature closely follows the behavior of the load, a little behind because of the oil temperature time constant that introduces a small

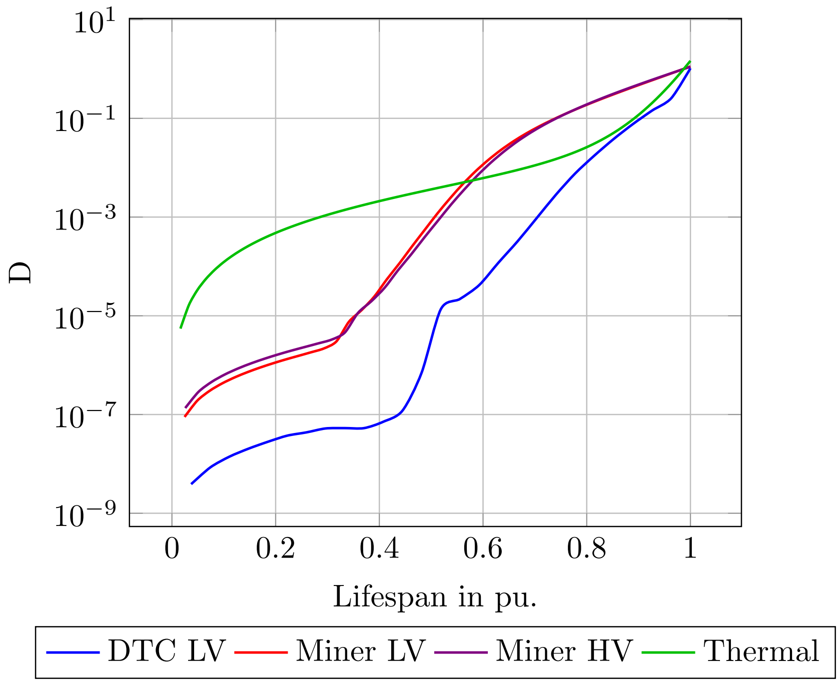


Figure 4.10: Evolution of the index D in regard to the lifespan and the thermal model.

difference in time. It can be seen that at 19:00 hours, when the peak load is located, the temperature is also in the way towards reaching also its peak.

The thermal model was applied to find the curve presented in Figure 4.10. When applying this same model to the whole year, until the loss of life reached 100%, there were 61 years of lifespan. As in the cases before, a yearly increment of 3% was included.

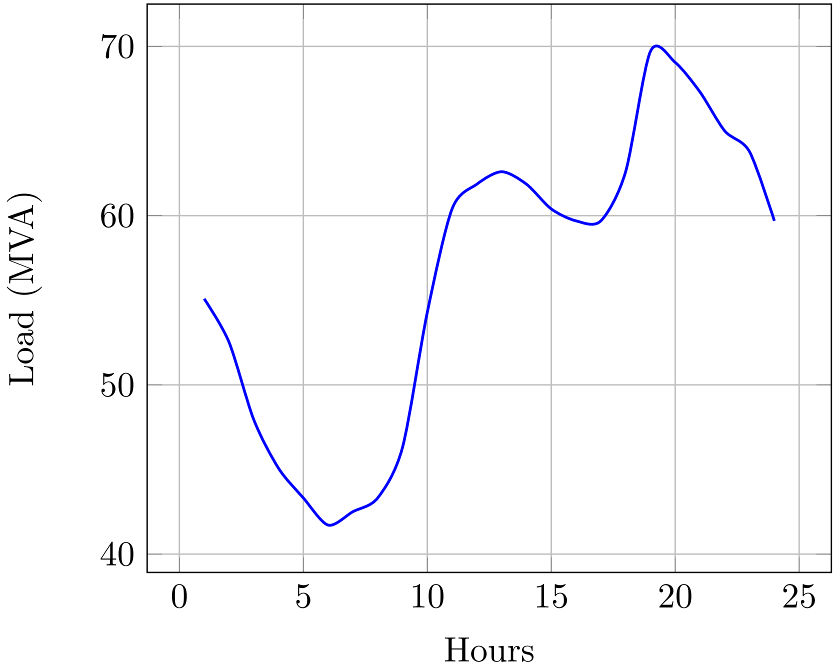


Figure 4.11: Load in the transformer for the first day of the year.

4.5.2.2 Conductor lifespan compared with insulation lifespan

Figure 4.13 shows the copper lifespan for different values of ultimate strength, and the insulation lifespan calculated with the thermal model. When the ultimate strength is less than 30 MPa, the winding conductor lifespan is lower than that of the insulation. Usually, the transformer is designed considering the maxim fault current it will face. If that is the case, the maximum mechanical stress would be 15 MPa if the fault is calculated at the terminals of the transformer, which means that under this consideration, the copper would end its life before the insulation.

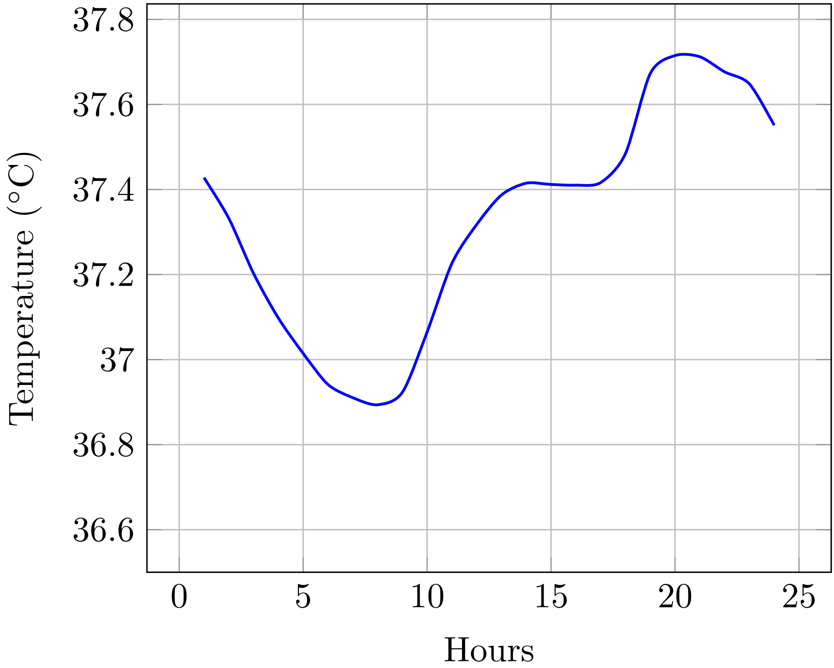


Figure 4.12: Load in the transformer for the first day of the year.

The result is more evident if the maximum fault current is calculated with the transformer as a part of the power system, which in this situation would give a stress of 5 MPa. The copper ages much faster than the inner insulation

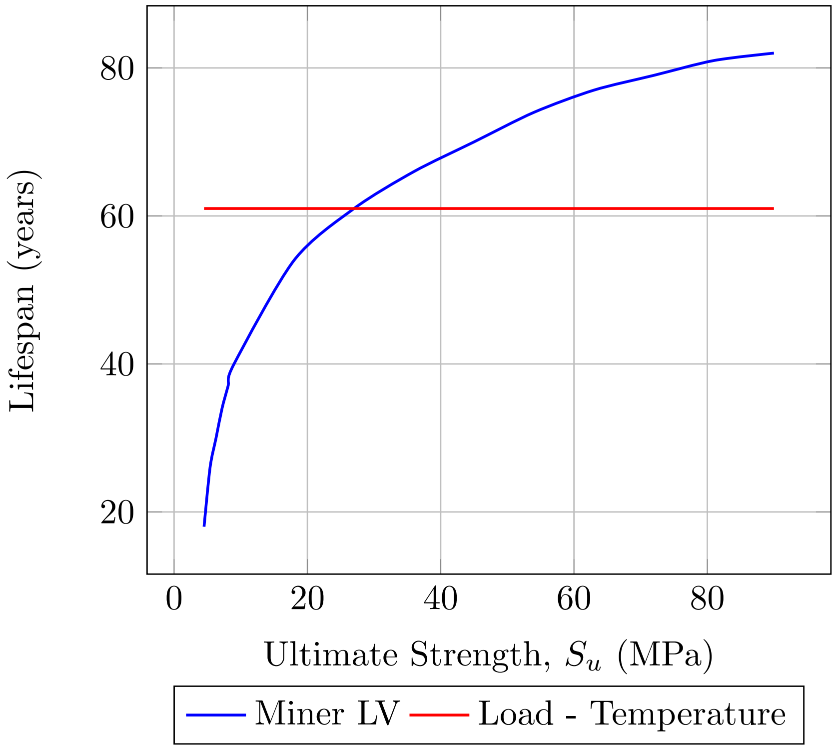


Figure 4.13: Lifespan in regard to the ultimate strength and lifespan according to the thermal model.

5 CONCLUSIONS AND RECOMMENDATIONS

5.1 CONCLUSIONS

The main contribution of this research was the development of a new method to determine the lifespan of the winding conductor. This method uses fatigue analysis to consider the continuos effect of mechanical stress in the conductor material. This takes a different approach to the usual lifespan analysis, which takes into account only the deterioration of insulation.

5.1.1 Effect of the Fatigue in the Lifespan of Power Transformers

Generally speaking, the transformer could be designed to support the electromagnetic forces produced by the highest possible fault current the transformer will face. The costumer is assured that the transformer will work well as long as the fault current does not increase to levels higher than planned. However, when analyzing the fatigue effect, it is evident that forces much less than those of design may cause a significant reduction in the lifespan of the power transformer.

The reason for this phenomenon is the continuous affectation that the fatigue has in the winding conductor. The fatigue effects are cumulative and irreversible. In that way, even those forces due to normal loading reduce the lifespan of the power transformer.

As an example of the degree of importance of this phenomenon, the life of the insulating paper was compared with the life of the winding conductor. For this purpose, studies were made in one of the transformers of the IEEE Reliability Test System, which allowed the use of different load curves according to the month, day and hour. Besides, it was possible to use the reliability indexes to generate faults in different parts of the power system.

In the example just described, the insulating paper reached a lifespan of 60 years. The maximum current fault produced a stress of 15 MPa. It was found that a winding conductor with an ultimate strength of until 50 MPa reached its end of life before 60 years. This means that, even a conductor designed to support more than three times the maximum expected stress, ages faster than the insulating paper.

Usually the fatigue effect has not become a problem because the winding conductor has been designed for higher levels of stress than the maximum expected. Nonetheless, with the reduction in costs, this positive criteria could change, and the quality of the conductor could be reduced. Moreover, when faults appear in conductors, such as buckling or tilting, it is generally attributed to external faults, but it is never investigated if the conductor reached a point of cumulative fatigue end of life. It should also be considered that, in some utilities, a post fault analysis is not performed, and the transformers are replaced without determining the cause of the final fault.

5.1.2 Fatigue Effects - Copper vs. Aluminum

Aluminum has a less ultimate strength than copper. However, due to the dissipation characteristics of each material, aluminum windings are designed to carry a current density half of that of copper windings. Hence, the aluminum conductors and disks are bigger and so, the mechanical stress is considerably reduced. In the case analyzed in this research, the ratio between the lifespans of aluminum and copper windings goes from tens to hundreds, not giving any space to doubt about the better performance of aluminum transformers.

5.1.3 Miner's Rule and Damage Transfer Concept

According to the Miner's rule, the transformer will work for longer times. The Damage Transfer Concept predicts lower windings lifespan. The difference between the results is very high. The main reason of this behavior is the limitation in the Miner's rule to differentiate the order of the mechanical stresses and its lack of consideration of the degree of change in the mechanical stress. For the Miner's rule, it is the same if a high mechanical stress comes before a lower one or if the lower mechanical stress comes first. Likewise, there is no difference if the next pulse is twice or ten times the current pulse, for the Miner's rule, those events are the same. On the other hand, the Damage Transfer Concept is very careful in this regard, by introducing an exponent μ which accelerates the index D towards one when the mechanical stress goes from a high to a low value. In fault transients, this cycle from high to low is the general behavior. As a consequence, the index D faster reaches the limit value.

A very important disadvantage of the Damage Transfer Concept method is the numerical instability. In common mechanical engineering problems, this defect is not evident because the mechanical stress usually has long cycles with one value before changing to a different mechanical stress. In those scenarios, the relation between the number of cycles n that the stress has been applied to the number of cycles that causes the destruction of the material N is not too low. For electrical transients, each mechanical stress is applied only for one mechanical cycle. Hence, the ratio $1/N$ is very low, mainly when the mechanical stress is much lower than the ultimate stress. For this reason, the ratio is in the order of 10^{-7} to 10^{-23}, which is the cause of the oscillation. Future researches could be performed to study how to overcome this difficulty when applying the method to mechanical stresses caused by electrical transients.

5.1.4 Random Forests Applied to Mechanical Stress

Although, initially, the application of machine learning was not planned to be a contribution of this research, in finding the most optimal method for a possible application of the fatigue analysis in real time, random forests turned out to be a good a feasible option.

Random forests have been proved to be very accurate in modeling mechanical stresses in winding conductors as a function of the currents circulating through the windings. The errors are less than 10% in most of the cases. Moreover, the design of the random forests model only required one hyperparameter to be tuned: the number of forests in the model. For this reason, the design process was relatively fast if compared with other machine learning techniques.

During the design of the random forests model, the error in the validation data set was lower than the error in the training set. This characteristic is not common to other machine learning models. Usually, the validation error is expected to be higher than the training set because the latter is used for the internal determination of the machine learning parameters. However, that is not the case for the model of this research, mainly because the training data were obtained using a finite element model without noise. If, for instance, the mechanical stresses were measured directly in the transformer, the training data would be modeled more accurately (noise included), and the validation data would be modeled more loosely with a higher error.

The error of the random forests model is higher when the mechanical stress has a lower value. Although the training data was normalized before training the model, the higher stresses had more influence in the determination of the internal parameters. For that reason, the random forests model was closer to the behavior of those higher stresses. Furthermore, since the main affectation of the winding conductors comes from higher stresses, there were more of them than those with lower values. Thus, lower stresses were not accurately modeled and had a high error.

5.1.5 Effect of the Inrush Current

The inrush current for the 400 MVA power transformer is less than 10% of the rated current, and does not have too much effect on the fatigue analysis. Due to the design and the size of the transformer, the magnetization effects are not important. A future study, where the excitation current, and hence the inrush current, is higher may give a better insight of the behavior of the mechanical stress for energization events.

5.2 RECOMMENDATIONS

It was exposed in the literature review that some fatigue effects have been found in distribution transformers due to the energization of the system after the presence of a fault. In order to have a better understanding of transformer faults, the utilities may include a complementary analysis of cumulative fatigue. Maybe the buckilng, tilting or any deformation found in the windings is not due exclusively to the mechanical stress of the last fault, but to the cumulative aging due to the fatigue of the material.

As long as the quality of the transformers is reduced due to the competition in the market, the fatigue analysis should be included in the design process. This implies data recollection improvements, since it will be needed to have the map of the expected operation and fault currents of the power transformer.

In regards to the application of machine learning in mechanical stress of windings, electromagnetic forces, etc., a future work may be to have a set of test systems. This set might have transformers with different capacities, dimensions, materials, etc. With that test system, future researches may be performed aiming towards the application of the methods in practical cases. In this research the database had to be generated, which was time consuming.

If there are new data or a new method for fatigue analysis considering mechanical non linear models, this study could be updated. The nonlinearity analysis is pesimistic, so, a better insight of the problem may be achieved. At this moment that was not possible due to the oscillations of the method.

6 REFERENCES

[1] J. Woodhouse, "Pas 55: Specification for the optimized management of physical infrastructure assets," in *PAS Workshop*, vol. 2007, 2004.

[2] W. Cigre, "Asset management decision making using different risk assessment methodologies," *Cigre Technical*, vol. 541, 2013.

[3] A. E. Abu-Elanien and M. M. A. Salama, "Asset management techniques for transformers," *Electric power systems research*, vol. 80, no. 4, pp. 456–464, 2010. DOI: 10.1016/j.epsr.2009.10.008.

[4] N. Petkova, P. Nakov, and V. Mladenov, "Real time monitoring of incipient faults in power transformer," in *Electricity Distribution*, Springer, 2016, pp. 221–240. DOI: 10.1007/978-3-662-49434-9_9.

[5] S. McArthur and E. Davidson, "Multi-agent systems for diagnostic and condition monitoring applications," in *IEEE Power Engineering Society General Meeting.*, IEEE, 2004, pp. 50–54. DOI: 10.1109/PES.2004.1372751.

[6] H. Zhou, K. Hong, H. Huang, and J. Zhou, "Transformer winding fault detection by vibration analysis methods," *Applied Acoustics*, vol. 114, pp. 136–146, 2016. DOI: 10.1016/j.apacoust.2016.07.024.

[7] P. Nirgude, B. Gunasekaran, Channakeshava, A. Rajkumar, and B. Singh, "Frequency response analysis approach for condition monitoring of transformer," in *The 17th Annual Meeting of the IEEE Lasers and Electro-Optics Society, 2004. LEOS 2004.*, 2004, pp. 186–189. DOI: 10.1109/CEIDP.2004.1364220.

[8] S. Ryder, "Diagnosing transformer faults using frequency response analysis," *IEEE Electrical Insulation Magazine*, vol. 19, no. 2, pp. 16–22, 2003. DOI: 10.1109/MEI.2003.1192032.

[9] M. Duval, "Dissolved gas analysis: It can save your transformer," *IEEE Electrical Insulation Magazine*, vol. 5, no. 6, pp. 22–27, 1989. DOI: 10.1109/57.44605.

[10] L. Lundgaard, "Partial discharge. XIII. acoustic partial discharge detection-fundamental considerations," *IEEE Electrical Insulation Magazine*, vol. 8, no. 4, pp. 25–31, 1992. DOI: 10.1109/57.145095.

[11] "IEEE guide for failure investigation, documentation, analysis, and reporting for power transformers and shunt reactors," *IEEE Std C57.125-2015 (Revision of IEEE Std C57.125-1991)*, pp. 1–84, 2015. DOI: 10.1109/IEEESTD.2015.8741317.

[12] J. Moubray, *Reliability-centered Maintenance*. New York, USA: Industrial Press, 2001, ISBN: 9780831131463. [Online]. Available: https://books.google.com.ec/books?id=bNCVF0B7vpIC.

[13] P. Jarman, Z. Wang, Q. Zhong, and T. Ishak, "End-of-life modelling for power transformers in aged power system networks," in *CIGRE 2009 6th Southern Africa Regional Conference, Cape Town, Southern Africa*, CIGRE, 2009.

[14] J. Rosenlind, “Lifetime modeling and management of transformers,” Ph.D. dissertation, KTH Royal Institute of Technology, 2013.

[15] J. Wang and L. Fu, “Power transformer life analysis based on Lambert W function,” in *E3S Web of Conferences*, EDP Sciences, vol. 252, 2021, p. 01 004.

[16] Rigatos and Gerasimos, *Intelligent renewable energy systems: modelling and control*, First. Switzerland: Springer, 2016, pp. 464–505. DOI: 10.1007/978-3-319-39156-4.

[17] M. Arshad, S. M. Islam, and A. Khaliq, “Power transformer asset management,” in *2004 International Conference on Power System Technology, 2004. PowerCon 2004.*, IEEE, vol. 2, 2004, pp. 1395–1398.

[18] A. Kachler and I. Hohlein, “Aging of cellulose at transformer service temperatures. Part 1: Influence of type of oil and air on the degree of polymerization of pressboard, dissolved gases, and furanic compounds in oil,” *IEEE Electrical Insulation Magazine*, vol. 21, no. 2, pp. 15–21, 2005. DOI: 10.1109/MEI.2005.1412215.

[19] J. Scheirs, G. Camino, M. Avidano, and W. Tumiatti, “Origin of furanic compounds in thermal degradation of cellulosic insulating paper,” *Journal of applied polymer science*, vol. 69, no. 13, pp. 2541–2547, 1998. DOI: 10.1002/(SICI)1097-4628(19980926)69:13<2541::AID-APP3>3.0.CO;2-A.

[20] T. Leibfried, M. Jaya, N. Majer, M. Schafer, M. Stach, and S. Voss, “Postmortem investigation of power transformers—profile of degree of polymerization and correlation with furan concentration in the oil,” *IEEE Transactions on Power Delivery*, vol. 28, no. 2, pp. 886–893, 2013. DOI: 10.1109/TPWRD.2013.2245152.

[21] R. Stebbins, D. Myers, and A. Shkolnik, "Furanic compounds in dielectric liquid samples: Review and update of diagnostic interpretation and estimation of insulation ageing," in *Proceedings of the 7th International Conference on Properties and Applications of Dielectric Materials (Cat. No.03CH37417)*, vol. 3, 2003, 921–926 vol.3. DOI: 10.1109/ICPADM.2003.1218572.

[22] R. Medina, "Desarrollo de indicadores para el análisis de riesgo en parques de transformadores de potencia dentro de un contexto de gestión de activos fisicos," in Spanish, Ph.D. dissertation, Universidad Nacional de San Juan, Jun. 2017.

[23] S. Forouhari and A. Abu-Siada, "Application of adaptive neuro fuzzy inference system to support power transformer life estimation and asset management decision," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 3, pp. 845–852, 2018. DOI: 10.1109/TDEI.2018.006392.

[24] "IEEE guide for loading mineral-oil-immersed transformers and step-voltage regulators," *IEEE Std C57.91-2011 (Revision of IEEE Std C57.91-1995)*, pp. 1–123, 2012. DOI: 10.1109/IEEESTD.2012.6166928.

[25] A. Franzén and L. Bertling, "State of the art-life time modeling and management of transformers," *KTH Electrical Engineering*, 2007.

[26] 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. 112 494–112 504, 2019.

[27] H.-M. Ahn, J.-Y. Lee, J.-K. Kim, Y.-H. Oh, S.-Y. Jung, and S.-C. Hahn, "Finite-element analysis of short-circuit electromagnetic force in power transformer," *IEEE Transactions on Industry Applications*, vol. 47, no. 3, pp. 1267–1272, 2011. DOI: 10.1109/TIA.2011.2126031.

[28] V. Behjat, A. Shams, and V. Tamjidi, “Characterization of power transformer electromagnetic forces affected by winding faults,” *J. Oper. Autom. Power Eng*, vol. 6, no. 1, pp. 40–49, 2018. DOI: 10.22098/joape.2018.2436.1210.

[29] C. Zhang, W. Ge, Y. Xie, and Y. Li, “Comprehensive analysis of winding electromagnetic force and deformation during no-load closing and short-circuiting of power transformers,” *IEEE Access*, vol. 9, pp. 73 335–73 345, 2021. DOI: 10.1109/ACCESS.2021.3068054.

[30] F. Friend, “Cold load pickup issues,” in *2009 62nd Annual Conference for Protective Relay Engineers*, IEEE, 2009, pp. 176–187.

[31] L. F. Blume, G. Camilli, S. B. Farnham, and H. A. Peterson, “Transformer magnetizing inrush currents and influence on system operation,” *Transactions of the American Institute of Electrical Engineers*, vol. 63, no. 6, pp. 366–375, 1944. DOI: 10.1109/T-AIEE.1944.5058946.

[32] N. Beniwal, H. Gupta, and D. Dwivedi, “Effect of creep on failure of distribution transformers: An experimental evaluation,” *International Journal of Performability Engineering*, vol. 6, no. 2, p. 171, 2010. DOI: 10.23940/ijpe.10.2.p171.mag.

[33] N. Micone, “Development of testing methodologies for the analysis of variable amplitude fatigue and corrosion-fatigue of offshore steels,” Ph.D. dissertation, Ghent University, Jan. 2017.

[34] S. M. Tavares and P. M. De Castro, “Damage tolerance of metallic aircraft structures: Materials and numerical modelling,” in Gewerbestrasse, Switzerland: Springer, 2019, ch. 1, pp. 3–16.

[35] J. F. Araujo, E. G. Costa, F. L. M. Andrade, A. D. Germano, and T. V. Ferreira, “Methodology to evaluate the electromechanical effects of elec-

tromagnetic forces on conductive materials in transformer windings using the von mises and fatigue criteria," *IEEE Transactions on Power Delivery*, vol. 31, no. 5, pp. 2206–2214, 2016. DOI: 10.1109/TPWRD.2016.2579165.

[36] "IEC ability to withstand short circuit," *IEC Std 60076-5*, pp. 1–45, 2006.

[37] R. A. Serway and J. W. Jewett, *Physics for scientists and engineers*, 10th ed. Boston, USA: Cengage learning, 2018.

[38] G. Bertagnolli, *Short-circuit Duty of Power Transformers: The ABB Approach*, 3rd ed. Zurich, Switzerland: ABB Trasformatori, 2007.

[39] J. A. Stratton, *Electromagnetic theory*. New Jersey: John Wiley & Sons, 2007, vol. 33.

[40] M. Renardy and R. C. Rogers, *An introduction to partial differential equations*. New York, USA: Springer Science & Business Media, 2006, vol. 13.

[41] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. USA: Wiley, 1998.

[42] J.-M. Jin, *The finite element method in electromagnetics*. Hoboken, New Jersey, USA: John Wiley & Sons, 2015.

[43] P. P. Silvester and R. L. Ferrari, *Finite elements for electrical engineers*, 3rd ed. Cambridge, UK: Cambridge university press, 1996.

[44] A. F. Peterson, S. L. Ray, R. Mittra, I. of Electrical, and E. Engineers, *Computational methods for electromagnetics*. USA: IEEE press New York, 1998, vol. 351.

[45] Z. Cendes, D. Shenton, and H. Shahnasser, "Magnetic field computation using delaunay triangulation and complementary finite element methods," *IEEE Transactions on Magnetics*, vol. 19, no. 6, pp. 2551–2554, 1983. DOI: 10.1109/TMAG.1983.1062841.

[46] S. Timoshenko, *Strength of materials Part 1*, 1st ed. USA: D. Van Nostrand Co., Inc, 1940.

[47] E. Alpaydin, *Introduction to machine learning*. Massachusetts: MIT press, 2020.

[48] H. J. Seltman, *Experimental design and analysis*, Reference material for the course Experimental Design for the Behavioral and Social Sciences, 2018. [Online]. Available: https://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf.

[49] S. Shalev-Shwartz and S. Ben-David, *Understanding machine learning: From theory to algorithms*. New York, USA: Cambridge university press, 2014.

[50] A. J. Hayter, *Probability and statistics for engineers and scientists*, 4th ed. Boston, USA: BROOKS Cole, Cengage Learning, 2012.

[51] W. C. Navidi, *Statistics for engineers and scientists*, 4th ed. New York, NY, USA: McGraw-Hill, 2015.

[52] G. Louppe, "Understanding random forests," Ph.D. dissertation, Cornell University Library, Jul. 2014.

[53] W.-L. Chao, *Machine learning tutorial*, Graduate Institute of Communication Engineering, National Taiwan University, Dec. 2011. [Online]. Available: http://disp.ee.ntu.edu.tw/~pujols/Machine%5C%20Learning%5C%20Tutorial.pdf.

[54] K. Tumer and J. Ghosh, "Estimating the Bayes error rate through classifier combining," in *Proceedings of 13th international conference on pattern recognition*, IEEE, vol. 2, 1996, pp. 695–699. DOI: 10.1109/ICPR.1996.546912.

[55] M. Claesen and B. De Moor, "Hyperparameter search in machine learning," in *MIC 2015: The XI Metaheuristics International Conference*, vol. 14, Agadir, Morocco, Jun. 2015, pp. 1–5.

[56] R. Ghawi and J. Pfeffer, "Efficient hyperparameter tuning with grid search for text categorization using knn approach with bm25 similarity," *Open Computer Science*, vol. 9, no. 1, pp. 160–180, 2019. DOI: 10.1515/comp-2019-0011.

[57] B. Shekar and G. Dagnew, "Grid search-based hyperparameter tuning and classification of microarray cancer data," in *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*, IEEE, 2019, pp. 1–8. DOI: 10.1109/ICACCP.2019.8882943.

[58] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization.," *Journal of machine learning research*, vol. 13, no. 2, 2012.

[59] A. Campbell, W. Chen, V. Stimper, J. M. Hernandez-Lobato, and Y. Zhang, "A Gradient Based Strategy for Hamiltonian Monte Carlo Hyperparameter Optimization," in *International Conference on Machine Learning*, PMLR, 2021, pp. 1238–1248.

[60] H. Jabbar and R. Z. Khan, "Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study)," *Computer Science, Communication and Instrumentation Devices*, pp. 163–172, 2015.

[61] L. Prechelt, "Early stopping-but when?" In *Neural Networks: Tricks of the trade*, Berlin, Heidelberg: Springer, 1998, pp. 55–69. DOI: 10.1007/3-540-49430-8_3.

[62] D. G. Altman and J. M. Bland, "Parametric v non-parametric methods for data analysis," *Bmj*, vol. 338, 2009. DOI: 10.1136/bmj.a3167.

[63] J. Xuan, J. Lu, and G. Zhang, "A survey on bayesian nonparametric learning," *ACM Comput. Surv.*, vol. 52, no. 1, Jan. 2019, ISSN: 0360-0300. DOI: 10.1145/3291044. [Online]. Available: https://doi.org/10.1145/3291044.

[64] D. Xu, Y. Shi, I. W. Tsang, Y.-S. Ong, C. Gong, and X. Shen, "Survey on multi-output learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 7, pp. 2409–2429, 2020. DOI: 10.1109/TNNLS.2019.2945133.

[65] S. Geman, E. Bienenstock, and R. Doursat, "Neural networks and the bias/variance dilemma," *Neural computation*, vol. 4, no. 1, pp. 1–58, 1992. DOI: 10.1162/neco.1992.4.1.1.

[66] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise reduction in speech processing*, Springer, 2009, pp. 1–4. DOI: 10.1007/978-3-642-00296-0_5.

[67] F. P. Beer, E. Johnston, J. DeWolf, and D. Mazurek, "Mechanics of materials," in 7th ed. New York, USA: McGraw-Hill, 2015, ch. 1, pp. 3–24.

[68] R. I. Stephens, A. Fatemi, R. R. Stephens, and H. O. Fuchs, *Metal fatigue in engineering*. John Wiley & Sons, 2000.

[69] G. T. Mase, R. E. Smelser, and G. E. Mase, *Continuum mechanics for engineers*, 2nd ed. CRC press, 2009.

[70] F. P. Beer, E. Johnston, J. DeWolf, and D. Mazurek, "Mechanics of materials," in 7th ed. New York, USA: McGraw-Hill, 2015, ch. 2, pp. 25–145.

[71] R. G. Budynas, J. K. Nisbett, *et al.*, "Shigley's mechanical engineering design," in 8th ed. New York, USA: McGraw-Hill, 2006, ch. 1, pp. 8–207.

[72] A. C. on Terminology, $ASTM$ *dictionary of engineering, science, and technology*, 10th ed. West Conshohocken, 2005.

[73] C. de Moura Branco, *Mecânica dos materiais*. Portugal: Fundacao Calouste Bulgenkian, 2011, In Portuguese.

[74] F. P. Beer, E. Johnston, J. DeWolf, and D. Mazurek, "Mechanics of materials," in 7th ed. New York, USA: McGraw-Hill, 2015, ch. 6, pp. 260–348.

[75] A. Zainab, D. Syed, A. Ghrayeb, *et al.*, "A Multiprocessing-Based Sensitivity Analysis of Machine Learning Algorithms for Load Forecasting of Electric Power Distribution System," *IEEE Access*, vol. 9, pp. 31 684–31 694, 2021. DOI: 10.1109/ACCESS.2021.3059730.

[76] H. E. Boyer *et al.*, *Atlas of fatigue curves*, 1st ed. Ohio, USA: ASM International, 1985.

[77] Y. Bai and W.-L. Jin, "Chapter 25 - Fatigue Capacity," in *Marine Structural Design (Second Edition)*, Y. Bai and W.-L. Jin, Eds., Second Edition, Oxford: Butterworth-Heinemann, 2016, pp. 489–507, ISBN: 978-0-08-099997-5. DOI: 10.1016/B978-0-08-099997-5.00025-3.

[78] A. Handbook, *Fatigue and Fracture, vol. 19*. USA: ASM International, 1996.

[79] R. Agrawal, R. Uddanwadiker, and P. Padole, "Low cycle fatigue life prediction," *International Journal of Emerging Engineering Research and Technology*, vol. 2, no. 4, pp. 5–15, 2014.

[80] O. Basquin, "The exponential law of endurance tests," in *Proc Am Soc Test Mater*, vol. 10, 1910, pp. 625–630.

[81] G. Noll and C. Lipson, "Allowable working stresses," *Society for Experimental Stress Analysis*, vol. 3, no. 2, pp. 89–109, 1946.

[82] R. G. Budynas, J. K. Nisbett, *et al.*, "Shigley's mechanical engineering design," in 8th ed. New York, USA: McGraw-Hill, 2006, ch. 6, pp. 260–349.

[83] Y. N. Loginov, S. Demakov, M. Ivanova, A. Illarionov, M. Karabanalov, and S. Stepatov, "Effect of annealing on properties of hot-rolled electrical copper," *The Physics of Metals and Metallography*, vol. 116, no. 4, pp. 393–400, 2015. DOI: 10.1134/S0031918X1502009X.

[84] E. B. Haugen, *Probabilistic mechanical design*. Michigan: Wiley-Interscience, 1980.

[85] M. Miner *et al.*, "Cumulative fatigue damage," *Journal of applied mechanics*, vol. 12, no. 3, A159–A164, 1945.

[86] A. McEvily, S. Ishihara, and M. Endo, "On the causes of deviation from the palmgren-miner rule," in *Fatigue Testing and Analysis Under Variable Amplitude Loading Conditions*, ASTM International, 2005. DOI: 10.1520/JAI19025.

[87] P. M. Subcommittee, "Ieee reliability test system," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 6, pp. 2047–2054, 1979. DOI: 10.1109/TPAS.1979.319398.

[88] I. Dasgupta, *Design of transformers*. Tata McGraw-Hill Education, 2002.

[89] J. J. Grainger, *Power system analysis*, 1st ed. New York, USA: McGraw-Hill, 1994.

[90] W. Li *et al.*, *Reliability assessment of electric power systems using Monte Carlo methods*. New York, USA: Springer Science & Business Media, 2013. DOI: 10.1007/978-1-4899-1346-3.

[91] L. Hewitson, M. Brown, and R. Balakrishnan, *Practical power system protection*. Oxford: Elsevier, 2004.

[92] D. Meeker, *Finite element method magnetics version 4.2: User's manual*, dmeeker@ieee.org, USA, 2010.

[93] F. Valencia, H. Arcos, and F. Quilumba, "Comparison of Machine Learning Algorithms for the Prediction of Mechanical Stress in Three-Phase Power Transformer Winding Conductors," *Journal of Electrical and Computer Engineering*, vol. 2021, DOI: 10.1155/2021/4657696.

[94] L. Buitinck, G. Louppe, M. Blondel, *et al.*, "API design for machine learning software: Experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 2013, pp. 108–122.

[95] A. Aeran, S. C. Siriwardane, O. Mikkelsen, and I. Langen, "A new non-linear fatigue damage model based only on sn curve parameters," *International Journal of Fatigue*, vol. 103, pp. 327–341, 2017. DOI: 10.1016/j.ijfatigue.2017.06.017.

[96] F. Valencia, H. Arcos, and F. Quilumba, "Prediction of stress in power transformer winding conductors using artificial neural networks: Hyperparameter analysis," *Energies*, vol. 14, no. 14, p. 4242, 2021. DOI: 10.3390/en14144242.

7 APPENDICES

7.1 AERAN FORMULATION AND THE DAMAGE TRANSFER CONCEPT ANALYSIS

7.1.1 Theoretical Foundations

Trying to solve the problem of the nonlinearity of the phenomenon of fatigue, Aeran et al. presented an alternative to perform a fatigue analysis [95]. The main objective was to obtain a model that only needs the SN curve. Aeran's model bases the diagnosis of the material in the same damage index D but formulated by (7.1). The exponent δ_i is found by (7.2).

$$D_i = 1 - \left[1 - \frac{n_i}{N_i}\right]^{\delta_i} \tag{7.1}$$

$$\delta_i = -\frac{1.25}{\ln N_i} \tag{7.2}$$

Aeran´s formulation is nonlinear, only uses the SN curve, and furthermore is closer to the experimental data. In order to better visualize the linear and nonlinear characteristics of Miner and Aeran's models respectively, Figure 7.1 shows the evolution of D for a constant stress.

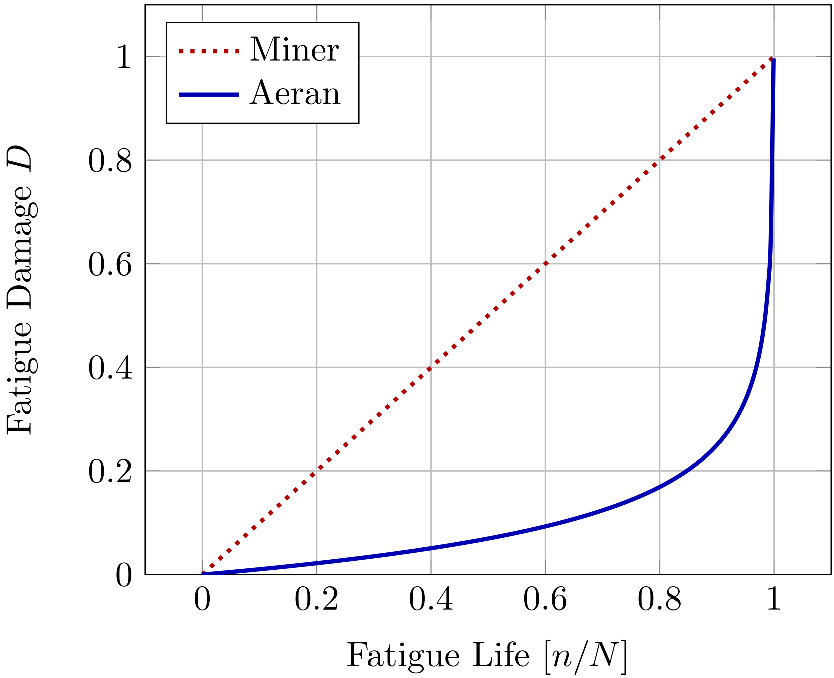


Figure 7.1: D determined by Miner and Aeran's models for a constant stress.

In addition to the formulation valid for a constant stress, Aeran et al. also developed a model for a fluctuating stress, which they called the Damage Transfer Concept (DTC), see Figure 7.2. The DTC assumes that once an initial stress σ_1 has acted on a body, and the system changes to a new state with a stress σ_2, a new curve must be constructed, where D_1 remains constant but with a new number of cycles n_{eff}, and with the same number of cycles to failure N_2 corresponding to the second stress. In other words, the relationship expressed by (7.3) is fulfilled. The term μ_{i+1} in the exponent represents the variation of the stresses, and is calculated by (7.4).

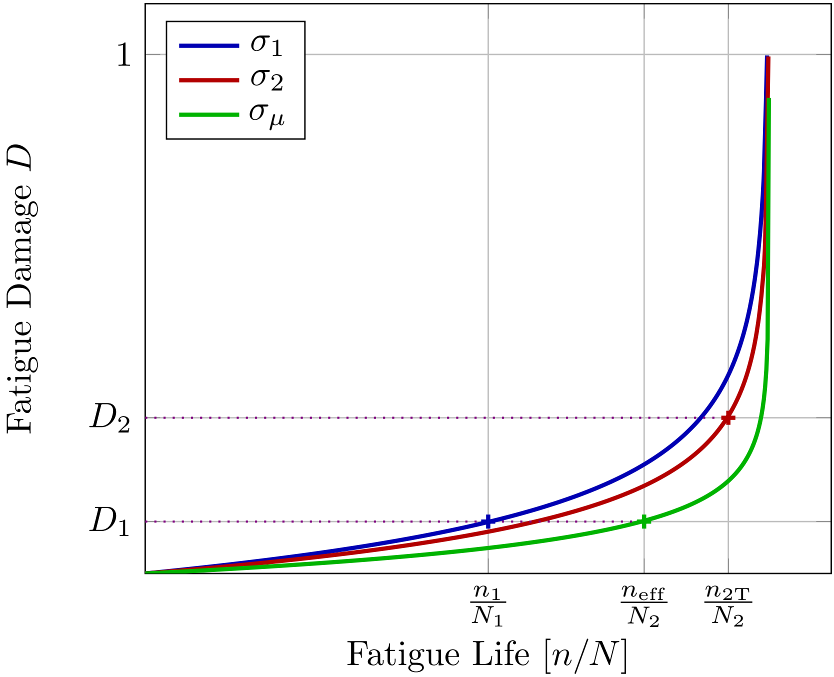


Figure 7.2: Relationships between the Fatigue Life and the Damage Index when the stress changes from σ_1 to σ_2 under a Damage Transfer Concept analysis.

$$D_i = 1 - \left[1 - \frac{n_{i+1,\text{eff}}}{N_{i+1}}\right]^{\frac{\delta_{i+1}}{\mu_{i+1}}} \tag{7.3}$$

$$\mu_{i+1} = \left(\frac{\sigma_i}{\sigma_{i+1}}\right)^2 \tag{7.4}$$

In practice, $n_{i+1,\text{eff}}$ is solved from (7.3), as in (7.5). Finally, the total number of equivalent cycles $n_{i+1,\text{T}}$ for the new stress is obtained by (7.6). $n_{i+1,\text{T}}$ is part of the σ_2 curve as seen in Figure 7.2, and is valid for a calculation of a new

D_{i+1}. The procedure continues until the damage index becomes one, i.e. when a fatigue failure has occurred.

$$n_{i+1,\text{eff}} = \left[1 - (1 - D_i)^{\frac{\mu_{i+1}}{\delta_{i+1}}}\right] \cdot N_{i+1} \tag{7.5}$$

$$n_{i+1,\text{T}} = n_{i+1} + n_{i+1,\text{eff}} \tag{7.6}$$

7.1.2 Oscillations of the Method

If a more pessimistic criterion is desired, it is apparent that using DTC is a better option over the Miner's rule. However, when introducing fault events, the DTC presents oscillations that do not allow having an accurate analysis.

For example, in Figure 7.3 the results of the 400 MVA power transformer subjected to random faults are shown. Note how the index D never rises constantly, which is expected since a continuous ageing constitutes a normal behavior. Even after the 15th year, the index is still close to 0 and has not been stabilized.

This oscillation brings up the question about the reason why this behavior only appears in the DTC model and not in the Miner's rule model. The problem comes when the number of mechanical stresses pulses is low. For instance, for the fatigue analysis during fault transients, 60 cycles have been chosen, which corresponds to 0.5 seconds, since each electrical cycle contains two pulses of force. Hence, a number n of 60 pulses of stress must be compared with a total number N of pulses previous to fault, which are in the order of 10^{20}. Therefore, the ratio n/N used in the determination of the index D is too small. This term

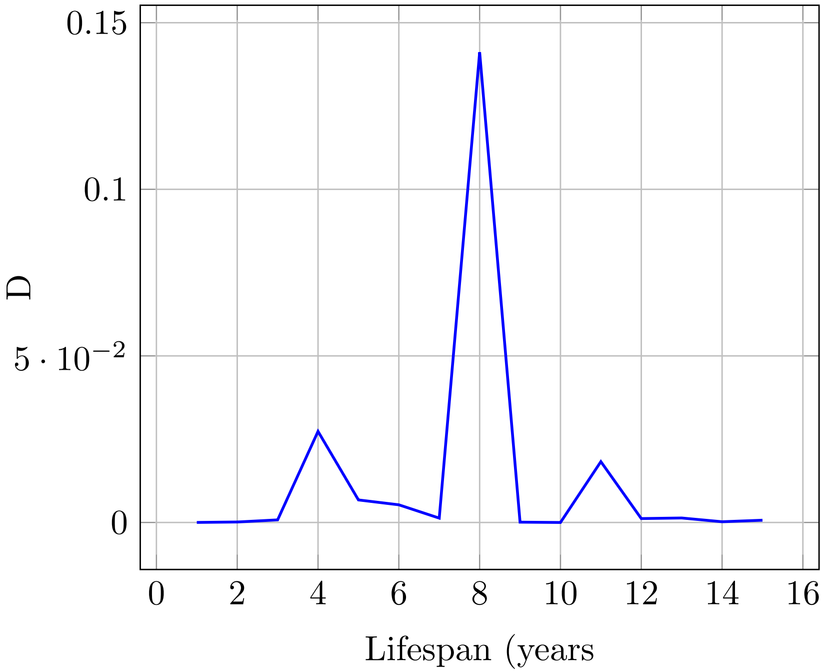


Figure 7.3: Oscillations of index D for DTC when the fatigue analysis considers fault events.

also exists in the Miner's rule method; however, in this case, the result adds up directly to the total D, hence the index D continues rising without problem. On the other hand, in the case of DTC, the n/N ratio is part of an expression that is raised to an exponent, which introduces the instability.

7.2 SCRIPTS DEVELOPED IN THIS RESEARCH

7.2.1 Stress Data Generation

Scripts to automate the FEM simulations to get the mechanical stresses in the middle disk of the windings for cooper and aluminum windings. They include a library that is a link between the FEMM software and Python.

Source Code 7.1: Automation of FEM simulations in FEMM

```
import femm

from time import time

def getStress(outputData, Winding, radius):
    stress=[];
    for ii in Winding:
        B = femm.mo_getb(ii[0],ii[1]);
        Jz = 1e6 * femm.mo_getj(ii[0],ii[1]);
        fxAux = -B[1] * Jz;
        #    f_perLength = fxAux * LV_Section;
        #    P = f_perLength * LV_radius;
        #    sigma = P/LV_Section;
        sigma = abs(fxAux * radius);
    #This is a simplified equation for the process as indicated in the
        three foregoing commented lines.
        stress.append(sigma); #force per volume unit
    outputData.extend(stress);

femm.openfemm()
femm.opendocument('transf_25_MVA.fem')
femm.mi_saveas("temp.fem")

```

```
inputData = [];

LV_y = 0.010 #y coordinate for the central LV disk
HV_y = 0.012 #y coordinate for the central HV disk

LV_x = [0.41, 0.83, 1.66] #x coordinates for the central LV disks
HV_x = [0.52, 0.72, 1.77] #x coordinates for the central HV disks

#Winding coordinates: A_right, A_left, B_right, B_left, C_right, C_left
LV_coordinates = [[-LV_x[2],LV_y],[-LV_x[1],LV_y],[-LV_x[0],LV_y],
                  [LV_x[0],LV_y],[LV_x[1],LV_y],[LV_x[2],LV_y]]

HV_coordinates = [[-HV_x[2],HV_y],[-HV_x[1],HV_y],[-HV_x[0],HV_y],
              [HV_x[0],HV_y],[HV_x[1],HV_y],[HV_x[2],HV_y]]

I_LV_max_1 = 148 #Current in the low voltage winding for the maximum
    point.
I_LV_2 = -74 #Current in the other two phases when phase 1 is at maximum
    .

I_HV_max_1 = I_LV_max_1 * 138 / 230
I_HV_2 = I_LV_2 * 138 / 230

file1 = open("temp0.txt","w")
currentFile = open("currentTemp.txt","w")

Ia = [];
Ib = [];
Ic = [];
start = time();

cases = 100;
I0 = 0.2
Istep = 0.01

```

```
LV_radius = (0.7761 + 0.8761) / 4
HV_radius = (0.9961 + 1.0969) / 4

for ii in range(cases):
    print(ii);
    outputData=[];
    I_c = (I0 + ii * Istep) * I_LV_max_1
    I_a = (I0 + ii * Istep) * I_LV_2
    I_b = I_a
    I_C = (I0 + ii * Istep) * I_HV_max_1
    I_A = (I0 + ii * Istep) * I_HV_2
    I_B = I_A
    femm.mi_modifycircprop('coil_a', 1, I_a);
    femm.mi_modifycircprop('coil_A', 1, I_A);
    femm.mi_modifycircprop('coil_b', 1, I_b);
    femm.mi_modifycircprop('coil_B', 1, I_B);
    femm.mi_modifycircprop('coil_c', 1, I_c);
    femm.mi_modifycircprop('coil_C', 1, I_C);
    femm.mi_analyze()
    femm.mi_loadsolution()
    mu = femm.mo_getmu(0,0);
    getStress(outputData, LV_coordinates, LV_radius)
    getStress(outputData, HV_coordinates, HV_radius)
    inputData = [I_a, I_b, I_c, I_A, I_B, I_C];
    L = [];
    for jj in range(len(outputData)):
        L.append(str(outputData[jj]));
        L.append(",");
    L.append(str(mu[0]));
    if ii != cases - 1:
        L.append('\n');
    file1.writelines(L);
    Lcurrent = [];
    for kk in range(len(inputData)):
        Lcurrent.append(str(inputData[kk]))
```

```
        Lcurrent.append(",");
    if ii != cases - 1:
        Lcurrent.append('\n');
    currentFile.writelines(Lcurrent);

file1.close();
currentFile.close()
end = time()
print(end - start); #Seconds
```

7.2.2 Random Forests Model

The script to train, evaluate and save the random forest model is shown below.

Source Code 7.2: Training of de Random Forests Model

```
import numpy as np
import pandas as pd
from sklearn.multioutput import MultiOutputRegressor
#from sklearn.pipeline import Pipeline
#from sklearn.preprocessing import PolynomialFeatures
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
import pickle

# Importing data files:

dfInputCurrent = pd.read_csv('current_data.txt', sep=",", header=None);
dfOutputStress = pd.read_csv('stress_data.txt', sep=",", header=None);

x_aux = dfInputCurrent.to_numpy()
y_aux = dfOutputStress.to_numpy()

```

```
x = np.delete(x_aux, -1, 1) #the last column was deleted to avoid the
    blank column generated during the collection of data
y = np.delete(y_aux, -1, 1) #last column deleted because it only has the
     permeability

# Standardization
# The following lines of code are used to standardize the data, so that
    all data is in the range $[0, 1]$
# The following equation was used for this purpose:
# $$
# x_{std} = \frac{x-x_{min}}{x_{max}-x_{min}}
# $$
#
# For the output, once the model is used for prediction, the new result
    must be calculated as following:
#
# $$
# y=y_{std} \cdot (y_{max}-y_{min})+y_{min}
# $$

x_max, x_min = x.max(0), x.min(0)
y_max, y_min = y.max(0), y.min(0)
deltaX = x_max - x_min
deltaY = y_max - y_min
x_std = x/deltaX - x_min/deltaX
y_std = y/deltaY - y_min/deltaY

# Random Forests implementation

regr_rf = RandomForestRegressor(n_estimators=100)
regr_rf.fit(x_std, y_std)

#save the model
filename = 'rf_model.sav'
pickle.dump(regr_rf, open(filename, 'wb'))
```

```
# # In[44]:


print(regr_rf.score(x_std,y_std))

# # # Testing the model in $R=1$ and $X_L=80$

# # In[5]:


# dfInput_val_1_80 = pd.read_csv('current_Phase_C_val_1_15.csv', sep=";"
    , header=None);
# dfOutput_val_1_80 = pd.read_csv('stress_Phase_C_val_1_15.csv', sep=";"
    , header=None);

# x_val_1_80 = dfInput_val_1_80.to_numpy();
# y_val_1_80 = dfOutput_val_1_80.to_numpy();

# #x_val_1_47 = x_val_1_80;
# #y_val_1_47 = y_val_1_47;


# # In[6]:


# x_std_val_1_80 = x_val_1_80/deltaX - x_min/deltaX
# y_std_val_1_80 = y_val_1_80/deltaY - y_min/deltaY


# # In[7]:


# filename = 'regr_rf.sav'
```

```
# model = pickle.load(open(filename, 'rb'))
# y_aux=model.predict(x_std_val_1_80)
# print(model.score(x_std_val_1_80,y_std_val_1_80))
```

7.3 POWER FLOW THROUGH THE POWER TRANSFORMER

Script to solve the power flow through the power transformer.

Source Code 7.3: Functions used in the power flow script

```
import numpy as np

def creationZbus():
    """Creates Zbus including the generators"""
    line_data = np.genfromtxt('IEEE_impedances.csv', delimiter=',')

    number_lines = 38
    number_buses = 24

    Y_line=np.zeros((number_lines), dtype=complex)
    B_line = np.zeros((number_lines), dtype=complex)

    for ii in range(number_lines):
        aux = complex(line_data[ii,2],line_data[ii,3])
        Y_line[ii]=aux**-1
        B_line[ii] = complex(0,line_data[ii,4]/2)

    Ybus = np.zeros((number_buses,number_buses), dtype=complex)

    #diagonal
    for jj in range(number_buses):
        for ii in range(number_lines):
            if line_data[ii,0] == jj + 1 or line_data[ii,1] == jj + 1:
```

```
                Ybus[jj ,jj] = Ybus[jj ,jj] + Y_line[ii] + B_line[ii]

    #Generators subsincronous reactance. =0.20 on each machine base.
    #It is converted to a 100 MVA base
    yd1 = (2/(20/20) + 2/(20/76))*(-1j)
    yd2 = yd1
    yd7 = 3/(20/100)*(-1j)
    yd13 = 3/(20/197)*(-1j)
    yd15 = (5/(20/12) + 1/(20/155))*(-1j)
    yd16 = 1/(20/155)*(-1j)
    yd18 = 1/(20/400)*(-1j)
    yd21 = yd18
    yd22 = 6/(20/50)*(-1j)
    yd23 = (2/(20/155) + 1/(20/350))*(-1j)

    yd = np.array([[1,yd1],[2,yd2],[7,yd7],[13,yd13],[15,yd15],[16,yd16
        ],
                    [18,yd18],[21,yd21],[22,yd22],[23,yd23]])

    for ii in yd:
        Ybus[int(ii[0]-1),int(ii[0]-1)] += ii[1]

    #off diagonal
    for ii in range(number_lines):
        bus1 = int(line_data[ii,0]-1)
        bus2 = int(line_data[ii,1]-1)
        Ybus[bus1,bus2] = Ybus[bus1,bus2] - Y_line[ii]
        Ybus[bus2,bus1] = Ybus[bus2,bus1] - Y_line[ii]

    print(Ybus[0,0])
    Zbus = np.linalg.inv(Ybus)

    return Zbus

def load_generation(week,day,hour):
```

```python

Base_Load = np.array([
    [1,108,22],[2,97,20],[3,180,37],[4,74,15],[5,71,14],
[6,136,28],[7,125,25],[8,171,35],[9,175,36],[10,195,40],
[13,265,54],[14,194,39],[15,317,64],[16,100,20],[18,333,68],
[19,181,37],[20,128,26]
                    ])

Weekly_Load = np.array
    ([[1,86.2],[2,90],[3,87.8],[4,83.4],[5,88],[6,84.1],[7,83.2],
[8,80.6],[9,74],[10,73.7],[11,71.5],[12,72.7],[13,70.4],[14,75],
[15,72.1],[16,80],[17,75.4],[18,83.7],[19,87],[20,88],[21,85.6],
[22,81.1],[23,90],[24,88.7],[25,89.6],[26,86.1],[27,75.5],
[28,81.6],[29,80.1],[30,88],[31,72.2],[32,77.6],[33,80],[34,72.9],
[35,72.6],[36,70.5],[37,78],[38,69.5],[39,72.4],[40,72.4],
[41,74.3],[42,74.4],[43,80],[44,88.1],[45,88.5],[46,90.9],
[47,94],[48,89],[49,94.2],[50,97],[51,100],[52,95.2]
                        ])

#The first data is monday, the second tuesday, etc.
Daily_Load = np.array([93,100,98,96,94,77,75])

if day in range(1,6):
    day_type = 1 #weekday
else:
    day_type = 2 #weekend

#Winter weeks: 1 to 8 and 44 to 52
Hourly_Load_Winter = np.array
    ([[1,67,78],[2,63,72],[3,60,68],[4,59,66],[5,59,64],[6,60,65],
[7,74,66],[8,86,70],[9,95,80],[10,96,88],[11,96,90],[12,95,91],
[13,95,90],[14,95,88],[15,93,87],[16,94,87],[17,99,91],[18,100,100],
[19,100,99],[20,96,97],[21,91,94],[22,83,92],[23,73,87],[24,63,81]])

#Summer weeks: 18 to 30
```

```
Hourly_Load_Summer = np.array
    ([[1,64,74],[2,60,70],[3,58,66],[4,56,65],[5,56,64],[6,58,62],
[7,64,62],[8,76,66],[9,87,81],[10,95,86],[11,99,91],[12,100,93],
[13,99,93],[14,100,92],[15,100,91],[16,97,91],[17,96,92],[18,96,94],
[19,93,95],[20,92,95],[21,92,100],[22,93,93],[23,87,88],[24,72,80]])

#spring weeks: 9 to 17. Fall weeks 31 to 43
Hourly_Load_Spring_Fall = np.array
    ([[1,63,75],[2,62,73],[3,60,69],[4,58,66],[5,59,65],[6,65,65],
[7,72,68],[8,85,74],[9,95,83],[10,99,89],[11,100,92],[12,99,94],
[13,93,91],[14,92,90],[15,90,90],[16,88,86],[17,90,85],[18,92,88],
[19,96,92],[20,98,100],[21,96,97],[22,90,95],[23,80,90],[24,70,85]])

Load = np.zeros((17,3))

weekly_factor = Weekly_Load[week-1,1]/100
daily_factor = Daily_Load[day-1]/100
hourly_factor = 1

if week in range(1,9) or week in range(44,53):
    hourly_factor = Hourly_Load_Winter[hour-1,day_type]/100

elif week in range(18,31):
    hourly_factor = Hourly_Load_Summer[hour-1,day_type]/100

elif week in range(9,18) or week in range(31,44):
    hourly_factor = Hourly_Load_Spring_Fall[hour-1,day_type]/100

else:
    print("wrong week")

Load[:,0] = Base_Load[:,0]
Load[:,1:] = Base_Load[:,1:] * weekly_factor * daily_factor *
    hourly_factor
```

```
    total_load = np.sum(Load, axis=0)

    Base_Generation = np.array
        ([[1,192],[2,192],[7,300],[13,591],[15,215],
                              [16,155],[18,400],[21,400],[22,300],[23,660]])


    total_generation = np.sum(Base_Generation, axis=0)

    Shared_Generation = np.zeros((10,2))

    Shared_Generation[:,0] = Base_Generation[:,0]
    Shared_Generation[:,1] = np.ceil(Base_Generation[:,1] * total_load
        [1] / total_generation[1])

    #Power Generated, Q Generated, P load, Q load, Voltage magnitud,
        Voltage angle, Type of Bus (0:slack, 1:PQ, 2:PV)
    number_buses = 24
    number_generators = 10
    number_loads = 17
    bus_data = np.zeros((number_buses,7))

    bus_data[:,6] = 1
    bus_data[0,6] = 0
    bus_data[:,4] = 1

    for ii in range(1,number_generators):
        bus_data[int(Shared_Generation[ii,0]-1),0] = Shared_Generation[
            ii,1]
        bus_data[int(Shared_Generation[ii,0]-1),6] = 2

    for ii in range(number_loads):
        bus_data[int(Load[ii,0]-1),2] = Load[ii,1]
        bus_data[int(Load[ii,0]-1),3] = Load[ii,2]

```

```
    return bus_data

def connectZbBetweenBuses(Zbus_original,Zb,busj,busk):
    """Return the modified Zbus once the line located between buses j and k
    has been eliminated."""
    #bus1, bus2 integers
    #Zb complex

    dimX = Zbus_original.shape[0]

    Zbus_aux = np.zeros((dimX + 1,dimX + 1),dtype=complex)
    Zbus_aux[0:-1,0:-1] = Zbus_original
    Zbus_aux[0:-1,-1] = Zbus_original[:,busj-1] - Zbus_original[:,busk-1]
    Zbus_aux[-1,0:-1] = Zbus_original[busj-1,:] - Zbus_original[busk-1,:]
    Zbb = Zbus_original[busj-1,busj-1] + Zbus_original[busk-1,busk-1] - \
        2*Zbus_original[busj-1,busk-1] + Zb
    Zbus_aux[-1,-1] = Zbb

    Zbus_aux2 = np.zeros((dimX,dimX),dtype=complex)
    for ii in range(dimX):
        for jj in range(dimX):
            Zbus_aux2[ii,jj] = Zbus_aux[ii,jj] - Zbus_aux[ii,-1] * \
                Zbus_aux[-1,jj] / Zbb

    return Zbus_aux2

def augmentNewBus(Zbus_original,Zb,busj):
    """"Return the modified Zbus once a new bus has been augmented through
    an impedance Zb connected to bus j."""
    #busj integer
```

```
    #Zb complex
    dimX = Zbus_original.shape[0]

    Zbus_aux = np.zeros((dimX + 1,dimX + 1),dtype=complex)
    Zbus_aux[0:-1,0:-1] = Zbus_original
    Zbus_aux[0:-1,-1] = Zbus_original[:,busj-1]
    Zbus_aux[-1,0:-1] = Zbus_original[busj-1,:]
    Zbus_aux[-1,-1] = Zbus_original[busj-1,busj-1] + Zb

    return Zbus_aux

def faultCurrentTransf(Zbus,busj,busk,Zb,alpha):
    """Returns a vector with the fault current amplitude circulating
        through
    the transformer, the resistance, and the reactance of the Thevenin
    equivalent. The faulted line is located between buses j and k at a
        fraction
    alpha of the line considered from bus j. Zbus is the original
        impednace
    matrix"""
    #Zbus: 2D array
    #busj and busk: integers
    #Zb: complex
    #alpha: float between 0 and 1

    currentVector = np.zeros(3)
    Ztransf = 0.0023 + 0.0839j

    if alpha == 0:
        print("Fault at bus 1.")
        Zpp = Zbus[busj-1,busj-1]
        Zjp = Zbus[8,busj-1]
        Zkp = Zbus[10,busj-1]
```

```
    elif alpha == 1:
        print("Fault at bus 2.")
        Zpp = Zbus[busk-1,busk-1]
        Zjp = Zbus[8,busk-1]
        Zkp = Zbus[10,busk-1]

    else:
        #Withdrawing the line from Zbus
        Zbus_aux = connectZbBetweenBuses(Zbus,-Zb,busj,busk)

        #Creating a new bus in the place of the fault
        Zbus_augm = augmentNewBus(Zbus_aux,alpha*Zb,busj)

        #Placing a line (1-alpha)Zb between the fault and bus k.
        busp = np.shape(Zbus_augm)[0]
        Zbus_fault = connectZbBetweenBuses(Zbus_augm,(1-alpha)*Zb,busk,
            busp)
        print(Zbus_fault[0,10])

        Zjp = Zbus_fault[8,busp-1]
        Zkp = Zbus_fault[10,busp-1]
        Zpp = Zbus_fault[busp-1,busp-1]

    I9_11 = -1/Ztransf * (Zjp - Zkp)/Zpp

    currentVector[0] = np.real(Zpp)
    currentVector[1] = np.imag(Zpp)
    currentVector[2] = abs(I9_11)

    return currentVector
```

Source Code 7.4: Power flow through the power transformer

```
# Program to generate the power flow data through the transformer for
    one year
```

```
# of 365 days with 24 hours each day. This means a total of 8760 hours.

import numpy as np
from numpy import genfromtxt
import power_flow_functions
import importlib

line_data = genfromtxt('IEEE_impedances.csv', delimiter=',')

number_lines = 38
number_buses = 24

S_transf = np.zeros((8760),dtype=float)
bus_trafo1 = 9
bus_trafo2 = 11

##################
#Creation of Ybus
##################

Y_line=np.zeros((number_lines), dtype=complex)
B_line = np.zeros((number_lines), dtype=complex)

for ii in range(number_lines):
    aux = complex(line_data[ii,2],line_data[ii,3])
    Y_line[ii]=aux**-1
    B_line[ii] = complex(0,line_data[ii,4]/2)

Ybus = np.zeros((number_buses,number_buses), dtype=complex)

#diagonal
for jj in range(number_buses):
    for ii in range(number_lines):
        if line_data[ii,0] == jj + 1 or line_data[ii,1] == jj + 1:
            Ybus[jj,jj] = Ybus[jj,jj] + Y_line[ii] + B_line[ii]
```

```
#off diagonal
for ii in range(number_lines):
    bus1 = int(line_data[ii,0]-1)
    bus2 = int(line_data[ii,1]-1)
    Ybus[bus1,bus2] = Ybus[bus1,bus2] - Y_line[ii]
    Ybus[bus2,bus1] = Ybus[bus2,bus1] - Y_line[ii]

index=0

for ww in range(52):

    print(ww)

    week = ww + 1

    print(week)

    for dd in range(7):
        day = dd

        for hh in range(24):

            hour = hh
            alpha = 1.6
            tol = 1e-5

            #importlib.reload(power_flow_functions)
            bus_data=power_flow_functions.load_generation(week,day,hour)

            #############
            #Gauss Seidel
            #############

            #Bus slack is bus 0. Vbase = 100 MVA. alpha: acceleration
```

```
        factor

    #Initial voltage on each bus
    Vbus = np.zeros((number_buses), dtype=complex)
    for ii in range(number_buses):
        x = bus_data[ii,4] * np.cos(bus_data[ii,5])
        y = bus_data[ii,4] * np.sin(bus_data[ii,5])
        Vbus[ii] = complex(x,y)

    #Initial scheduled power on each bus
    PQsch = np.zeros((number_buses), dtype=complex)
    for ii in range(number_buses):
        x = (bus_data[ii,0] - bus_data[ii,2]) / 100
        y = (bus_data[ii,1] - bus_data[ii,3]) / 100
        PQsch[ii] = complex(x,y)

    iter = 0
    epsR = np.ones(number_buses)
    epsR[0] = 0 #Slack bus
    eps = 1

    while eps > tol and iter < 200:
        for ii in range(number_buses):
            if bus_data[ii,6] != 0:
                if bus_data[ii,6] == 2:

                    sum = 0;
                    for kk in range(number_buses):
                        sum = sum + Ybus[3,kk] * Vbus[kk]

                    Qaux = -np.imag(np.conj(Vbus[3]) * sum)
                    PQsch[ii] = complex(np.real(PQsch[ii]),Qaux)

                sumV = 0
                for jj in range(number_buses):
```

```
                            if jj != ii:
                                sumV = sumV + Ybus[ii,jj]*Vbus[jj]

                        Vaux = 1/Ybus[ii,ii]*(np.conj(PQsch[ii])/np.conj(
                            Vbus[ii]) - sumV)

                        if bus_data[ii,6] == 1:
                            aux = Vbus[ii]
                            Vbus[ii] = Vbus[ii] + alpha * (Vaux - Vbus[ii])
                            epsR[ii] = np.real(Vbus[ii]) - np.real(aux)

                        elif bus_data[ii,6] == 2:
                            aux = Vbus[ii]
                            Vbus[ii] = abs(Vbus[ii])/abs(Vaux) * Vaux
                            epsR[ii] = np.real(Vbus[ii]) - np.real(aux)

                eps = np.amax(abs(epsR))
                #print(eps)
                iter = iter + 1

            # Power flow through the power transformer

            I_in_trafo = -Ybus[bus_trafo1-1,bus_trafo2-1] * (
                        Vbus[bus_trafo1-1] - Vbus[bus_trafo2-1])
            I12 =  I_in_trafo + Vbus[bus_trafo1-1] * B_line[13]
            I21 = -I_in_trafo + Vbus[bus_trafo2-1] * B_line[13]
            S9_11 = Vbus[bus_trafo1-1] * np.conj(I12)
            S11_9 = Vbus[bus_trafo2-1] * np.conj(I21)
            S_transf[index] = abs(S9_11)
            #print(np.sum(bus_data,axis=0))
            index = index + 1

#np.savetxt("transf_load.csv", S_transf, delimiter=",")
```

7.3.1 Power Transformer Life

The script below was developed to automize the calculation of the lifespan of the power transformer using fatigue analysis and the random forest model.

Source Code 7.5: Calculation of lifespan of the power transformer

```
import numpy as np
import pandas as pd
import pickle
import csv

def find_life(R_load,X_load,faults_year,fact_Su):
    Su = fact_Su * Su_base
    S_90 = 0.9 * Su
    S_30 = 0.3 * Su * k_sn
    b = (np.log10(S_30) - np.log10(S_90)) / 4
    a = 10 ** (np.log10(S_30) - b * 7) #408.33 a 90 MPa

    R_total = R_trafo + R_load
    X_total = X_trafo + X_load

    Imax_pu = 1/(R_total ** 2 + X_total ** 2) ** 0.5
    Imax = I_rated_lv * Imax_pu
    alpha = R_total / X_total * w

    i_input = []

    """
    Transient Current
    """
    for ii in range(61):
```

```python
        t = ii / 120
        i_lv = Imax * (np.cos(w*t) + np.exp(-alpha * t))
        i_hv = i_lv * RT
        i_input.append([i_lv, -0.5 * i_lv, -0.5 * i_lv,
                        i_hv, -0.5 * i_hv, -0.5 * i_hv])

    #Rated current
    i_input.append([I_rated_lv, -0.5 * I_rated_lv, -0.5 * I_rated_lv,
                    I_rated_hv, -0.5 * I_rated_lv, -0.5 * I_rated_lv])

    i_std = i_input / deltaX - x_min / deltaX

    stress_std = model.predict(i_std)
    stress = stress_std * deltaY + y_min

    N = (stress * 1e-6 / a) ** (1/b)

    N_aux = 1 / N
    D = N_aux.sum(axis=0) * faults_year + N_aux[-1] * number_cycles
    return 1/D.max(axis=0)

w = 2 * 3.1416 * 60
step = 1 / 120
RT = 138 / 230
I_rated_lv = 1.4142 * 400e6/(1.732 * 138e3)
I_rated_hv = RT * I_rated_lv

R_trafo = 0.009
X_trafo = 0.09

number_cycles = 365 * 24 * 3600 * 120 #each second, two hits of force

filename = 'regr_rf.sav'
model = pickle.load(open(filename, 'rb'))

```

```
"""
Data preparation:
    1. Pandas is used to import the data that was used to train the ML.
    2. The imported data is converted to numpy matrices.
    3. The last column is deleted because that is not part of this
        algorithm
    and may cause wrong results.
    4. The fourth block corresponds to the standardization procedure.
        These
    results are used in the ML model.
"""
dfInputA = pd.read_csv('current_Phase_A.csv', sep=";", header=None);
dfInputB = pd.read_csv('current_Phase_B.csv', sep=";", header=None);
dfInputC = pd.read_csv('current_Phase_C.csv', sep=";", header=None);

dfOutputA = pd.read_csv('stress_Phase_A.csv', sep=";", header=None);
dfOutputB = pd.read_csv('Stress_Phase_B.csv', sep=";", header=None);
dfOutputC = pd.read_csv('Stress_Phase_C.csv', sep=";", header=None);

x_A, x_B, x_C = dfInputA.to_numpy(), dfInputB.to_numpy(), dfInputC.
    to_numpy()
y_A, y_B, y_C = dfOutputA.to_numpy(), dfOutputB.to_numpy(), dfOutputC.
    to_numpy()

x = np.concatenate((x_A, x_B, x_C))
y = np.concatenate((y_A, y_B, y_C))


# dfInputCurrent = pd.read_csv('current_data.txt', sep=",", header=None)
    ;
# dfOutputStress = pd.read_csv('stress_data.txt', sep=",", header=None);

# x_aux = dfInputCurrent.to_numpy()
# y_aux = dfOutputStress.to_numpy()

```

```
# x = np.delete(x_aux, -1, 1)
# #last column was deleted to avoid the blank column generated during
    the
# #collection of data
# y = np.delete(y_aux, -1, 1)
# #last column deleted because it only has the permeability

x_max, x_min = x.max(0), x.min(0)
y_max, y_min = y.max(0), y.min(0)
deltaX = x_max - x_min
deltaY = y_max - y_min
x_std = x/deltaX - x_min/deltaX
y_std = y/deltaY - y_min/deltaY

##Data preparation finished here!
"""
Curva SN: S = a N**b
"""
#faults_year = 12 #faults per year
#R_load = 0
#X_load = 0
#fact_Su = 1
Su_base = 21 #Based on the maximum current at the terminals of the
    transformer
#Modifying factors
k_sn = 1.162 * 1.02 * 0.689 * 0.85 * 0.5

R_load = np.linspace(0,1,11)
X_load = np.linspace(0,1,11)
faults_year = np.linspace(1,50,11)
fact_Su = np.linspace(1,5,11)

#life = find_life(R_load, X_load, faults_year, fact_Su)

life = [[R,X,fy,fact,find_life(R, X, fy, fact)]
```

```
        for R in R_load
        for X in X_load
        for fy in faults_year
        for fact in fact_Su
        ]

with open('life_25', 'w') as f:
    write = csv.writer(f)
    write.writerows(life)
```

7.4 PUBLICATIONS FROM THE RESEARCH

There are two publications performed during this research that detail the process of finding a machine learning model to find the mechanical stresses in the windings of the power transformer. The first one presents the use of deep learning, the main difficulties and the criteria that must be followed during the design of the model [96]. The second publication compares the performance of different machine learning tools [93].

energies

Article

Prediction of Stress in Power Transformer Winding Conductors Using Artificial Neural Networks: Hyperparameter Analysis

Fausto Valencia *, Hugo Arcos and Franklin Quilumba

Faculty of Electrical Engineering, Escuela Politecnica Nacional, Ladrón de Guevara 253, Quito 170517, Ecuador; hugo.arcos@epn.edu.ec (H.A.); franklin.quilumba@epn.edu.ec (F.Q.)
* Correspondence: fausto.valencia@epn.edu.ec

Abstract: The purpose of this research is the evaluation of artificial neural network models in the prediction of stresses in a 400 MVA power transformer winding conductor caused by the circulation of fault currents. The models were compared considering the training, validation, and test data errors' behavior. Different combinations of hyperparameters were analyzed based on the variation of architectures, optimizers, and activation functions. The data for the process was created from finite element simulations performed in the FEMM software. The design of the Artificial Neural Network was performed using the Keras framework. As a result, a model with one hidden layer was the best suited architecture for the problem at hand, with the optimizer Adam and the activation function ReLU. The final Artificial Neural Network model predictions were compared with the Finite Element Method results, showing good agreement but with a much shorter solution time.

Keywords: artificial neural networks; deep learning; power transformers; stress; finite element method; electromagnetic forces

check for updates

Citation: Valencia, F.; Arcos, H.; Quilumba, F. Prediction of Stress in Power Transformer Winding Conductors Using Artificial Neural Networks: Hyperparameter Analysis. *Energies* **2021**, *14*, 4242. https://doi.org/10.3390/en14144242

Received: 9 June 2021
Accepted: 8 July 2021
Published: 14 July 2021

1. Introduction

Electromagnetic forces and stresses in conductors are usually determined with the Finite Element Method (FEM) [1–5]. As a numerical method, FEM needs the discretization of the medium, which results in the creation of nodes, each one represented as one row and one column in a matrix [6]. This process is not unique because FEM internally looks for the discretization that presents a slight field variation between nodes [7]. This is achieved through an iterative process, which lasts longer when the parameters of the problem have a nonlinear behavior—such as the permeability of the transformer core.

As a consequence, FEM could spend excessive time in obtaining one solution. When the transformer design requires the electromagnetic forces, this is not a problem because just a small number of calculations need to be performed. For example, if the procedure of evaluation of design recommended in [8] is followed, only the worst force and stress must be calculated, which implies one FEM simulation. Even if the dynamical analysis of the same reference is considered, 250 simulations should be performed at most, which correspond to 2 s at 60 Hz, at the highest value of short circuit current and the most pessimistic transient conditions.

However, if the forces must be calculated multiple times, FEM could become cumbersome. For that reason, finding a method that quickly determines the stress in winding conductors of a power transformer, such as an artificial neural network (ANN), may be very useful.

The interaction between artificial intelligence and electromagnetic forces is not widely studied in the research community.

Some investigations are focused on the analysis of possible damage inside the equipment. For example, in [9], ANNs were applied to classify the condition of power transformers as normal, degraded, or anomalous. The novelty of the research was the identification of nonlinearities in the response of the vibrations before electrical load variations. In the experimental setup, the vibrations due to electromagnetic forces in windings were separated from those due to magnetostriction in the core. The ANN classification results were compared with those of a Naive Bayesian Classifier, giving better accuracy.

Other related investigations are not focused on forces but on the determination of the magnetic field. In [10], magnetic fields were calculated with the use of Convolutional Neural Networks (CNN). In that research, a transformer with variable dimensions of the core and windings was included in the analysis. The authors used pictures of the magnetic field distribution as input for training. Each pixel was an input for the CNN. The training process was of the supervised kind. The authors used dropout layers to increase the generalization of the network. The dropout neurons were probabilistically chosen through a Bayesian Monte Carlo Technique. This same technique was used to detect if the problem was not suitable to be modeled with a CNN. An important feature of the research is the use of dilated filters, which fit well with the behavior of the magnetic field, which is influenced by the surrounding fields. The main limitation of the research is that the electrical current was not varied when creating the training data. This restriction limits the application of the CNN model when the goal is its use in practical problems. The study is limited to only one current, which could become a restriction when the fields and the forces are required for different transformer operating conditions.

A similar approach to using CNN—treating the magnetic field like a picture—was used in [11]. In this case, the problem is the determination of electromagnetic scattering when there is an incident wave. Finite differences in the frequency domain were used to generate the training data corresponding to wave and scattering pictures. The novelty of the paper is the use of deep residual networks to improve the network accuracy despite the number of hidden layers. The main difference with the problem of the fields in power transformers is the characteristics of the internal components' dimensions. While the core has lengths in the order of one meter, the winding conductors or disks could be as thin as millimeters. This wide range of dimension values brings difficulties in treating the pictures required to use this technique.

This research presents an ANN design to determine the stress in the conductors of a 400 MVA, three-phase power transformer produced by the circulation of three-phase fault currents.

FEM simulations are used to train the ANN, where a set of possible fault currents is the input data, and the corresponding stresses in the conductors are the output data. As discussed, each FEM simulation takes a long time. Thus, if the objective is to monitor the effects of the stress on the power transformer in real-time, the FEM simulations would not be the most appropriate option. On the other hand, when developing the ANN model, the long FEM simulation time is needed just during the data collection. After that, the ANN model can predict the stress due to any electrical current inside the range that was established during the creation of the model. As a consequence, there would be a handy ANN model fast enough to get results in real-time, and that could be used during the transformer's lifetime.

In regards to the design of the ANN model an objective is to find the best combination of alternatives based on three characteristics: architecture, activation functions, and optimizers used in the ANN model. The best combination is assessed by checking three attributes: the error for the training data, the generalization with the validation data, and the convergence time of the process.

2. Artificial Neural Networks Characteristics

Two concepts are commonly used in ANN literature: parameters and hyperparameters. Parameters change during the training process, i.e., they are controlled by the computational program—for example, the weights on each connection. Hyperparameters are defined outside the training process and usually are predefined, for example, the number of ANN layers [12]. Hence, the designer's task is to find the appropriate hyperparameters to get the lowest difference between the actual and the calculated value.

The difference between the calculated output value $\hat{y}$ and the actual value y is the error function of the model L, determined in Equation (1). This function is minimized during the training process.

$$\text{Minimize} \quad L = \sum (y - \hat{y})^2 \tag{1}$$

The objective of the ANN model design is to obtain the minimum error. The hyperparameters that the designer can change are the architecture, the optimizer, and the activation function. The architecture includes the number of layers and the number of neurons on each layer. The optimizer is the method to minimize the error function. The activation function introduces nonlinearities in the model.

2.1. Optimizers

Each iteration updates the weights to reach a point as close as possible to the error function's minimum. This optimization problem is multidimensional (one dimension per weight); hence, the best direction of change for the weights is given by the function's gradient to be optimized, i.e., the error function. This research has compared four optimizers: SGD, SGD with Momentum (SGDmom), RMSprop, Adam, whose implementations are shown from Equation (2) to Equation (5). α is the learning rate, η is the momentum, and v is the velocity of change. ρ, ϵ, m, β_1, and β_2 are factors according to each optimizer.

- SGD

$$\bar{w}_j \leftarrow \bar{w}_j - \alpha \nabla_w L \tag{2}$$

- SGD with Momentum (SGDmom) [13]

$$\begin{aligned} v_{t+1} &= \eta \cdot v_t + \nabla_w L \\ w_{t+1} &= w_t - \alpha \cdot v_{t+1} \end{aligned} \tag{3}$$

- RMSprop [14]

$$\begin{aligned} v_{t+1} &= \rho \cdot v_t + (1 - \rho) \cdot \nabla_w L^2 \\ \Delta w_{t+1} &= \eta \cdot \Delta w_t + \frac{\alpha}{\sqrt{v_{t+1} + \varepsilon}} \cdot \nabla_w L \\ w_{t+1} &= w_t - \Delta w_{t+1} \end{aligned} \tag{4}$$

- Adam [15]

$$\begin{aligned} m_{t+1} &= \beta_1 \cdot m_t + (1 - \beta_1) \cdot \nabla_w L \\ v_{t+1} &= \beta_2 \cdot v_t + (1 - \beta_2) \cdot \nabla_w L^2 \\ b_{t+1} &= \frac{\sqrt{1 - \beta_2^{t+1}}}{1 - \beta_1^{t+1}} \\ \Delta w_{t+1} &= \alpha \cdot \frac{m_{t+1}}{\sqrt{v_{t+1}} + \epsilon} \cdot b_{t+1} \\ w_{t+1} &= w_t - \Delta w_{t+1} \end{aligned} \tag{5}$$

2.2. Activation Functions

The activation functions that are part of the analysis in the present research are

- ReLU

$$f(x) = \max[0, x] \tag{6}$$

- Leaky ReLU [16,17]

$$f(x) = \begin{cases} x & \text{if} \quad x \geq 0 \\ \alpha x & \text{if} \quad x < 0 \end{cases} \tag{7}$$

- PReLU [18]: The factor *a* is determined adaptively during training [19].

$$f(x) = \begin{cases} x & \text{if} \quad x \geq 0 \\ ax & \text{if} \quad x < 0 \end{cases} \tag{8}$$

- ELU

$$f(x) = \begin{cases} x & \text{if} \quad x \geq 0 \\ \alpha(e^x - 1) & \text{if} \quad x < 0 \end{cases} \tag{9}$$

3. Description of the case

3.1. Power Transformer Characteristics

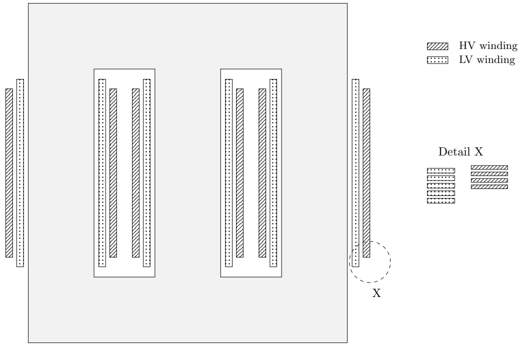
The transformer under analysis is a two-winding, three-phase power transformer. Table 1 presents its characteristics, Table 2 indicates the dimensions, and Figure 1a shows its internal structure. The windings are of a disk type. The transformer core material is US Steel Type 2-S, 0.018 inch thickness, whose saturation curve is represented in Figure 1b. This material is part of the FEMM software library [20].

Table 1. Power transformer under analysis.

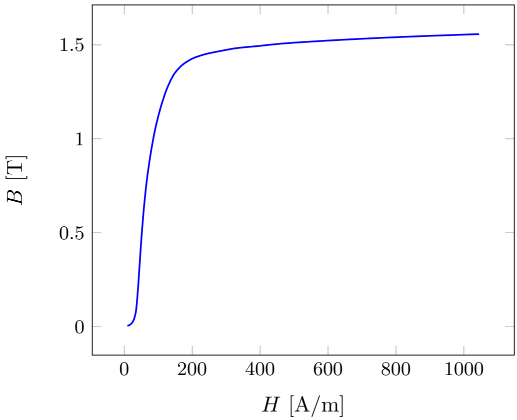
Variable	Value	Unit
Power	400	MVA
High Voltage	230	kV
Low Voltage	138	kV
Frequency	60	Hz
Group of Connection	Yy0	
Impedance (own base)	33.61	%
Number of Low-Voltage Disks	100	u
Number of High Voltage Disks	105	u

Table 2. Transformer internal dimensions.

Item	Length (m)
Core Diameter	0.9582
Core window height	3.034
Limb–limb separation	1.853
Low-voltage winding inside diameter	1.108
Low-voltage winding outside diameter	1.309
High-voltage winding inside diameter	1.429
High-voltage winding outside diameter	1.625
Low-voltage disk height	0.0175
High-voltage disk height	0.0136
Spacer block (located between disks)	0.010



(**a**)



(**b**)

Figure 1. Power transformer characteristics. (**a**) Internal structure; (**b**) Core saturation curve.

3.2. Currents in Windings

The currents considered to calculate the conductors' stress are those of three-phase faults seen by the transformer. The range goes from fault currents on the transformer terminals to fault currents with values lower than the rated current.

The peak values are taken from the transient currents. For simplicity, the origin of the fault is located when one phase current (A, B, or C) crosses through zero. In Equation (10), the general form to calculate the currents in the three phases of the low-voltage windings is shown [21]. The current previous to the fault has not been considered when generating the training samples; however, in the final results, it will be seen that the model is accurate even for preloaded cases. Table 3 shows the correspondence between the phase currents and those of Equation (10), according to the instant of the fault.

$$\begin{aligned} i_1(t) &= I_{LV} \cdot [\sin(\omega t - \phi) + \sin(\phi) \cdot \exp(-\lambda t)] \\ i_2(t) &= I_{LV} \cdot [\sin(\omega t - \theta - \phi) + \sin(\theta + \phi) \cdot \exp(-\lambda t)] \\ i_3(t) &= I_{LV} \cdot [\sin(\omega t + \theta - \phi) + \sin(-\theta + \phi) \cdot \exp(-\lambda t)] \end{aligned} \tag{10}$$

Table 3. Correspondence between the phase currents and Equation (10).

Phase Crossing the Zero Current	i_1	i_2	i_3
Phase A	i_a	i_b	i_c
Phase B	i_b	i_c	i_a
Phase C	i_c	i_a	i_b

I_{LV} is the maximum fault current when it has reached the steady-state. It depends on the impedance Z seen by the transformer during the fault—i.e., on the resistance R and the inductive reactance X_L—as shown in Equation (11), where V_{LV} is the peak line-to-neutral voltage of the low-voltage winding (112.7 kV for this study).

$$\begin{aligned} I_{LV} &= V_{LV}/Z \\ Z &= \sqrt{R^2 + X_L^2} \end{aligned} \tag{11}$$

where ω is the angular frequency ($2 \cdot \pi \cdot 60$ for a 60 Hz power frequency). θ is the angle to control the fault's time; one phase will see the fault at zero seconds or zero radians, whereas the other two phases will see the fault at $\pm 120°$, depending on the phase. ϕ is the angle of the impedance Z, i.e., $\cos\phi = R/Z$. λ is the time constant of the electromagnetic transient, i.e., $\lambda = R/L$, where L is the inductance of the inductive reactance X_L.

The transformer's impedance is the lowest limit, which, in the generation of input data, represents the current at the transformer's terminals. The leakage reactance is 33.61%, which is 16 Ω in the transformer base. The highest limit has been set until a reactance of 80 Ω, as Table 4 shows.

Table 4. Range of impedances used in the calculation of the phase currents.

Location of Fault	Resistance	Reactance
Transformer terminals	1 Ω	15 Ω
Far away from transformer	5 Ω	80 Ω

3.3. Electromagnetic Forces

Electromagnetic forces act on the conductors of a winding when the current interacts with the surrounding magnetic field [22]. In mathematical form, this is expressed in Equation (12); f is the force per volume, J is the current density, and B is the magnetic induction [23].

$$f = J \times B \tag{12}$$

Therefore, the main task is to find the magnetic induction B since J is part of the input data. From Electromagnetic Theory, B can be obtained by first calculating the magnetic vector potential A in Equation (13) (Laplace equation for magnetics) and then B in Equation (14) [24]. However, this process requires the treatment of partial differential equations (PDE), which usually do not have analytical solutions unless the problem has a simple geometry [25]. That is not the case for power transformers, which, on the contrary, have a complex geometry [26].

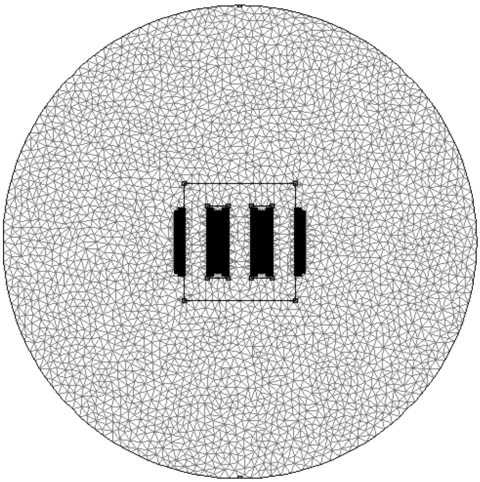
$$\nabla^2 A = -\mu J \tag{13}$$

$$\nabla \times B = A \tag{14}$$

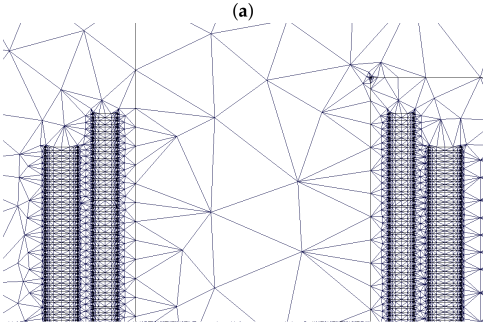
FEM is a numerical method used to solve partial differential equations that have the form of Poisson or Laplace equations. Hence, the method can be used to find the magnetic field B in power transformers. FEM discretizes the medium where the field must be determined.

Like any partial differential equation, Equation (13) also needs a boundary condition to have a unique solution. This condition is considered inside FEM and must be set during the simulation. FEM simulations for power transformers use the Dirichlet boundary condition, which sets a null value for A in some predetermined border. The ideal case would be to have a null vector magnetic potential at infinity because it works as a reference for the rest of the system. As it is not physically possible to have this condition, an appropriate border is chosen by following the convergence criteria for computational methods [27], which recommends performing simulations until no high variation is seen in the values of the potential inside the power transformer. Moreover, the boundary conditions have been located symmetrically around the equipment to have a similar influence in the internal fields.

Figure 2 presents the discretization of the 400 MVA power transformer under analysis, generated by the software FEMM. Note that around the disks, the discretization is denser. This characteristic of the triangulation process allows finding a more detailed field where it is suspected that the magnetic induction will have a higher change.



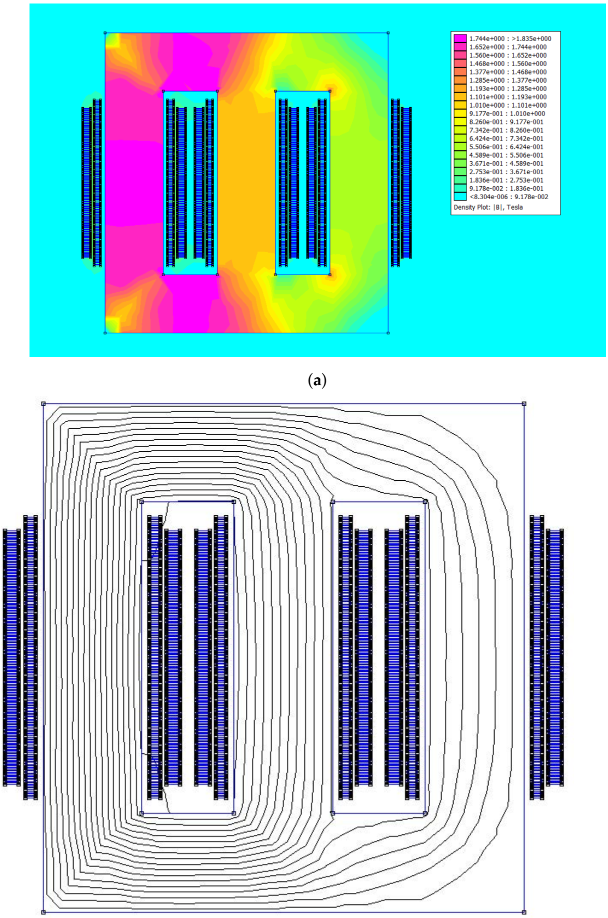
(a)



(b)

Figure 2. Discretization of 400 MVA power transformer. (**a**) Complete transformer; (**b**) Discretization around the disks.

Figure 3 shows the magnetic induction when the transformer is working under rated conditions. The current on Phase A is at its peak. The magnetic induction is the highest in the first limb, which belongs to Phase A. The same behavior is seen in the flux lines. Note that some lines cross the disks of the transformer. They represent the magnetic flux that, together with the current density, will produce a force in the disks.

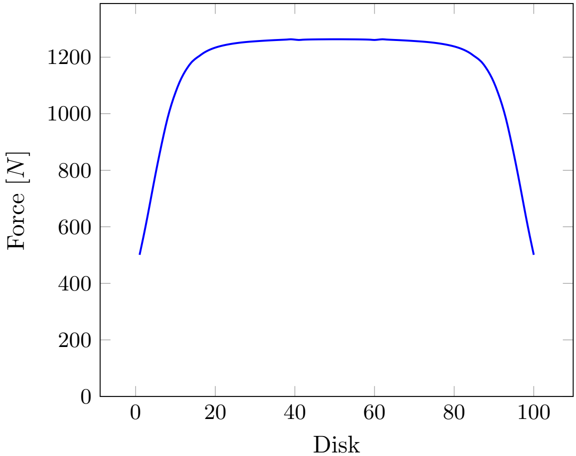


(**a**)

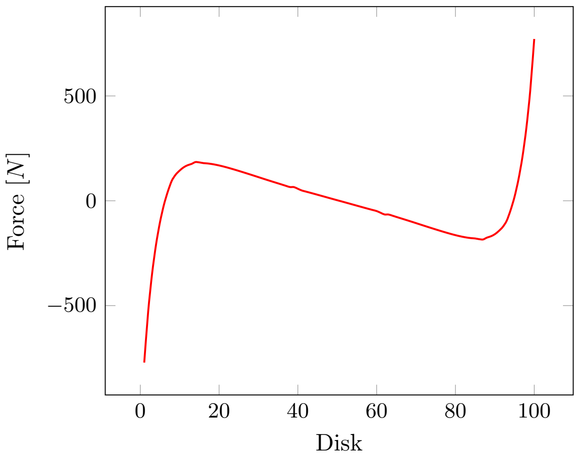
(**b**)

Figure 3. Magnetic induction for rated current. (**a**) Magnetic induction; (**b**) Flux lines.

For this same case, Figure 4 shows the radial and axial forces on the low-voltage winding. Disk 1 is the highest of the winding, and Disk 100 is the lowest. The radial forces are all positive, which means they are directed towards the core. The maximum radial force is located around the middle of the winding. The axial forces are negative in the highest disk and positive in the lowest disk. This characteristic means that they are directed towards the middle of the winding.



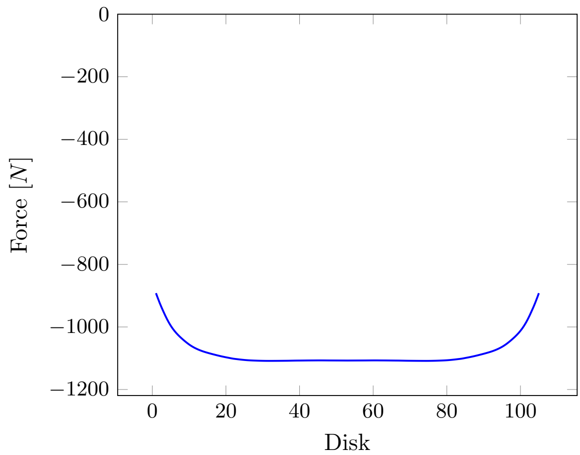
(**a**)



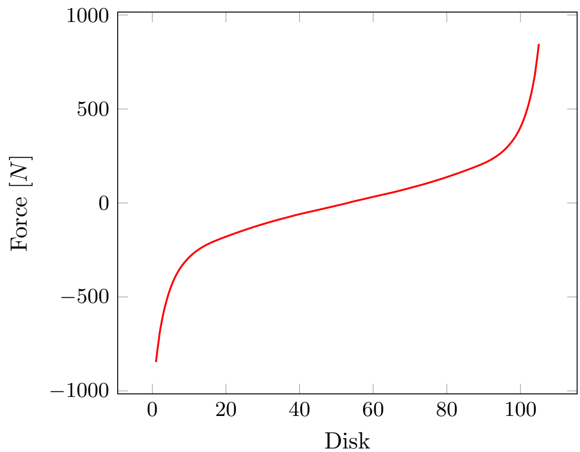
(**b**)

Figure 4. Forces on Phase A low-voltage winding. (**a**) Radial force; (**b**) Axial force.

In the high-voltage winding (see Figure 5), the radial forces are negative. They are directed away from the core. The axial forces have similar behavior to the low-voltage radial forces, which means that they point towards the middle of the winding.



(**a**)



(**b**)

Figure 5. Forces on Phase A high-voltage winding. (**a**) Radial force; (**b**) Axial force.

3.4. Winding Conductors Stress

The stresses in the winding conductors are the outputs of the ANN. For the sake of simplicity, the ANN design considers only the middle winding conductors, which have the highest value of stress and force. Thus, for the low-voltage winding, the stress is stored from disk 48 to disk 51; for the high-voltage winding, it is held from disk 49 to disk 51.

In addition, to reduce the simulation time, each disk of the winding is taken as one conductor. This simplification in the model is possible because the stress is defined as a per area magnitude, which does not have a high difference among the disk conductors when there is no considerable variation in the force per volume. The procedure developed in this research has taken the middle conductor of the disk to analyze the average value of stress.

The model considers each disk as a ring. Only the radial force is counted towards calculating the stress because it is the only one that can have a value without causing displacement of the ring. The axial force is nullified in the winding structure, either by another disk or by the winding supports; if that were not the case, the disk would have vertical movements [28].

Figure 6 shows a ring subject to a radial force. The wire radius is much lower than the radius R. The radial force f_r causes P, which is perpendicular to the wire section. Due to the force P, there is a stress σ acting on the wire. In that way, the stress σ is calculated according to Equation (15) [29]. The force f_r is the radial component of the force per unit length.

$$\begin{aligned} P &= f_r \cdot R \\ \sigma &= \frac{P}{Sc} \end{aligned} \tag{15}$$

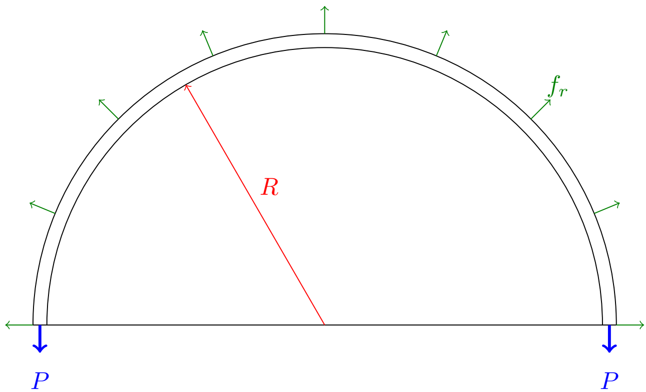
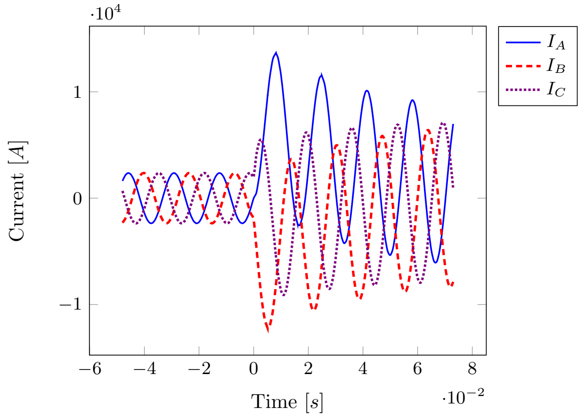


Figure 6. Ring subject to the action of a radial force f and a force P perpendicular to its section.

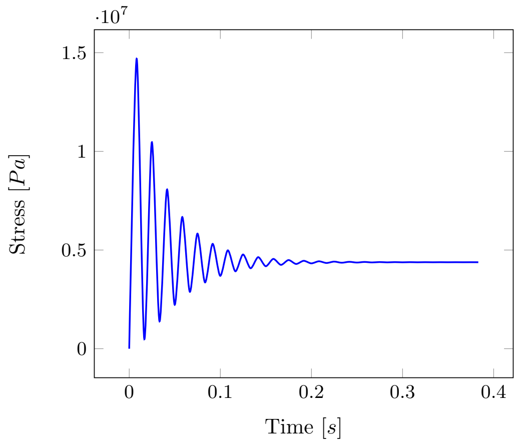
The following example uses a case with $R = 1\ \Omega$ and $X_L = 15\ \Omega$ to explain the approach to find the conductors' stress. Table 5 presents the values used in Equation (10) to calculate the current as a function of time. Figure 7a shows the three phases' current shape for a fault starting when the Phase A current is zero.

Table 5. Variables for the $R = 1\ \Omega$ and $X_L = 15\ \Omega$ case.

Variable	Value	Unit
α	25.13	s^{-1}
Z	15.03	Ω
ϕ	1.504	radians
V max	112.67	kV
I max	7495	A
θ	2.094	radians



(**a**)



(**b**)

Figure 7. Current and stress for a three-phase fault with a system equivalent impedance $R = 1$ and $X_L = 15$. (**a**) Current on the three phases; (**b**) Stress on phase A.

The current and the transformer internal geometry are the input data for the FEM simulations. The output of the FEM simulations is the magnetic field. Then, the stress is calculated as described before. Figure 7b shows the stress for the example.

4. ANN Design

The input data of the ANN are the currents circulating through the windings and the output data are the stresses on the conductors, as shown in Figure 8.

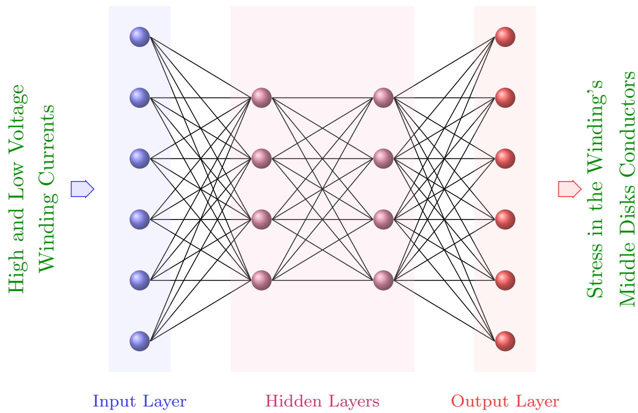


Figure 8. Artificial neural network used to find the stress in the conductors of the windings as a function of the input currents.

The design of the ANN follows three steps:

- Tune the architecture of the model;
- Tune the optimizer;
- Tune the activation functions.

SGD models are developed to obtain a set of possible architectures. The learning rate used during the process is 0.1 because a value less than that would take too much time and too many epochs to converge. The following architectures had the best performance and are part of the next design steps.

One hidden layer:

126	251	501	1000
1995	3981	7943	10,000

Two hidden layers:

[16 1000]	[32 501]	[32 1000]	[63 126]
[63 251]	[63 501]	[63 1000]	[126 32]
[126 63]	[126 126]	[126 251]	[126 501]
[126 1000]	[251 32]	[251 63]	[251 126]
[251 251]	[251 501]	[251 1000]	[501 16]
[501 32]	[501 63]	[501 126]	[501 251]
[501 501]	[501 1000]	[1000 16]	[1000 32]
[1000 63]	[1000 126]	[1000 251]	[1000 501]
[1000 1000]			

4.1. Tuning the Momentum and Epsilon

Three models from SGD are used to find the appropriate momentum. They represent the first, second, and third quartile: 3981, [501 126], and [63 126]. The usual behavior is to have a lower error as long as the momentum rises. The momentum with the lowest error is 0.9, as Table 6a shows in the mean value for the three cases tested. Therefore, a momentum of 0.9 is used when testing the optimizers.

For tuning epsilon in RMSprop, the first, second, and third quartiles SGDmom are taken, i.e., [501 251], 1000, [126 251]. An epsilon of 10^{-4} yields the lowest error for the three cases. This result is verified in the mean values of Table 6b.

The same quartiles of SGDmom are used to tune epsilon in Adam. The lowest error appears for the epsilon 10^{-4}. Table 6c confirms this result.

Table 6. Mean value for the error according to the momentum and epsilon.

(a) Momentum for SGDmom	
Momentum	Mean value
0.1	$6.3774 \cdot 10^{-6}$
0.3	$6.6169 \cdot 10^{-6}$
0.5	$6.2652 \cdot 10^{-6}$
0.7	$5.8841 \cdot 10^{-6}$
0.9	$4.8939 \cdot 10^{-6}$
(b) Epsilon for RMSprop	
Epsilon	Mean value
10^{-5}	$3.52 \cdot 10^{-6}$
10^{-4}	$2.90 \cdot 10^{-6}$
10^{-3}	$4.15 \cdot 10^{-6}$
10^{-2}	$5.01 \cdot 10^{-6}$
10^{-1}	$5.52 \cdot 10^{-6}$
1.0	$7.03 \cdot 10^{-6}$
(c) Epsilon for ADAM	
Epsilon	Mean value
10^{-7}	$6.60 \cdot 10^{-6}$
10^{-6}	$5.91 \cdot 10^{-6}$
10^{-5}	$4.14 \cdot 10^{-6}$
10^{-4}	$3.11 \cdot 10^{-6}$
10^{-3}	$3.80 \cdot 10^{-6}$

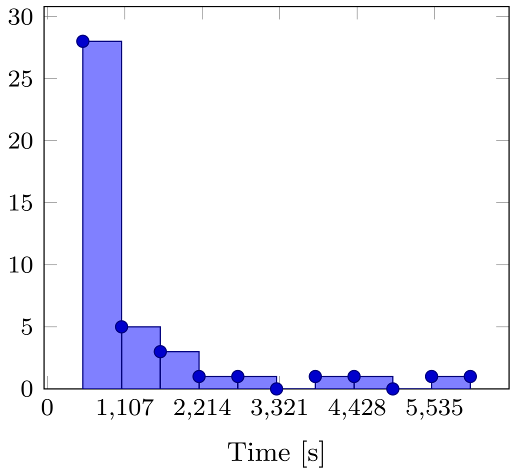
4.2. Comparison of Optimizers

Table 7 summarizes the hyperparameters used in the process for each optimizer. RMSprop and Adam have a lower learning rate than SGD and SGDmom because they are faster to converge.

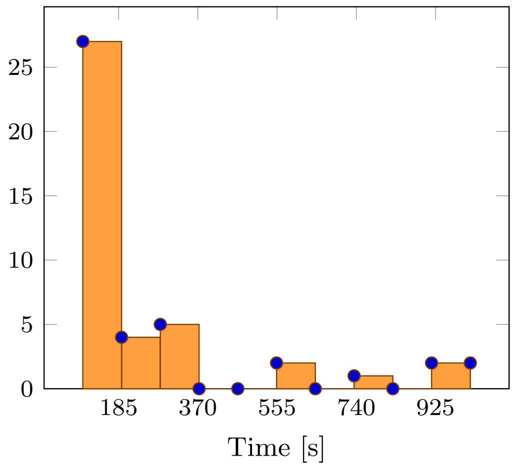
Table 7. Range of each optimizer hyperparameter.

Optimizer	**Hyperparameter**	**Value**
SGD	α	0.1
SGDmom	α	0.1
	η	0.9
RMSprop	α	0.01
	ρ	0.9
	η	0.9
	ϵ	10^{-4}
Adam	α	0.001
	β_1	0.9
	β_2	0.999
	ϵ	10^{-4}

Figure 9 shows the training time for each optimizer. By far, SGD needs the longest time to converge, even with a higher learning rate. It is about forty times greater than RMSprop and Adam times.

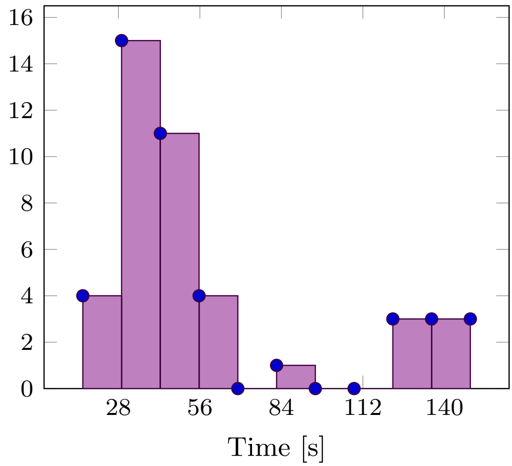


(**a**) SGD

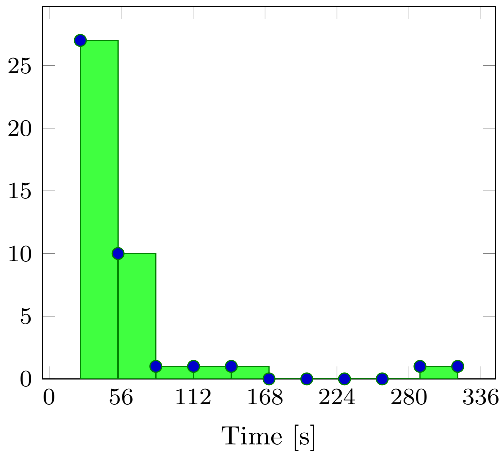


(**b**) SGD with momentum

Figure 9. *Cont.*



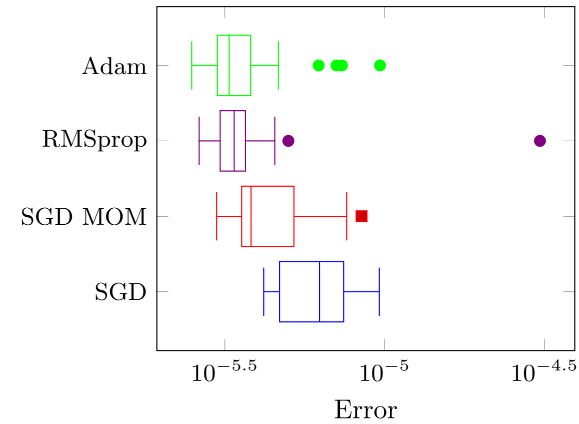
(**c**) RMSprop



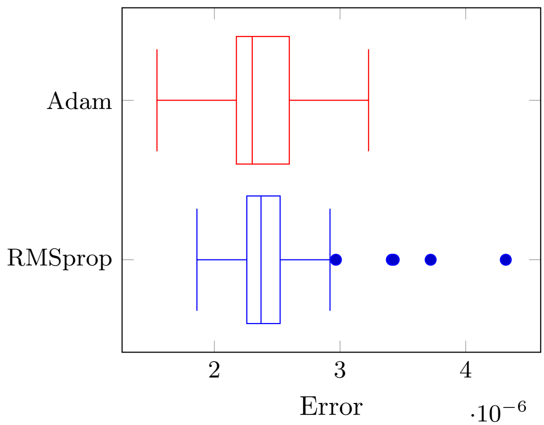
(**d**) Adam

Figure 9. Histograms for the convergence time of each optimizer.

In Figure 10a, the performance for the optimizers analyzed in this research is presented for a patience 10. RMSprop and Adam have the lowest error; hence, they are chosen for the analysis with patience 40, which is shown in Figure 10b. Both optimizers have similar errors, though Adam has a slightly lower value; therefore, it is chosen for the ongoing analysis.



(**a**) Training with Patience 10



(**b**) Training with Patience 40

Figure 10. Comparison of errors for different optimizers.

4.3. Choosing the Activation Function

Similar to the procedure to get the optimizers' hyperparameters, the first, second, and third quartiles obtained in the Adam process with patience 40 are used to analyze the activation functions. The related architectures are [251 32], [251 1000], and 126, respectively.

Table 8a shows the behavior of the error considering the variation of Alpha for Leaky ReLU. Alpha 0.1 presents the lowest error and is used in the detailed analysis. Table 8b shows the behavior of loss for alpha with the ELU activation function. The value of alpha 0.3 presents the lowest error.

Table 8. Mean value for the error according to Alpha.

a. Leaky ReLU with Adam	
Alpha	Mean value
0.1	$2.80 \cdot 10^{-6}$
0.2	$3.21 \cdot 10^{-6}$
0.3	$4.02 \cdot 10^{-6}$
0.4	$7.78 \cdot 10^{-6}$
0.5	$2.25 \cdot 10^{-5}$
b. ELU with Adam	
Alpha	Mean value
0.1	$5.12 \cdot 10^{-6}$
0.3	$3.42 \cdot 10^{-6}$
0.5	$7.05 \cdot 10^{-6}$
0.7	$8.20 \cdot 10^{-6}$
0.9	$1.83 \cdot 10^{-5}$
1.0	$1.29 \cdot 10^{-5}$

Figure 11 presents the loss distribution for the four activation functions: ReLU, Leaky ReLU, ELU, and PReLU. The behavior is similar for all of them. ReLU and Leaky ReLU have the lowest losses. ReLU is the activation function chosen because of its natural simplicity.

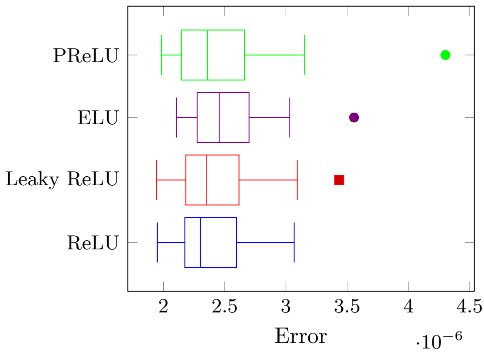


Figure 11. Comparison of activation functions.

To conclude, the best suited ANN has the following characteristics:

- Optimizer—Adam;
- Activation Function—ReLU.

4.4. Choosing the Final Architecture

Until this point, the design procedure has considered all the architectures described at the beginning of Section 4. Once the optimizer and the activation function have been settled, along with the related hyperparameters, the architectures included in the first quartile are used to continue with the design:

One hidden layer: 501, 1000, 1995, 3981, 7943, 10,000.

Two hidden layers: [16 1000], [63 1000], [501 16], [1000 251], [1000 1000].

Figure 12 shows the error for each architecture. The best four, one-hidden-layer architectures have a better performance than any two-hidden-layer architectures. In addition, they have a similar value of error. The architecture with one hidden layer and 1995 neurons is chosen for the simplicity of the model. According to the training processes, the model does not require additional tuning since the error will not decrease significantly.

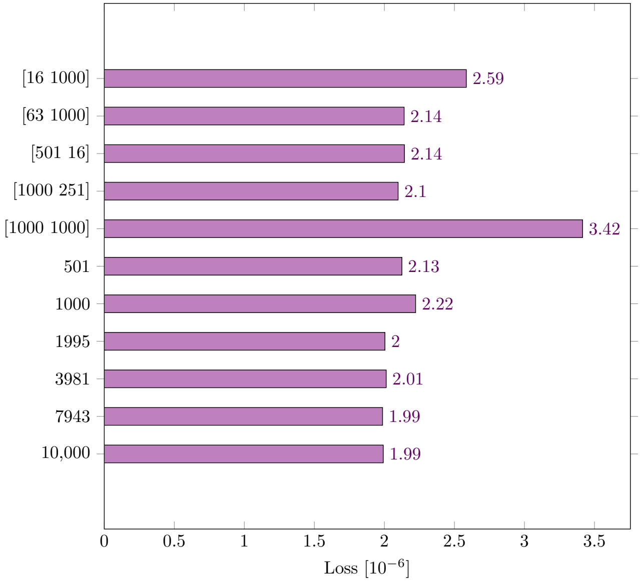


Figure 12. Errors for Patience 100, Optimizer—Adam.

Figure 13 presents the process of training the final design with a patience 100. The error consistently reduces until the number of epochs is close to the thousands. Then, the error is almost constant. There is a noise, which means that the algorithm has reached a minimum and is skipping from one point to another.

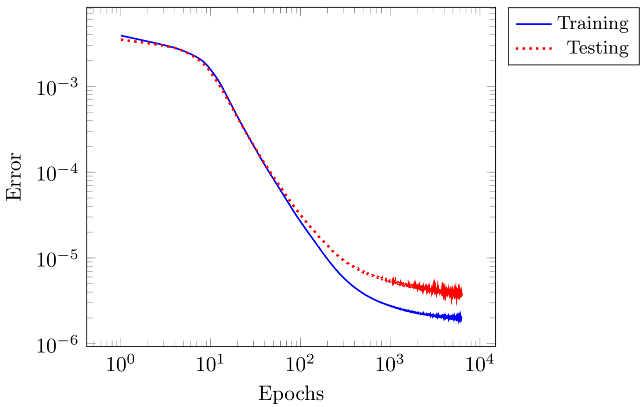


Figure 13. Behavior of the training and testing errors for the 1995 neurons, one-hidden-layer model.

The final design is validated with data that were not used for training nor testing during the design process. The error with the validation data was $7.8136 \cdot 10^{-8}$. The validation data has 795 samples from different values of resistance and reactance combinations.

5. Comparison to the Finite Element Method

5.1. Results in Actual Measurement Units

The model is applied to the transformer where the input data are the current through the windings in amperes, and the output data are the stresses on the winding's middle disks in Pascals. For the simulated cases, the stress is calculated at the points where the currents (hence, the forces and the stresses) have an extreme value. Figure 14 illustrates this point for one cycle. For Phase A, the stress is calculated in π and in 2π. For Phase B, the stress is calculated in $2\pi/3$ and in $5\pi/3$. Finally, for Phase C, the stress is calculated in $\pi/3$ and in $4\pi/3$. The process is repeated for the rest of the analyzed time.

All of the cases simulate a fault that begins when Phase A is crossing through zero, and the power transformer has a prefault rated current.

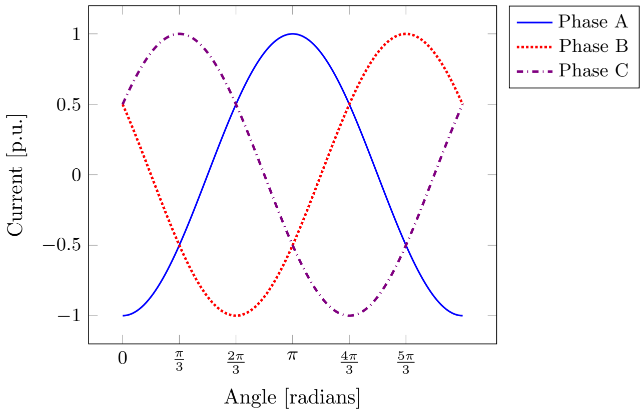
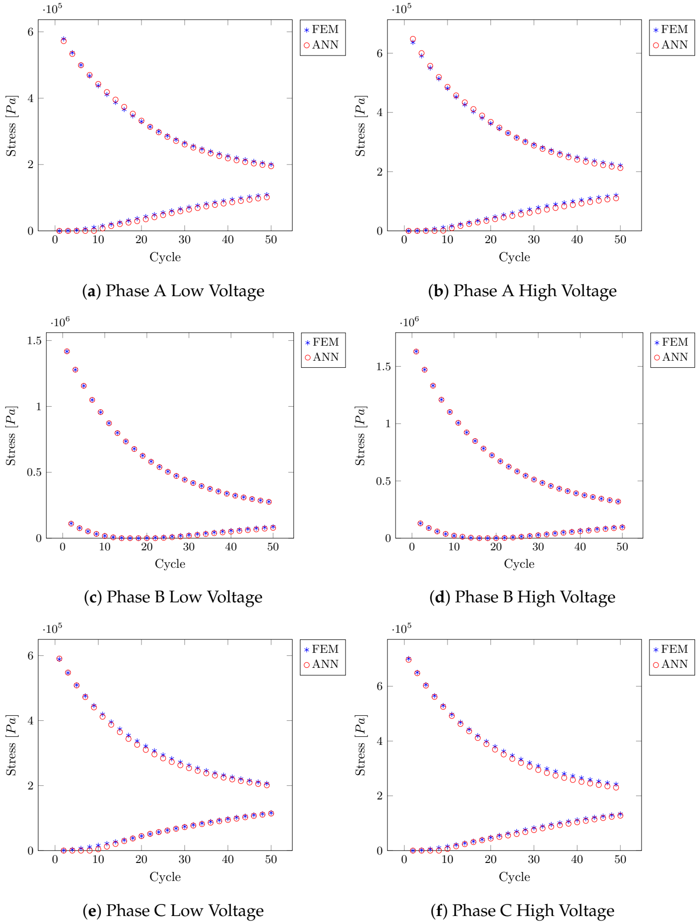


Figure 14. Extreme values location of current for each phase in one cycle.

The first case simulates a fault with resistance $R = 1\ \Omega$ and a reactance $X_L = 80\ \Omega$. This impedance represents a fault located far away from the transformer's terminals. Figure 15 shows the results for FEM and ANN simulations. The high X_L/R ratio extends the transient of the phenomenon. Thus, the cycles with high stress do not touch the cycles with low stress within the fifty cycle simulations. The results for FEM and ANN simulations are very close.



(**a**) Phase A Low Voltage

(**b**) Phase A High Voltage

(**c**) Phase B Low Voltage

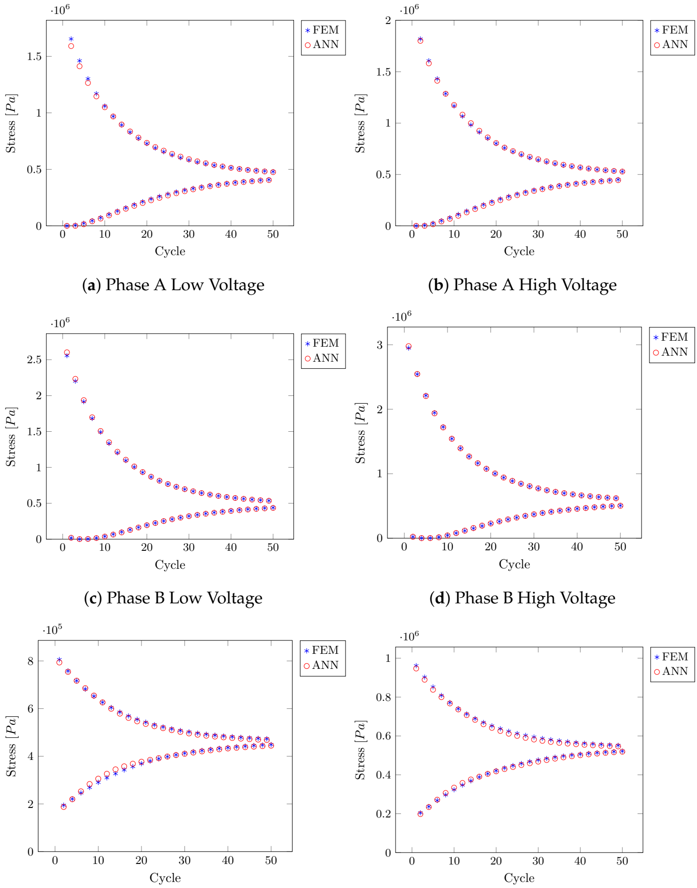
(**d**) Phase B High Voltage

(**e**) Phase C Low Voltage

(**f**) Phase C High Voltage

Figure 15. FEM and ANN simulations for a fault impedance of $R = 1\ \Omega$ and $X_L = 80\ \Omega$.

The second event is a fault with a resistance $R = 1\ \Omega$ and a reactance $X_L = 47\ \Omega$ (see Figure 16). The impedance is lower than in the first case, which causes higher stress on the conductors. The resistance/reactance ratio is also lower; hence, the transient ends faster and the difference between cycles of high stress and low stress is reduced at the end of the fifty cycles period. Some differences in the results of FEM and ANN are seen at the beginning of the phenomenon, mainly when the stress is high. Nonetheless, such difference is negligible.



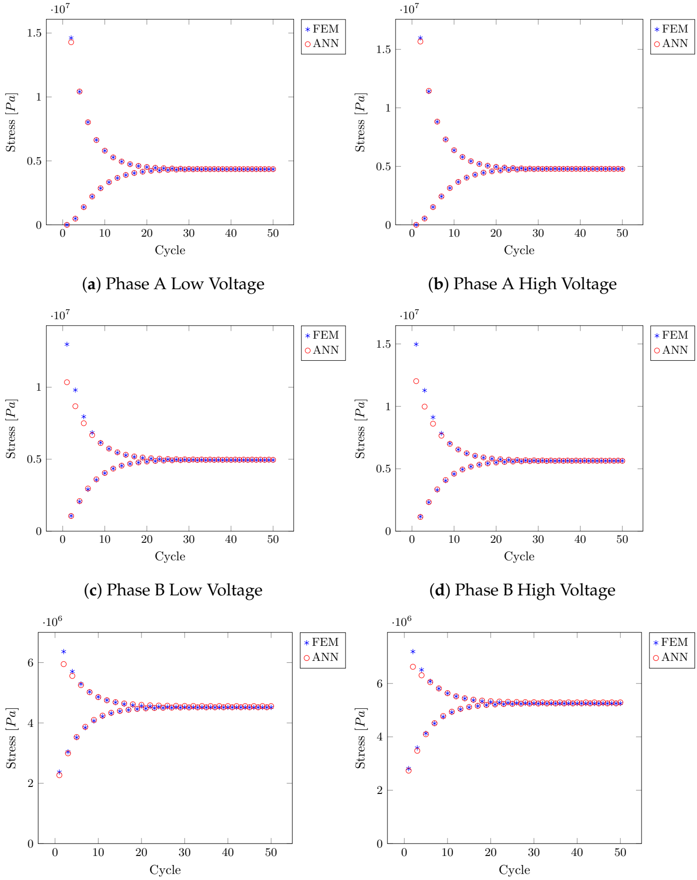
(**a**) Phase A Low Voltage (**b**) Phase A High Voltage

(**c**) Phase B Low Voltage (**d**) Phase B High Voltage

(**e**) Phase C Low Voltage (**f**) Phase C High Voltage

Figure 16. FEM and ANN simulations for a fault impedance of $R = 1\ \Omega$ and $X_L = 47\ \Omega$.

Finally, a fault at the terminals of the transformer, with a resistance $R = 1\ \Omega$ and $X_L = 15\ \Omega$, is analyzed (see Figure 17). The transient ends in about 20 cycles. The highest stress is present in Phase A, which agrees with the fault time when Phase A crosses zero. Most of the results are very close for FEM and ANN. A few cases where an appreciable difference exists are located at the beginning of the transient in Phase B and Phase C, for the highest stress values.



(**a**) Phase A Low Voltage (**b**) Phase A High Voltage

(**c**) Phase B Low Voltage (**d**) Phase B High Voltage

(**e**) Phase C Low Voltage (**f**) Phase C High Voltage

Figure 17. FEM and ANN simulations for a fault impedance of $R = 1\ \Omega$ and $X_L = 15\ \Omega$.

Table 9 presents the mean absolute percentage error for the foregoing cases. The model has the least error when the fault is at the terminals of the power transformer. As long as the fault is farther away, the error increases. In each case, the error is greater when the stress is lower. This behavior agrees with the training characteristics of ANNs that are more influenced by higher values than by lower ones.

Table 9. Mean absolute percentage error for the ANN model.

Phase	Winding	MAPE(%)
a. $R = 1$ and $X_L = 15$		
A	Low Voltage High Voltage	0.33 0.26
B	Low Voltage High Voltage	0.92 0.74
C	Low Voltage High Voltage	0.66 0.65
b. $R = 1$ and $X_L = 47$		
Phase	Winding	MAPE(%)
A	Low Voltage High Voltage	5.21 5.30
B	Low Voltage High Voltage	2.76 1.81
C	Low Voltage High Voltage	1.47 1.51
c. $R = 1$ and $X_L = 80$		
Phase	Winding	MAPE(%)
A	Low Voltage High Voltage	15.10 15.47
B	Low Voltage High Voltage	18.06 12.90
C	Low Voltage High Voltage	13.40 14.85

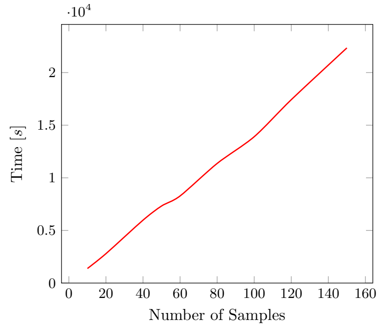
5.2. Simulation Time

A fault with $R = 1\ \Omega$ and $X_L = 15\ \Omega$ was simulated with FEM and the ANN model to obtain the time of the simulation.

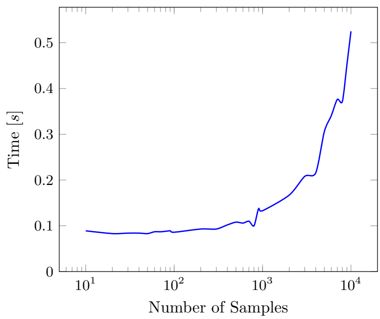
Figure 18a presents the time spent in FEM simulations. The relationship between the number of samples and the time of simulation is linear. Each sample takes about 143 s for its simulation. For this reason, the simulation of a phenomenon with 150 pulses of stress required about six hours.

On the other hand, ANN simulations require about 0.53 s for 10,000 samples, see Figure 18b. The reduction of time due to the direct application of the model is evident. This reduction was expected since the model contains only matrix operations and the use of the activation function. The computational system limits the time reduction; therefore, the minimum simulation time is constant (0.085 s) from the tens to the hundreds of samples.

The high number of samples that can be simulated in less than one second opens the possibility of using the model in real-time applications. Considering only the peak values of stress, there are six values per cycle (two per phase); for a power frequency of 60 Hz, that means a sample each $2.78 \cdot 10^{-3}$ s. In real-time, 10,000 samples would require 27.8 s. Consequently, the 0.53 s are enough for a real-time simulation of this kind. The limit would be a batch of 40 samples that has an average of about 470 samples per second to be simulated, which is higher than the 360 samples per second needed for stress analysis. This analysis was performed in a personal computer with an Intel i5, 64 bit, 2.50 GHz processor and 4.00 GB of RAM.



(**a**) Training with Patience 10



(**b**) Training with Patience 40

Figure 18. Simulation time according to the number of samples.

6. Conclusions

This research shows the possibility of modeling the stresses in winding conductors of a 400 MVA, three-phase power transformer using artificial neural networks. For this purpose, an analysis of the combination of hyperparameters that best reduces the error between the proposed model and the finite element analysis results was performed.

For the conductor stress problem, the single-layer models performed better than the two-layer ones. Additionally, more complex models took longer to minimize the error function, which is a disadvantage during the design process. In other words, it is not always true that a more complex model is better.

As the optimizers were more elaborated or had more hyperparameters, the algorithm found the minimum of the loss function faster. RMSprop and Adam had a better convergence speed than SGD or SGDmom. As a consequence, lower learning rates could be used, and the model achieved lower losses.

As for this research's problem, the activation functions did not have too much influence on the model's accuracy. ReLU was chosen, in the end, because of its simplicity since it does not have any hyperparameter and its implementation is straightforward.

With groups of data greater than 40 samples, it is even possible to use the artificial neural network model in real-time applications when the stress is needed only in peak values, such as in fatigue analysis.

A possible future investigation could be the design of an artificial neural network model for stresses caused by inrush currents. The problem has particular characteristics that make it challenging for an artificial neural network model. Mainly, the nonlinear features of the core could be an issue in modeling the phenomenon overall because, in this case, the transformer might be working under saturated conditions.

Author Contributions: Conceptualization, F.V.; methodology, F.V.; software, F.V.; validation, H.A. and F.Q.; formal analysis, F.Q.; investigation, F.V.; resources, H.A.; data curation, F.V. and F.Q.; writing—original draft preparation, F.V.; writing—review and editing, H.A. and F.Q.; visualization, F.V.; supervision, H.A.; project administration, H.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Behjat, V.; Shams, A.; Tamjidi, V. Characterization of power transformer electromagnetic forces affected by winding faults. *J. Oper. Autom. Power Eng.* **2018**, *6*, 40–49.
2. Moradnouri, A.; Vakilian, M.; Hekmati, A.; Fardmanesh, M. HTS transformer windings design using distributive ratios for minimization of short circuit forces. *J. Supercond. Nov. Magn.* **2019**, *32*, 151–158. [CrossRef]
3. Mikhak-Beyranvand, M.; Rezaeealam, B.; Faiz, J.; Rezaei-Zare, A. Impacts of ferroresonance and inrush current forces on transformer windings. *IET Electr. Power Appl.* **2019**, *13*, 914–921. [CrossRef]
4. Najafi, A.; Iskender, I. Electromagnetic force investigation on distribution transformer under unbalanced faults based on time stepping finite element methods. *Int. J. Electr. Power Energy Syst.* **2016**, *76*, 147–155. [CrossRef]
5. Ge, W.; Zhao, J.; Wang, Y. Analysis of the residual flux influence on inrush current and electromagnetic force in large power transformer. *J. Eng.* **2019**, *2019*, 2426–2429. [CrossRef]
6. Jin, J.M. *The Finite Element Method in Electromagnetics*; John Wiley & Sons: Hoboken, NJ, USA 2015
7. Peterson, A.F.; Ray, S.L.; Mittra, R.; Institute of Electrical and Electronics Engineers. *Computational Methods for Electromagnetics*; IEEE: New York, NY, USA, 1998; Volume 351.
8. IEC. *IEC 60075 Power Transformers-Part 5: Ability to Withstand Short Circuit*; International Electrotechnical Commission: Geneva, Switzerland, 2006.
9. Hong, K.; Lin, G. State classification of transformers using nonlinear dynamic analysis and Hidden Markov models. *Measurement* **2019**, *147*, 106851. [CrossRef]
10. Khan, A.; Ghorbanian, V.; Lowther, D. Deep learning for magnetic field estimation. *IEEE Trans. Magn.* **2019**, *55*, 1–4. [CrossRef]
11. Qi, S.; Wang, Y.; Li, Y.; Wu, X.; Ren, Q.; Ren, Y. 2D Electromagnetic Solver Based on Deep Learning Technique. *IEEE J. Multiscale Multiphy. Comput. Tech.* **2020**, *5*, 83–88. [CrossRef]
12. Diaz, G.I.; Fokoue-Nkoutche, A.; Nannicini, G.; Samulowitz, H. An effective algorithm for hyperparameter optimization of neural networks. *IBM J. Res. Dev.* **2017**, *61*, 9:1–9:11. [CrossRef]
13. Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the importance of initialization and momentum in deep learning. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1139–1147.
14. Hinton, G.; Srivastava, N.; Swersky, K. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. *Cited* **2012**, *14*, 2.
15. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
16. Laurent, T.; Brecht, J. The multilinear structure of ReLU networks. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; pp. 2908–2916.
17. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. *Proc. Icml. Citeseer* **2013**, *30*, 3.
18. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1026–1034. [CrossRef]
19. Xu, B.; Wang, N.; Chen, T.; Li, M. Empirical evaluation of rectified activations in convolutional network. *arXiv* **2015**, arXiv:1505.00853.
20. Meeker, D. *Finite Element Method Magnetics Version 4.2: User's Manual*; IEEE: New York, NY, USA, 2010.
21. Greenwood, A. *Electrical Transients in Power Systems*; John Wiley and Sons Inc.: New York, NY, USA, 1991.
22. Bertagnolli, G. *Short-Circuit Duty of POWER Transformers*; ABB: Zürich, Switzerland, 2013.
23. Vanderlinde, J. *Classical Electromagnetic Theory*; Fundamental Theories of Physics; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006.

24. Stratton, J.A. *Electromagnetic Theory*; IEEE Press Series on Electromagnetic Wave Theory; John Wiley & Sons: Hoboken, NJ, USA, 2007; Volume 33.
25. Jackson, J. *Classical Electrodynamics*; Wiley: Hoboken, NJ, USA, 2012.
26. Kulkarni, S.V.; Khaparde, S. *Transformer Engineering: Design and Practice*; CRC Press: Boca Raton, FL, USA, 2004; Volume 25.
27. Rylander, T.; Ingelström, P.; Bondeson, A. *Computational Electromagnetics*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
28. Flügge, W. *Stresses in Shells*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
29. Timoshenko, S. *Strength of Materials Part 1*; D. Van Nostrand Company Incorporated: New York, NY, USA, 1940

Hindawi
Journal of Electrical and Computer Engineering
Volume 2021, Article ID 4657696, 9 pages
https://doi.org/10.1155/2021/4657696

Research Article

Comparison of Machine Learning Algorithms for the Prediction of Mechanical Stress in Three-Phase Power Transformer Winding Conductors

Fausto Valencia, Hugo Arcos, and Franklin Quilumba

School of Electrical and Electronics Engineering, National Polytechnic School, Ladrón de Guevara 253, Quito 170517, Ecuador

Correspondence should be addressed to Fausto Valencia; fausto.valencia@epn.edu.ec

Received 5 October 2021; Accepted 13 November 2021; Published 26 November 2021

Academic Editor: Renato Procopio

This research compares four machine learning techniques: linear regression, support vector regression, random forests, and artificial neural networks, with regard to the determination of mechanical stress in power transformer winding conductors due to three-phase electrical faults. The accuracy compared with finite element results was evaluated for each model. The input data were the transient electrical fault currents of power system equivalents with impedances from low to high values. The output data were the mechanical stress in the conductors located in the middle of the winding. To simplify the design, only one hyperparameter was varied on each machine learning technique. The random forests technique had the most accurate results. The highest errors were found for low-stress values, mainly due to the high difference between maximum and minimum stresses, which made the training of the machine learning models difficult. In the end, an accurate model that could be used in the continuous monitoring of mechanical stress was obtained.

1. Introduction

The determination of electromagnetic forces due to electrical faults in power transformer windings is a crucial activity developed during the design stage [1–3]. This problem is generally solved through numerical methods because of the transformer geometry characteristics and the partial equations involved. Nowadays, the finite element method (FEM) is the most used technique for this task in [4–7].

Internally, FEM divides the medium into triangular or rectangular elements inside which the magnetic field is considered constant [8]. The more variation the field has, the more elements are needed for better accuracy. Since the field value is unknown at the beginning, the first attempt at the disposition of elements is performed according to the geometry of the problem. After that, the elements are iteratively divided and relocated depending on the results of the simulations [9].

Therefore, the whole process takes a long time to find the solution for one set of high and low voltage currents. This time consuming process is not a problem for power transformer design because the simulation is limited to a few cases corresponding to the worst case scenario. However, when conducting a continuous analysis of the effect of the forces or mechanical stress in transformer windings, such as a stress analysis [10, 11], using FEM is unfeasible mainly because the results are needed almost in real-time.

An alternative path to solve this problem is the method of images since it reduces the algebraic operations needed to find the forces. The method of images for magnetic fields was presented by Hammond [12] based on the works introduced by Thomson and Kelvin [13]. Kulkarni and Khaparde suggested its use for the determination of the transformer reactance based on the magnetic field [14]. Minhas utilized the method of images for the determination of forces and winding vibrations in a single phase transformer [15]. In another investigation, the forces acting in the transformer terminal are calculated through the use of the method of images [16]. The method of images assumes that the permeability of the core is constant, and for each winding, the magnetic material extends towards infinite. As an advantage, this method directly gives the forces without the need to calculate the magnetic field. However, it has not been applied to three-phase transformers; hence, it is unclear how the images must be located, mainly for the windings that embrace the middle limb of the core. Another issue is the constant permeability, which does not represent the problem when the transformer phases affect each other.

In a previous work, the application of artificial neural networks (ANNs) for the determination of mechanical stress has been analysed [17]. For training the ANNs, FEM simulations were used, and a good approximation was obtained. The advantage of the method is the time reduction in getting the results. Although FEM is still necessary to get the training data, it is no longer used for the rest of the power transformer lifetime after the model is obtained. The drawback is the difficulty of training the ANNs. They have many hyperparameters that affect the model accuracy [18]. Finding the best combination of hyperparameters could become cumbersome.

This research explores the use of four machine learning techniques for the determination of mechanical stress: linear regression (LR), support vector regression (SVR), random forests (RF), and ANN. The objective is to compare each technique's accuracy when varying only one hyperparameter, thus simplifying the model design and implementation.

The worst accuracy result found when using ANNs was chosen for comparison purposes, i.e., when the electrical fault faces a high impedance [17].

2. Method

There are two stages in the development of the machine learning tools:

(i) Generation of data: the mechanical stress is found for different electrical currents circulating through the windings

(ii) Training of the machine learning model: the pair electrical currents-mechanical stress is used to train the model

2.1. Mechanical Stress. The first step towards finding the mechanical stress is the determination of the magnetic induction $\mathbf{B}$ around the winding conductor. For this, recall that the magnetic vector potential $\mathbf{A}$ acting in a point obeys the Poisson equation (1), where $\mathbf{J}$ is the current density circulating through the point and μ is the permeability of the medium. Then, $\mathbf{B}$ can be found by equation (2).

$$\nabla^2 \mathbf{A} = -\mu \mathbf{J}, \quad (1)$$

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (2)$$

For the solution of equation (1), the software FEMM [19], which implements FEM, was used. Table 1 presents the transformer characteristics, while Table 2 presents the geometry entered in FEMM.

Table 1: Power transformer technical characteristics.

Variable	Value	Units
Power	400	MVA
High voltage	230	kV
Low voltage	138	kV
Frequency	60	Hz
Group of connection	Yy0	
Impedance (own base)	33.61	%
Number of low voltage disks	100	U
Number of high voltage disks	105	U

Table 2: Power transformer geometry.

Variable	Value (m)
Core diameter	0.9582
Core window height	3.034
Limb-limb separation	1.853
Low voltage winding inside diameter	1.108
Low voltage winding outside diameter	1.309
High voltage winding inside diameter	1.429
High voltage winding outside diameter	1.625
Low voltage disk height	0.0175
High voltage disk height	0.0136
Spacer block (located between disks)	0.010

The electrical currents considered correspond to electromagnetic transients of electrical faults that face impedances from $z = 1 + j15\Omega$ to $z = 5 + j80\Omega$. A total of 7839 and 795 training and validation cases were generated, respectively. They belong to faults starting when Phase A, Phase B, and Phase C cross zero. Equation (3) shows the formulation for the low voltage winding transient current, I_{LV}, where ω is the angular frequency of the system, t is the time, ϕ is the angle representing the fault starting point, θ is the angle between phases (120° in a three-phase balanced system), and $\lambda = \omega r/x_l$, where r and x_l are the equivalent resistance and inductive reactance seen by the fault. The high voltage current is determined with the ratio of the respective windings.

$$\begin{aligned} i_a(t) &= I_{\mathrm{LV}} \cdot [\sin(\omega t - \phi) + \sin(\phi) \cdot \exp(-\lambda t)], \\ i_b(t) &= I_{\mathrm{LV}} \cdot [\sin(\omega t - \theta - \phi) + \sin(\theta + \phi) \cdot \exp(-\lambda t)], \\ i_c(t) &= I_{\mathrm{LV}} \cdot [\sin(\omega t + \theta - \phi) + \sin(-\theta + \phi) \cdot \exp(-\lambda t)]. \end{aligned} \quad (3)$$

Once $\mathbf{B}$ was found, the force per volume unit on the conductor was calculated by the following equation:

$$\mathbf{f} = \mathbf{J} \times \mathbf{B}. \quad (4)$$

The radial force in the middle conductors of the windings represents the highest value [17]. Moreover, this force component is the only one that could affect the conductor continuously because the winding internal structure nullifies the axial force effects. Therefore, a simplified model of the winding conductor can be used to determine the stress, where the conductor is modelled as a ring with radius $\mathrm{Rad}_{\mathrm{ring}}$ and cross-sectional area S_c. Thus, the force P normal to the section of the conductor subjected to a radial force per length F_r and the stress σ are calculated by the following equations, respectively:

$$P = F_r \cdot \mathrm{Rad}_{\mathrm{ring}}, \tag{5}$$

$$\sigma = \frac{P}{S_c}. \tag{6}$$

An additional simplification was to take the whole disk as a conductor. In that way, the FEM discretization was reduced and so was the simulation time to create the training and validation samples.

2.2. Machine Learning Models. The design and implementation of the machine learning models were performed in the Python library scikit-learn 0.24.2. For each technique, Table 3 shows the scikit-learn libraries employed and the modified hyperparameters. Two libraries are needed in LR because the LinearRegression library can model only a straight line. With PolynomialFeatures, a higher degree polynomial, including the products of all the input variables, can be used for the model.

Previous to training, the input and output data were standardized (see equation (7)) so that all the values are in the range from zero to one.

$$\begin{aligned} x_{\mathrm{std}} &= \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \\ y_{\mathrm{std}} &= \frac{y - y_{\min}}{y_{\max} - y_{\min}}. \end{aligned} \tag{7}$$

The most suitable hyperparameter value is determined by evaluating the coefficient of determination R^2, defined by equation (8), where y_{true} is the true value of the output variable, y_{pred} is the output value predicted by the model, and $\overline{y}_{\mathrm{true}}$ is the mean value of the output variable:

$$\begin{aligned} R^2 &= 1 - \frac{u}{v}, \\ u &= \sum \left(y_{\mathrm{true}} - y_{\mathrm{pred}}\right)^2, \\ v &= \sum \left(y_{\mathrm{true}} - y_{\mathrm{pred}}\right)^2. \end{aligned} \tag{8}$$

3. Results

3.1. Hyperparameters for the Highest Accuracy. For LR, the polynomial degree was varied from one to nine. Figure 1 shows the accuracy variation. The validation test had the best behaviour between degrees five and eight. The polynomial of degree five was considered for the sake of simplicity.

Figure 2 shows the error for SVR, when the hyperparameter C varied from 1 to 400. The minimum error value is located at $C = 250$. The training error was nearly constant for high values of C, but the validation error slightly increased, which could indicate overfitting.

In RF, the number of trees was modified (see Figure 3). The error decreased until a number of 100 trees. After that, the error was constant both in the training data and in the validation test. The maximum number of trees was set at 1000 because the training time was too long after that. Moreover, no improvement was seen from 100 to 1000 trees. Hence, 100 trees were chosen for the model.

Figure 4 shows the variation of the error for a different number of units in the hidden layer for ANN. Only one hidden layer was considered because that architecture has the best accuracy for the ANN-based models [17] in the mechanical stress problem. The number of units had a small effect on the training data error. However, the error decreased almost uniformly in the validation test, reaching a minimum at 1625 units. This machine learning technique accentuated the different behaviour between the training data and the validation test. It was clear that increasing the number of units in the layer did not decrease the model's error.

A comparison of the lowest error is seen in Figure 5. RF had the best behaviour, LR and SVR have similar values, and ANN has the highest error.

3.2. Error Comparison for a Transient Fault Current. The highest error of the ANN model in the determination of the mechanical stress was found for electrical faults with the highest impedance in a previous work. In particular, the case of $r = 1\Omega$ and $x_l = 80\Omega$ was seen as having the worst accuracy [17]. Therefore, that case was analysed in this research for the comparison of the machine learning techniques.

Figure 6 shows the mean absolute percentage error (MAPE) for the determination of stress in the low voltage winding conductors. LR and RF had the lowest MAPE, whereas it was the highest for SVR and ANN models. In general, Phase B has the highest error.

In Figure 7, the MAPE for the mechanical stress in the high voltage winding is shown. LR and RF still present the lowest error. The SVR model has the highest error in Phase B.

To clarify where the differences between the machine learning and FEM models are located, Figures 8 to 11 present the results for the worst cases of each model. All of them belong to the high voltage winding simulations. In general, there is a high difference in low values of the mechanical stress. This may be due to the lower importance that these cases have for the machine learning tool. For the practical use of the model, this behaviour is not an issue because the low mechanical stresses have little effect on the deterioration of the winding conductor.

The ANN model has a particular behaviour, as shown in Figure 11. The error prevails even when close to the steady-state of the transient. The error is high for the first cycles with high and low stresses. For practical analysis, the ANN model has the worst outcome.

4. Discussion

The mechanical stress and the electromagnetic forces in windings have a strong dependence on the electrical currents. For that reason, although the validity tests have cases outside those used in the training process, each machine learning

Table 3: Scikit-learn libraries and the hyperparameters for each machine learning techniques.

Machine learning technique	Library	Hyperparameter
Linear regression	Preprocessing.PolynomialFeatures Linear_model.LinearRegression	Degree
Support vector regression	Svm.SVR	C
Random forests	Ensemble.RandomForestRegressor	n_estimators
Artificial neural network	Neural_network.MLPRegressor	Hidden_layer_sizes

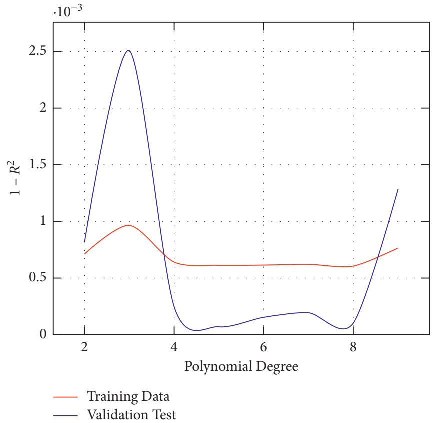


Figure 1: Accuracy variation according to the polynomial degree.

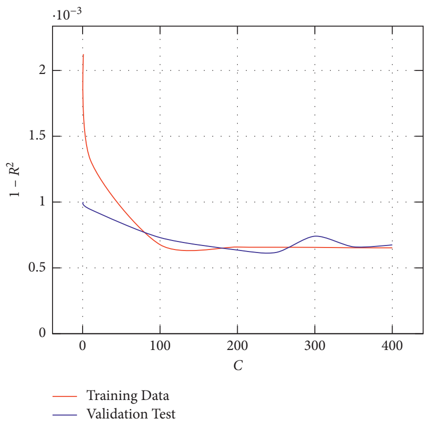


Figure 2: Accuracy variation according to C in support vector regression.

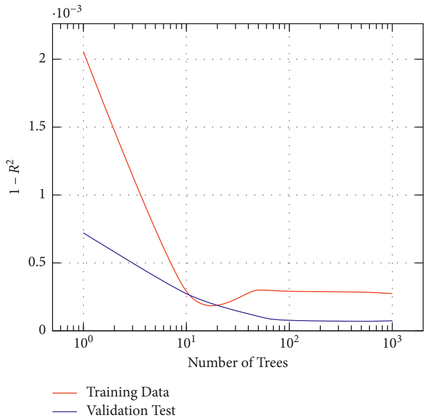


Figure 3: Accuracy variation according to the number of trees.

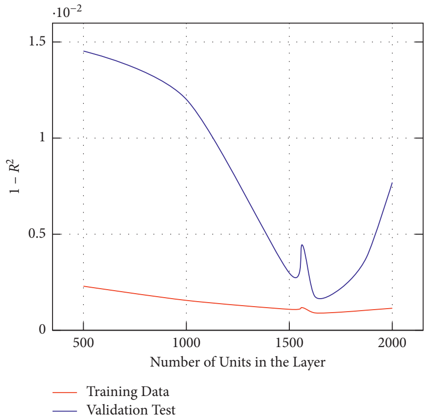


Figure 4: Accuracy variation according to the number of units in the layer.

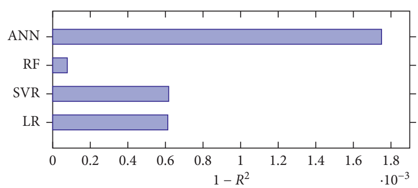


Figure 5: Comparison of the machine learning techniques according to the lowest $1 - R^2$ value.

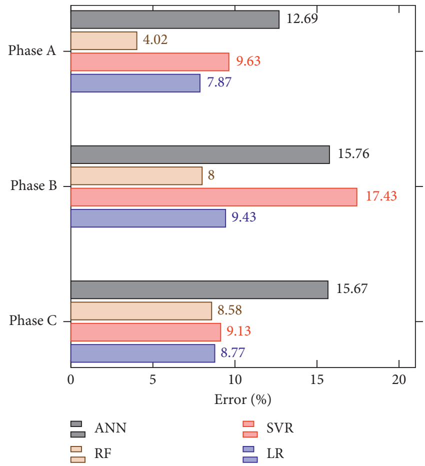


Figure 6: Error comparison for the low voltage windings.

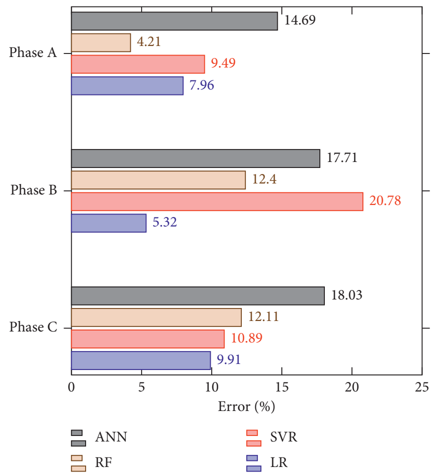


Figure 7: Error comparison for the high voltage windings.

model has presented high accuracy, showing a low level of overfitting. This means that the behaviour seen by the model during training is much related to the behaviour of the validation set of data. Therefore, except for the ANN model, the rest of the validation cases have given even lower errors than those of the training simulations. As is usual in machine learning models, overfitting is possible when the model is more complex, e.g., when augmenting the polynomial degree in LR or the value of hyperparameter *C* in SVR.

In the variation of hyperparameters, the models have a homogeneous behaviour in the validation test. When the model is simple, the error is high and lowers when the model increases its complexity. The error reaches a minimum at some value and then begins to rise, showing signs of overfitting. The exception to this behaviour is ANN, which has skipped in the flow of error when the model is more complex. This characteristic makes it difficult to find the optimal point in the model design. The multiple relations

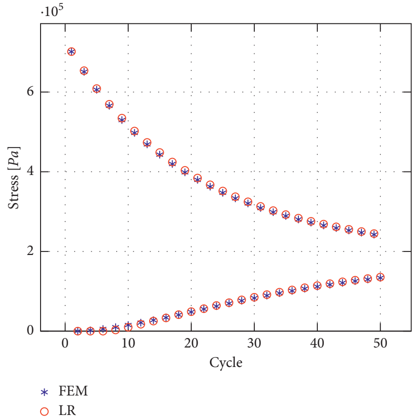


Figure 8: FEM and linear regression results for Phase C of the high voltage winding.

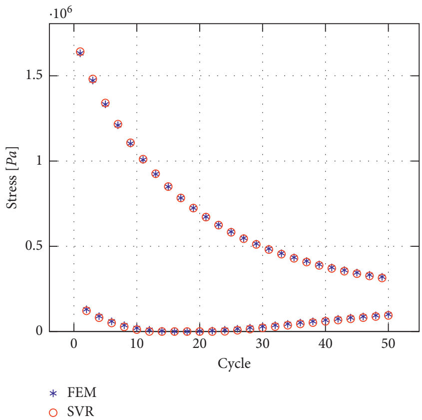


Figure 9: FEM and support vector regression results for Phase B of the high voltage winding.

that exist inside the ANN model, with all the weights and activation functions, might be the reason for that unusual behaviour.

The RF model has the best accuracy. This result is seen in the particular case of high impedance simulation for the low voltage winding as well as in the validation test results. The chosen model has 100 trees.

The second technique with the lowest error is LR, which was modelled with a five-degree polynomial. As a result, cross-terms between the six input currents are included in the internal structure of the model. This input data interconnection has allowed the model to represent the nonlinearities of the phenomenon and mainly the influence of the different input currents. It opens the possibility of

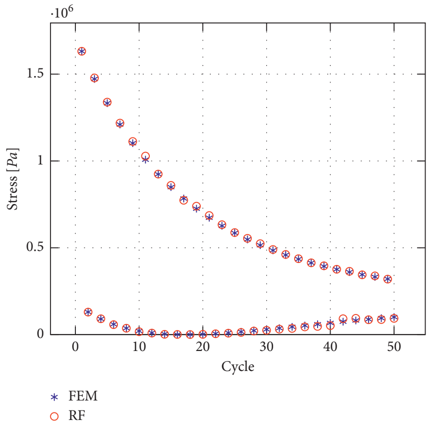


FIGURE 10: FEM and random forest regression results for Phase B of the high voltage winding.

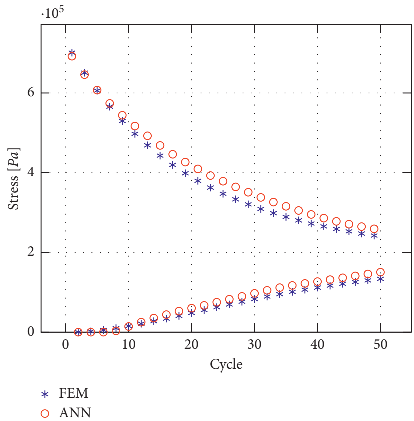


FIGURE 11: FEM and artificial neural network regression results for Phase B of the high voltage winding.

improving the predictions for other machine learning models if some combination of the winding currents is included as input data.

5. Conclusions

Four machine learning techniques have been compared with regard to accuracy. Only one hyperparameter has been varied for each technique so that the design process is simplified.

This study demonstrates that the appropriate machine learning technique improves the accuracy of the model. For the determination of mechanical stress in transformer windings, the random forest proved to be the best model, even for the high impedance electrical fault, which was the worst case in our previous research presented in [17].

All the models present the highest error in low values of stress. There is a high difference in the stress values during the transient period. This affects the behaviour of the machine learning models even though the input data are standardized before the training process. Nonetheless, this is not such an issue since the mechanical stress with the highest value is the one that could cause damage to the conductor.

By finding the right machine learning technique, the utilities might have a powerful tool that allows the continuous monitoring of the mechanical stress behaviour.

Thus, in the future, policies of fatigue analysis to determine the deterioration of the winding conductor could be established.

Data Availability

The CSV files with the training and validation data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] International Electrotechnical Commission 60075, *Power Transformers-Part 5: Ability to Withstand Short Circuit*, International Electrotechnical Commission, Geneva, Switzerland, 2006.

[2] G. Bertagnolli and AAB Management Services, *The ABB Approach to Short-Circuit Duty of Power Transformers*, AAB Management Services, Zurich, Switzerland, 2007.

[3] S. Gopalakrishna, K. Kumar, B. George, and V. Jayashankar, "Design margin for short circuit withstand capability in large power transformers," in *Proceedings of the 2007 International Power Engineering Conference*, pp. 1262–1267, Singapore, December 2007.

[4] A. C. De Azevedo, I. Rezende, A. C. Delaiba, J. C. De Oliveira, B. C. Carvalho, and S. Bronzeado Herivelto De, "Investigation of transformer electromagnetic forces caused by external faults using FEM," in *Proceedings of the 2006 IEEE/PES Transmission & Distribution Conference and Exposition: Latin America*, pp. 1–6, Caracas, VE, USA, August 2006.

[5] H.-Mo Hyun-Mo Ahn, Ji-Y. Ji-Yeon Lee, J.-K. Yeon-Ho Oh, S.-Y. Jung, and S.-C. Hahn, "Finite-element analysis of short-circuit electromagnetic force in power transformer," *IEEE Transactions on Industry Applications*, vol. 47, no. 3, pp. 1267–1272, 2011.

[6] H.-M. Ahn, B.-J. Lee, C.-J. Kim, H.-K. Shin, and S.-C. Hahn, "Finite element modeling of power transformer for short circuit electromagnetic force analysis," in *Proceedings of the 2012 15th International Conference on Electrical Machines and Systems (ICEMS)*, pp. 1–4, Saporo, Japan, October 2012.

[7] C. Yan, Z. Hao, S. Zhang, B. Zhang, T. Zheng, and Z. Li, "Computation and analysis of power transformer winding damage due to short circuit fault based on 3-D finite element method," *International Journal of Applied Electromagnetics and Mechanics*, vol. 51, no. 4, pp. 405–418, 2016.

[8] J.-M. Jin, *The Finite Element Method in Electromagnetics*, John Wiley & Sons, New Jersey, NJ, USA, 2015.

[9] T. Rylander, P. Ingelstrom, and B. Anders, *Computational Electromagnetics*, Springer Science & Business Media, New York, NY, USA, 2012.

[10] H. Zhang, B. Yang, W. Xu et al., "Dynamic deformation analysis of power transformer windings in short-circuit fault by FEM," *IEEE Transactions on Applied Superconductivity*, vol. 24, no. 3, pp. 1–4, 2014.

[11] J. F. Araujo, E. G. Costa, F. L. M. Andrade, A. D. Germano, and T. V. Ferreira, "Methodology to evaluate the electromechanical effects of electromagnetic forces on conductive materials in transformer windings using the von mises and fatigue criteria," *IEEE Transactions on Power Delivery*, vol. 31, no. 5, pp. 2206–2214, 2016.

[12] P. Hammond, "Electric and magnetic images," *Proceedings of the IEEE Part C: Monographs*, vol. 107, no. 12, pp. 306–313, 1960.

[13] W. Thomson and W. T. B. Kelvin, *Reprint of Papers on Electrostatics and Magnetism*, Macmillan, London, UK, 1872.

[14] S. V. Kulkarni and S. A. Khaparde, *Transformer Engineering: Design, Technology, and Diagnostics*, CRC Press, Boca Raton, FL, USA, 2017.

[15] M. S. A. Minhas, *Dynamic behaviour of transformer winding under short circuits*, PhD thesis, University of the Witwatersrand, Johannesburg, South Africa, 2007.

[16] M. Moghaddami, A. Moghadasi, and A. I. Sarwat, "An algorithm for fast calculation of short circuit forces in high current busbars of electric arc furnace transformers based on method of images," *Electric Power Systems Research*, vol. 136, pp. 173–180, 2016.

[17] F. Valencia, H. Arcos, and Q. Franklin, "Prediction of stress in power transformer winding conductors using artificial neural networks: hyperparameter analysis," *Energies*, vol. 14, no. 14, Article ID 4242, 2021.

[18] D. Choi, C. J. Shallue, Z. Nado, J. Lee, C. J. Maddison, and G. E. Dahl, "On empirical comparisons of optimizers for deep learning," 2019, https://arxiv.org/abs/1910.05446.

[19] D. Meeker, *Finite Element Method Magnetics Version 4.2: User's Manual*, IEEE, Piscataway, NJ, USA, 2010.