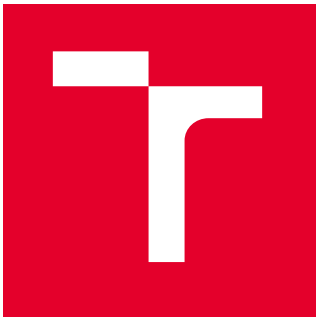


BRNO UNIVERSITY OF TECHNOLOGY

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF POWER ELECTRICAL AND ELECTRONIC ENGINEERING

ÚSTAV VÝKONOVÉ ELEKTROTECHNIKY A ELEKTRONIKY

INNOVATIVE TRANSFORMER DESIGN OPTIMIZATION COMBINED WITH MACHINE LEARNING LOSS ESTIMATION

INOVATIVNÍ OPTIMALIZACE NÁVRHU TRANSFORMÁTORŮ V KOMBINACI S VÝPOČTEM ZTRÁT POMOCÍ STROJOVÉHO UČENÍ

DOCTORAL THESIS
DIZERTAČNÍ PRÁCE

AUTHOR — **Ing. Miroslav Mrajca**
AUTOR PRÁCE

SUPERVISOR — **doc. Ing. Jan Bárta, Ph.D.**
ŠKOLITEL

BRNO 2025

ABSTRACT

This doctoral thesis presents a fully automated, data-driven methodology for optimizing transformer design to meet stringent efficiency standards, such as the EcoDesign requirements of the European Union, while controlling material costs. The approach integrates simulated annealing optimization with conventional analytical calculation models, systematically generating technically compliant and cost-effective designs. Gaussian process regression is employed as a surrogate model to accurately predict no-load losses, improve design margins, and reduce material use in all transformer components. A review of industrial design practices highlighted the continued reliance on empirical corrections, simplified models, and designer expertise, underscoring the need for automated, optimization-driven workflows. The methodology accounts for both deterministic and stochastic factors that influence no-load losses, providing a foundation for reliable data-driven optimization. Simulated annealing efficiently navigates complex, non-linear, and constrained design spaces, while Gaussian process regression reduces computational effort and enhances predictive accuracy. The workflow was validated on multiple transformer ratings (50 kVA, 630 kVA, 1,600 kVA), demonstrating the rapid generation of feasible and cost-effective designs with reduced dependence on the designer's intuition. Surrogate modeling using Gaussian process regression on 4,503 transformers in 322 designs enabled accurate prediction of losses, including uncertainty quantification, outperforming alternative machine learning methods. In general, the thesis establishes a practical, repeatable, and scalable framework for automated transformer design, combining classical analytical techniques, robust optimization, and advanced surrogate modeling, supporting high-efficiency, cost-effective transformers and paving the way for future integration of optimization and machine learning in industrial engineering.

KEYWORDS

Gaussian process regression, no-load losses, transformer core losses, machine learning, surrogate modeling, simulated annealing, transformer design optimization, automation of manual calculations, manufacture cost minimization, EcoDesign

ABSTRAKT

Tato dizertační práce představuje plně automatizovanou metodiku pro optimalizaci návrhu transformátorů, zpřesněnou změřenými daty, která umožňuje splnění přísných požadavků na účinnost, například podle směrnice EkoDesign Evropské unie, s důrazem na minimalizaci nákladů na materiál. Postup využívá optimalizaci metodou simulovaného žíhání s využitím známých analytických výpočetních modelů a systematicky hledá návrhy, které splňují technické požadavky a zároveň jsou hospodárné. Výpočet ztrát naprázdno je proveden pomocí náhradního modelu založeného na regresi Gaussovskými procesy, což snižuje bezpečnostní rezervy a umožňuje úsporu materiálu ve všech částech transformátoru. Současné výpočetní postupy v průmyslové praxi stále často spoléhají na empirické vzorce, zjednodušené modely a zkušenosti návrhářů, což zdůrazňuje potřebu automatizovaného optimalizačního přístupu. Navržená metodika zohledňuje jak deterministické, tak stochastické faktory ovlivňující ztráty naprázdno a poskytuje základ pro spolehlivou optimalizaci založenou na změřených datech. Simulované žíhání efektivně prohledává komplexní nelineární prostor s ohledem na konstrukční omezení, zatímco regrese Gaussovskými procesy zvyšuje přesnost výpočtu bez navýšení výpočetní náročnosti. Nově představený postup byl ověřen na několika typech transformátorů (50 kVA, 630 kVA a 1 600 kVA) a výsledky potvrzují rychlé nalezení návrhu splňujícího všechny požadavky při minimalizaci nákladů, nezávisle na intuici návrháře. Náhradní modelování pomocí regrese Gaussovskými procesy, založené na 4 503 transformátorech zahrnutých v 322 návrzích, umožňuje přesnější výpočet ztrát včetně odhadu nejistoty a překonává tak alternativní metody strojového učení. Obecně práce vytváří praktickou, ověřitelnou a přizpůsobitelnou metodiku pro automatizovaný návrh transformátorů, která kombinuje klasické analytické výpočty, robustní optimalizaci a pokročilé náhradní modelování, při minimalizaci nákladů a zajištění vysoké účinnosti. Připravuje tak půdu pro další integraci optimalizace a strojového učení do průmyslového inženýrství.

KLÍČOVÁ SLOVA

Regrese Gaussovskými procesy, ztráty naprázdno, ztráty v jádře transformátoru, strojové učení, náhradní modelování, simulované žíhání, optimalizace návrhu transformátorů, automatizace manuálních výpočtů, minimalizace výrobních nákladů, EkoDesign

MRAJCA, Miroslav. *Innovative Transformer Design Optimization Combined with Machine Learning Loss Estimation*. Doctoral thesis. Brno: Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Power Electrical and Electronic Engineering, 2025. Advised by doc. Ing. Jan Bárta, Ph.D.

Author's Declaration

Author:	Ing. Miroslav Mrajca
Author's ID:	195393
Paper type:	Doctoral thesis
Academic year:	2024/25
Topic:	Innovative Transformer Design Optimization Combined with Machine Learning Loss Estimation

I declare that I have written this paper independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the paper and listed in the comprehensive bibliography at the end of the paper.

As the author, I furthermore declare that, with respect to the creation of this paper, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll. of the Czech Republic, Section 2, Head VI, Part 4.

Brno ....

author's signature*

*The author signs only in the printed version.

ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to my supervisor, Associate Professor Jan Bárta, for his expert guidance, invaluable insights throughout the development of this thesis, and his significant contributions to my most important published scientific papers, as well as for securing their financial support.

I express my sincere gratitude *in memoriam* to Associate Professor Čestmír Ondrůšek for inspiring me to pursue doctoral studies and guiding the direction of this research, and to Associate Professor Radoslav Cipín for his valuable guidance as interim supervisor.

My special thanks go to my colleague Vladimír Bílek for his collaboration on scientific publications and for sharing his expertise on our joint projects. His contributions were instrumental and this thesis would not have been possible without his support.

I am deeply grateful to my nephew, Petr Šilling, a doctoral student at the Faculty of Informatics, Brno University of Technology (BUT), for his collaboration on key developments in optimization, which formed a central focus of this work.

Many thanks are also due to my colleagues in the department for their valuable assistance and advice in completing projects related to my doctoral studies.

I would like to extend my sincere thanks to the transformer manufacturer Elpro-Energo Transformers s.r.o. for providing essential data and valuable engineering experience, which served as a foundation for this research, and for their support throughout the work.

I also wish to express my special gratitude to Lemi Trafo JSC in Bulgaria for their hospitality during my international traineeship and for offering me the opportunity to further develop my expertise, not only in the field of transformer design.

This research was conducted at the Centre for Research and Utilization of Renewable Energy (CVVOZE) with the support of the Ministry of Education, Youth and Sports of the Czech Republic under institutional support and the BUT specific research programme (project No. FEKT-S-23-8430). It was further funded by the project Optimization of Power Transformer Design Utilizing Simulated Annealing Algorithm within the BUT interfaculty specific research programme (project No. FEKT/FIT-J-24-8611), and by the Linz Center of Mechatronics GmbH (LCM) K2 Center for Symbiotic Mechatronics through the Austrian COMET-K2 funding program, co-financed by the Austrian and Upper Austrian governments. The author gratefully acknowledges their support.

I acknowledge the use of artificial intelligence tools, including ChatGPT and DeepL, for grammar checking, sentence rephrasing, and improving clarity. All intellectual content, ideas, and analyses in this thesis are my own.

Finally, I am deeply grateful to my dear parents, family and close friends for their unwavering encouragement, patience, and support throughout my doctoral studies.

Contents

List of Figures

List of Tables

1 Introduction

A transformer is a fundamental component in the transmission and distribution of electrical energy. It plays a critical role throughout the entire electricity supply chain by enabling the sequential conversion of voltage levels—from the generation source to the point of consumption. Initially, transformers are used to increase voltage levels for efficient long-distance transmission, and subsequently to reduce them for safe and practical end-use. They are deployed at multiple stages, including power plants, high-voltage transmission substations, sub-transmission systems, and distribution networks close to consumers [1].

Although they are among the most energy efficient devices—typically exhibiting efficiencies exceeding 99%—transformers are installed in vast numbers worldwide and operate continuously. On average, electrical power passes through four to five transformers before reaching the end user. Consequently, even small inefficiencies accumulate significantly, with transformer losses accounting for approximately 5% of global electricity consumption, as reported in [2].

Therefore, there is an increasing demand to improve transformer efficiency globally, particularly in the European Union (EU), where strict limits on load and no-load losses are mandated for both distribution and power transformers, as outlined in [3]. These standards are commonly referred to as the *EcoDesign* Regulations. The first tier came into force on 1st July 2015, and the requirements were further tightened in the second tier, effective from 1st July 2021, which remains in effect as of now. However, additional reductions in allowable losses are already under discussion, as indicated in [4].

According to [5], these regulations have had a significant impact on transformer prices within the European market. Loss reductions can only be achieved by increasing the quantity of core and winding materials, such as aluminium, copper, and, in the case of transformer cores, higher-grade electrical steel. These key materials constitute a major portion of the transformer's final cost. In addition to higher prices, the stricter efficiency requirements have led to substantial increases in both the size and weight of transformers, occasionally resulting in difficulties when replacing older units in installations originally designed for more compact equipment. Furthermore, the increase in physical dimensions leads to greater consumption of other essential materials used in transformer manufacturing, such as transformer oil and steel tank with its cover (in the case of oil-immersed units), steel clamping structures, and insulation paper (or cast resin in dry-type units).

Hence, based on the stated facts, there is increasing pressure on transformer design to be as cost-effective as possible while still complying with the EcoDesign requirements. Given the current state of most distribution and power transformer

manufacturers, two primary areas for improvement are identified. The first is the optimization of the design process itself, aiming to minimize production costs while maintaining the required characteristics of transformer performance. The second is the accurate estimation of transformer efficiency during the design phase. This precision enables the reduction of design margins reserved for losses, as the EcoDesign regulations impose zero tolerance for exceeding specified loss limits during routine testing of transformer units. This approach helps to avoid unnecessary oversizing of conductor and core cross-sections. This doctoral thesis proposes a solution to address these challenges, and the objectives of the thesis are accordingly defined.

1.1 Objectives and Contributions of the Thesis

This doctoral dissertation presents research and development work carried out between 2021 and 2025, focusing on optimizing the transformer design process in a real manufacturing environment with the overarching goal of achieving highly cost-effective transformer designs. To this end, two main objectives are pursued. The first objective is to develop a fully automated transformer design workflow capable of operating without human intervention. This workflow integrates an optimization algorithm, in particalar simulated annealing (SA), with analytical calculation models to systematically generate transformer configurations that meet all technical and regulatory requirements, adhere to manufacturing constraints, and minimize material costs. This approach replaces traditional manual calculations based on experience with a rigorous, repeatable, and scalable optimization framework that provides a solid foundation for modern transformer design.

A second, closely related objective is to improve the accuracy of transformer no-load losses modeling through the application of machine learning techniques combined with experimentally measured data, in particular Gaussian process regression (GPR). Conventional analytical methods often necessitate oversizing of the magnetic core to ensure compliance with strict no-load loss limits, as no tolerance is permitted for exceeding specified thresholds. However, oversizing the core significantly affects every aspect of transformer design and contributes to increased production costs. By developing a more precise prediction method for no-load losses, this dissertation aims to eliminate unnecessary design margins, thereby enabling both core and overall transformer optimization. Together, these two research directions provide a comprehensive strategy for achieving high-performance and cost-effective transformer designs that meet current industrial requirements and regulatory standards.

The structure and contributions with the two main objectives corresponding to this dissertation are summarized below.

- **Chapter 2**—The **first objective** of the thesis is to map the current state of transformer design practice and provide a structured overview of the design process, emphasizing the interplay of analytical methods, empirical corrections, and simulation tools. This chapter is to formalize the design workflow, define key variables and constraints, and lay the foundation for the optimization framework developed in subsequent chapters.
- **Chapter 3**—The **second objective** is to review transformer no-load loss mechanisms and the state-of-the-art in conventional calculation methods, with emphasis on additional losses caused by manufacturing tolerances, material variations, and assembly techniques. Its purpose is to categorize predictable and variable loss components and to justify the adoption of advanced modeling and data-driven approaches to improve prediction accuracy and achieve cost-effective, high-efficiency transformer designs.
- **Chapter 4**—The **third objective** is to present the theoretical foundations of the optimization and machine learning techniques applied to automate and improve transformer design. Its purpose is to show how SA combined with GPR surrogate models can efficiently manage interdependent design variables and constraints, reduce computational effort, and enable fully automated and technically compliant transformer configurations with minimal material costs.
- **Chapter 5**—The **fourth objective** is to present and validate a fully automated transformer design workflow based on the SA optimization algorithm. Its purpose is to demonstrate how the algorithm can systematically explore the design space, handle constraints and cost evaluation, and reliably generate feasible, cost-effective transformer designs, outperforming conventional manual methods under real-world manufacturing conditions.
- **Chapter 6**—The **fifth objective** of the thesis is to develop and validate measured data-driven surrogate models to estimate transformer no-load losses, with a particular focus on GPR. Its purpose is to demonstrate how data-driven models, built from extensive experimental measurements, can accurately predict average losses and their variability, quantify uncertainty, and provide a robust and practical tool for supporting cost-effective and reliable transformer design.

Some of the results of the research have been published in scientific papers, which are listed in the following section. The thesis is closed with a **Conclusion**, where the acquired results are discussed, along with suggestions for improvement of the developed methodologies.

1.2 List of Authors Publications

The most significant contributions of this doctoral thesis, along with related research topics, have been published in the following scientific papers:

1. **Submitted for peer review in a scientific journal**
   - [6] M. Mrajca, V. Bilek, and J. Barta, "Transformer no-load losses estimation using machine learning from measured data", *IEEE Transactions on Industry Applications*, 2025
2. **Conference papers**
   - [7] M. Mrajca and P. Silling, "Advanced transformer design optimization via simulated annealing", in *Proceedings I of the 31st Conference STUDENT EEICT 2025*, Brno, Czech Republic: Brno University of Technology, Faculty of Electrical Engineering and Communication, 2025, pp. 311–315, ISBN: 978-80-214-6321-9. [Online]. Available: `https://www.eeict.cz/eeict_download/archiv/sborniky/EEICT_2025_sbornik_1.pdf`
   - [8] M. Mrajca, V. Bilek, J. Barta, *et al.*, "Transformer no-load losses calculation using machine learning", in *2024 International Conference on Electrical Machines (ICEM)*, Torino, Italy: IEEE, 2024, pp. 1–7, ISBN: 979-8-3503-7060-7. DOI: `10.1109/ICEM60801.2024.10700285`. [Online]. Available: `https://ieeexplore.ieee.org/document/10700285/`
   - [9] M. Mrajca, "Eddy currents additional losses comparison in oil distribution transformer windings", in *Proceedings I of the 29th Conference STUDENT EEICT 2023*, Brno, 2023, pp. 327–331, ISBN: 978-80-214-6153-6. [Online]. Available: `https://www.eeict.cz/eeict_download/archiv/sborniky/EEICT_2023_sbornik_1.pdf`
   - [10] M. Mrajca, "Advanced turowski's method for eddy loss calculation in transformer foil winding", in *Proceedings I of the 29th Conference STUDENT EEICT 2023*, Brno, 2023, pp. 332–336, ISBN: 978-80-214-6153-6. [Online]. Available: `https://www.eeict.cz/eeict_download/archiv/sborniky/EEICT_2023_sbornik_1.pdf`
   - [11] M. Mrajca and R. Cipín, "Axial channel effect on leakage reactance in split-winding transformer", in *Proceedings I of the 28th Conference STUDENT EEICT 2022*, Brno, 2022, pp. 331–335, ISBN: 978-80-214-6029-4. [Online]. Available: `https://www.eeict.cz/eeict_download/archiv/sborniky/EEICT_2022_sbornik_1_v2.pdf`

1.3 Summary of Chapter 1

This chapter introduced the fundamental role of transformers in the transmission and distribution of electrical energy, emphasizing the importance of efficiency and the impact of losses on global electricity consumption. It highlighted the challenges posed by strict efficiency regulations, such as the European Union EcoDesign standards, which underscore the need for optimized transformer designs. The main objectives of this thesis were presented: (i) the development of a fully automated transformer design workflow that integrates simulated annealing with analytical models to systematically generate cost-effective and compliant designs, and (ii) the improvement of no-load loss prediction using Gaussian process regression to reduce unnecessary design margins and material usage. Finally, the chapter outlined the structure of the thesis, demonstrating how the proposed methodologies collectively provide a rigorous, repeatable, and practical framework for modern transformer design that balances performance, efficiency, and minimal costs.

2 Transformer Design Process

This chapter provides an overview of the fundamental principles underlying the transformer design process, the mathematical models employed, and a general description of how this process is implemented in industrial transformer manufacturing. However, due to the complexity and multi-disciplinary nature of transformer design, a comprehensive, step-by-step explanation cannot be fully presented within the scope of this chapter. A detailed discussion of the complete design methodology, including the full analytical procedure for a three-phase, hermetically sealed oil-immersed distribution transformer, with explicit numerical calculations and manufacturing details, is provided in the author's master's thesis [12]. Understanding the manufacturing process is particularly important, as the technology used in core fabrication, winding, insulation, and assembly directly defines both the capabilities and limitations of practical transformer design.

The methods and analytical approaches proposed in this thesis are equally applicable to dry-type cast-resin transformers, as their active part (meaning core, clamping beams and windings) design closely resembles that of oil-immersed units. The primary differences lie in the design of the high-voltage winding and the thermal management system, which, in the case of dry-type transformers, relies solely on air cooling without the presence of oil and a sealed tank. A comprehensive explanation of the design methodology and manufacturing practices for dry-type cast-resin transformers is available in the author's bachelor's thesis [13].

2.1 Transformer Design Process in Practice

Most transformer manufacturers worldwide, including prominent companies such as SGB-SMIT GmbH [14] from Germany, Swiss-Swedish ABB Ltd. [15], Končar – Elektroindustrija d.d. [16] from Croatia, Lemi Trafo JSC [17] from Bulgaria, and Elpro-Energo Transformers s.r.o. [18] from the Czech Republic, as well as numerous other global producers, still predominantly rely on manual calculations or, at best, semi-automated software tools, such as those presented in [19]. Another commonly used tool in these companies is the brute-force calculator, which attempts to evaluate all possible combinations of design variables. In practice, these tools are typically limited to a small subset of the most critical variables due to computational constraints. The primary purpose of such calculators is to provide a preliminary starting point for the designer, who must subsequently refine and adjust the results to account for all relevant factors, including practical manufacturing limitations and secondary design considerations.

In these traditional design workflows, the human engineer plays the central role, iteratively exploring design variables in a mathematical model to arrive at a satisfactory solution. Consequently, the designer remains the primary decision-maker, effectively managing and steering the entire design process. The quality and efficiency of the final transformer design are heavily dependent on the expertise of the designer and the time allocated to explore different design alternatives. In practice, a single transformer specification can be satisfied by multiple valid configurations, each meeting performance and safety requirements. The designer's challenge is to identify and evaluate these configurations, ultimately selecting the most suitable option for manufacturing. Several perspectives can guide this selection of an "optimal" design, including:

- minimizing the overall external dimensions of the transformer;
- minimizing total mass to reduce material and transportation costs;
- minimizing the volume of insulating oil;
- selecting materials already in stock to streamline production;
- prioritizing designs compatible with underutilized manufacturing equipment to balance factory load.

Despite these multiple objectives, the decisive criterion in most industrial scenarios is minimizing the overall production cost. This cost is dominated by the expenses of electrical steel in core, aluminium/copper windings, insulating components, and steel tank with its cover. By contrast, processing, labor, and overhead costs typically constitute only about 10%–20% of the total manufacturing cost, as indicated by internal industrial data from Elpro-Energo Transformers s.r.o. [18]. This cost structure emphasizes the need for efficient material utilization, making manual and experience-driven optimization increasingly inadequate in an era where computational approaches could deliver substantial gains.

It is important to note that the calculation of power transformers, particularly for small distribution and medium-power units with ratings up to approximately 10 MVA, is in most cases performed using purely analytical equations derived from Maxwell's equations and other well-known physical expressions, simplified for practical engineering use, as stated in [20]. Examples include (3.7) for calculating the flux density in the core, as well as analytical methods for estimating short-circuit impedance and performing thermal calculations. These analytical approaches are supplemented with numerous empirical formulas derived from industry experience, along with corrective coefficients that adjust the results to better reflect real-world performance. Such corrections are essential for improving estimates of no-load and load losses, as well as for accurately predicting the final winding dimensions.

For larger power transformers, however, it becomes necessary to integrate finite element method (FEM) simulations into the design process to achieve sufficiently

accurate results. According to [20]–[22] FEM analysis is particularly critical for evaluating short-circuit forces and mechanical structural integrity, performing detailed thermal calculations to determine winding hot-spot temperatures and oil distribution within the tank, and quantifying additional eddy current losses in windings, steel clamping structures, and the tank itself.

A key distinction between these two design approaches lies in the nature of production: manufacturers specializing in large transformers typically produce only a few dozen units per year. This market segment faces less competition and lower overall demand compared to the distribution transformer industry, where manufacturers may produce thousands of units annually to remain economically viable. Although the application of FEM in the design of distribution transformers could yield more optimized and accurate results, the constraints of time, skilled workforce availability, and budget often make FEM impractical for standard designs. Consequently, FEM is generally reserved for special or custom transformer types. For the design of standard distribution transformers, the priority is to develop highly accurate analytical and empirical mathematical models that balance precision with efficiency in the design process.

2.2 Mathematical Model

This section outlines the principles of the mathematical model employed for transformer design. The most effective way to present this process is through the block diagram in Fig. 2.1, which illustrates all the essential steps relevant to both purely analytical calculations and those incorporating FEM simulations. These blocks demonstrate how the overall calculation is divided into distinct operations, reflecting the logical structure of transformer components and the fields that require analysis. The following subsections describe the function of each block, specify their inputs and outputs, and explain how data are processed throughout the design workflow.

2.2.1 Input Parameters

The input parameters, representing the initial block in the diagram shown in Fig. 2.1, define the starting point of the design process and are determined by the technical specification. This specification must always be provided when an order is placed by the customer, as it establishes the framework for all subsequent calculations and design decisions. At a minimum, it should contain the set of values and detailed information listed in Table 2.1.

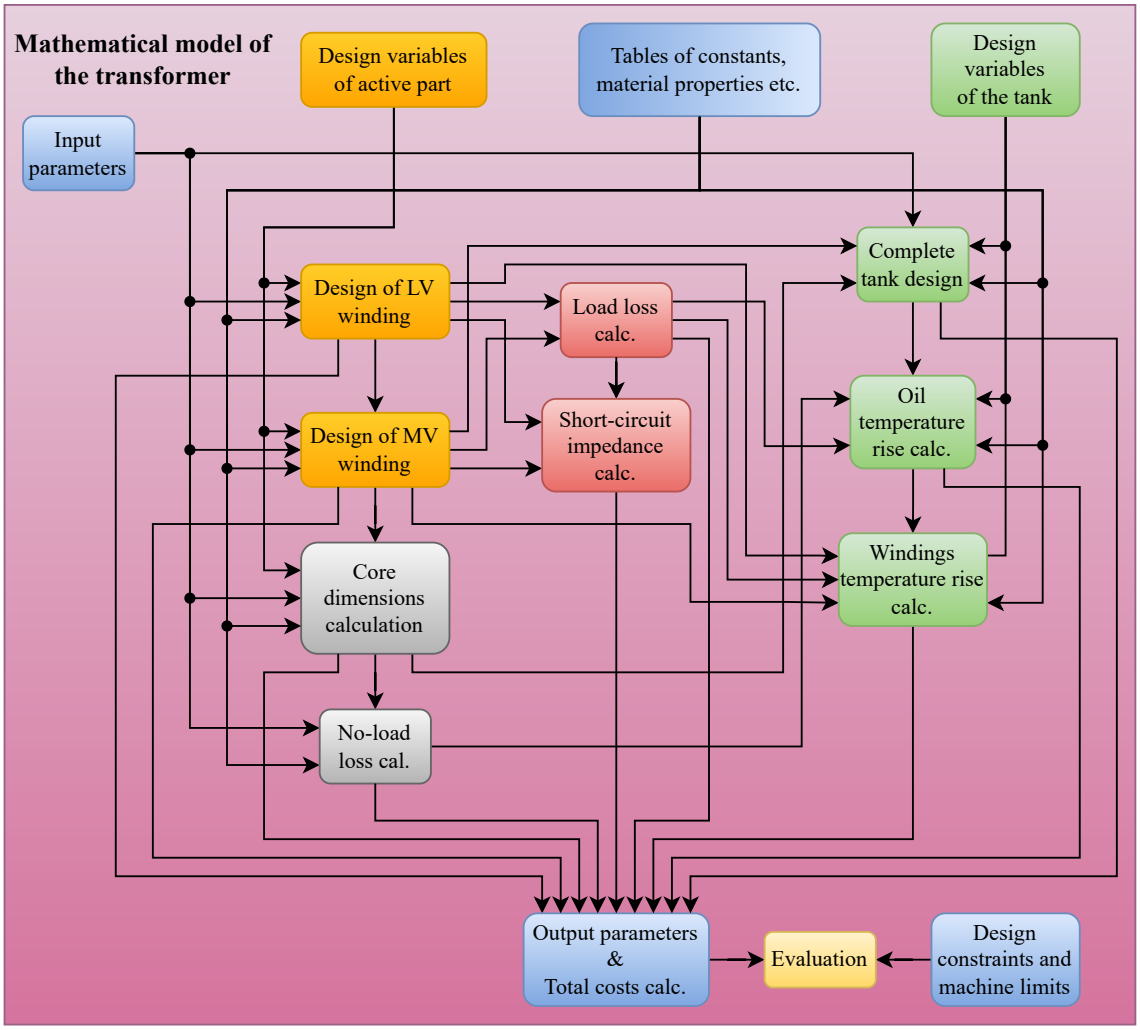


Fig. 2.1: Block diagram of the mathematical model used for oil-immersed transformers calculations.

2.2.2 Tables of Constants, Material Properties etc.

This block in Fig. 2.1 encompasses all fixed parameters that enter the calculations of the transformer mathematical model. These parameters represent essential boundary conditions, reference values, and technological constraints that must be taken into account during the design process. Typical categories of information included in this block are listed below:

- *Material properties*—fundamental physical and electrical properties of conductors (aluminium/copper), electrical steels (both magnetic and structural), insulation materials (papers, varnishes, and oils), including parameters such as electrical conductivity, density, and thermal conductivity.
- *Material databases*—tables specifying the range of variants available from suppliers. For example, round conductors are provided only within defined ranges

Table 2.1: Technical Specification Characteristics Serving as Inputs in Transformer Design

Technical characteristic	**Unit**	**Note / options**
Rated power	kVA	Max. continuous transformer rating
Primary rated voltage	V	Higher rated voltage
Secondary rated voltage	V	Lower rated voltage
Vector group	-	Star, delta, zig-zag; with/without neutral; clock-hour notation
Taps	%	Typically 5 positions ($2 \times \pm 2.5$)
Insulation levels ($U_{\mathrm{m}}; AC; BIL$)	kV	Highest voltage; power-frequency applied voltage; lightning impulse
Nominal frequency	Hz	Typically 50, 60, or 16.7 (railway)
Cooling method	-	ONAN, ONAF, etc. oil natural air natural/forced
Short-circuit impedance	%	4–6% (distribution), 8–10% (power)
Max. no-load losses	W	EcoDesign limits (in EU)
Max. load losses	W	EcoDesign limits (in EU)
Peak efficiency index (PEI)	%	Max. efficiency at rated load
Sound power level	dB(A)	Not IEC-specified; customer-specific
Max. outer dimensions	mm	Max. length, width, height
Max. total weight	kg	Optional specification
Max. oil weight	kg	Optional, for safety pit sizing
Max. ambient temperature	°C	Typically 40; may include avg. yearly/monthly temperature
Installation altitude	m.a.s.l.	Typically ≤ 1000
Overload	-	Overload class; affects cooling
Others	-	Pollution l., corrosion c., accessories

of diameters, foils within specific widths and thicknesses, and insulation paper or strips in standardized dimensions.

- *Pre-prepared parts*—catalogued data of standardized components, such as bus-bar leads from foil windings (depending on the current rating), clamping beams dimensioned by core weight, tank wall thickness as a function of height, or standardized conductor cross-sections for interconnections.
- *Accessories*—similar to the previous group but with less impact on electrical performance and more influence on mechanical layout and external dimensions. Examples include tables of bushings (porcelain and plug-in types), tap-changer

options classified by voltage and current ratings, as well as standardized chassis and roller dimensions.

- *Correction coefficients*—empirical factors used to approximate real behavior and manufacturing deviations, e.g., for more accurate prediction of no-load losses, load losses, or manufactured winding dimensions.
- *Economic data*—prices of materials (including LME quotations), exchange currency rates, processing fees, and cost tables for materials per kilogram or per meter, as well as prices of auxiliary accessories.

2.2.3 Design Variables of the Active Part

Design variables represent the parameters actively adjusted and selected by the designer to achieve the desired transformer performance. These variables interact with all relevant equations and functions within the transformer mathematical model, incorporating the input parameters, constants, and material properties defined in preceding sections. The primary design variables for a typical two-winding active part of oil-immersed distribution and medium-power transformers are summarized in Tab. 2.2.3. For dry-type cast-resin transformers, the set of variables remains largely similar. The main difference lies in the medium-voltage winding design, where alternative manufacturing technologies are used compared to oil-immersed designs, leading to some variation in specific design variables. However, foil windings in dry-type transformers follow essentially the same principles as in oil-filled units.

2.2.4 Design of Low-Voltage Winding

The inner winding of a typical distribution transformer (or medium power transformer) with two concentric windings on a single core limb is, in most cases, designed for low voltage (LV). For this application, foil winding technology is generally the most efficient manufacturing method due to the large conductor cross-section required to accommodate the high currents associated with low-voltage operation. In this process, a continuous conductive foil is wound together with insulation paper, which is slightly wider than the foil. The resulting edge gaps are filled with insulation strips to achieve complete dielectric separation. Structurally, this type of winding can be compared to a large cylindrical capacitor, as illustrated in the manufacturing process in Fig. 2.2(a). In oil-immersed transformers, the insulation paper is typically cellulose-based and coated with a heat-curable epoxy resin in a diamond-shaped pattern, commonly referred to as diamond pattern paper (DPP) [23], [24]. This resin serves to reinforce the winding mechanically, enhancing its resistance to short-circuit forces. In dry-type transformers, insulation is typically achieved using polyester film, pre-impregnated with a heat-resistant epoxy resin, commonly referred to as Prepreg.

Table 2.2: Transformer Design Variables of the Active Part Grouped by Basic Components

Group	**Variable**	**Note / options**
Core	Core shape	Oval or round cross-section
	Steel sheet grade	Vary by specific losses and sheet thickness
	Number of core steps	Typically 3–10, depending on core size
	[a]Length of straight part	For oval cores—defines center section
Lower voltage winding	Type of winding	Foil, round wire, or rectangular wire
	Conductor material	Aluminium or copper
	[b]Number of turns/layers	In foil winding, one turn is full layer
	[a]Turns per layer	Applies only to wire windings
	[b]Foil/wire width	Tables of options or flattened round wire
	[b]Foil/wire thickness	Tables of options or flattened round wire
	[a]Parallel wires	One or multiple wires simultaneously
	Cooling channel	None to multiple axial cooling channels
Higher voltage winding	Conductor material	Aluminium or copper
	[b]Foil/wire width	Tables of options or flattened round wire
	[b]Foil/wire thickness	Tables of options or flattened round wire
	[a]Parallel wires	One or multiple wires simultaneously
	Cooling channel	None to multiple axial cooling channels
Distance adaption	Core-to-winding	Increased for safety or machines' limits
	Between windings	Increased for safety or higher reactance
	Between phases	Increased for safety
	Yoke-to-winding end	Increased for safety or machines' limits

[a] The availability of this variable depends on the selection of other variables.

[b] The available options for this variable are determined by the selection of other variables.

Although the inner winding is predominantly used for low-voltage applications, it can also be designed for medium-voltage operation in certain cases. However, positioning the lower-voltage winding closer to the core is generally preferred, as it simplifies insulation requirements and improves manufacturability. In foil winding applications, placing the winding on the outer diameter is more complex due to the bus-bar outlets formed from the foil. These outlets tend to distort the winding geometry; such deformation is more easily managed when the inner winding's bus-bar emerges toward the core, where the lamination thickness provides structural tolerance. These design considerations are discussed in more detail in [12].

If the inner winding's voltage rating exceeds a certain threshold, a multilayer winding composed of flattened round or rectangular wire is often employed, as described in the following subsection.

Within the transformer mathematical model, this block in Fig. 2.1 is responsible for defining the final geometrical dimensions of the inner winding, including its diameters, height, and insulation structure, based on the selected conductor width and thickness, number of turns, and cooling channel configuration. Additional outputs calculated in this step include the total conductor length, conductor and insulation weight, and final DC resistance, among other performance parameters.

2.2.5 Design of Medium-Voltage Winding

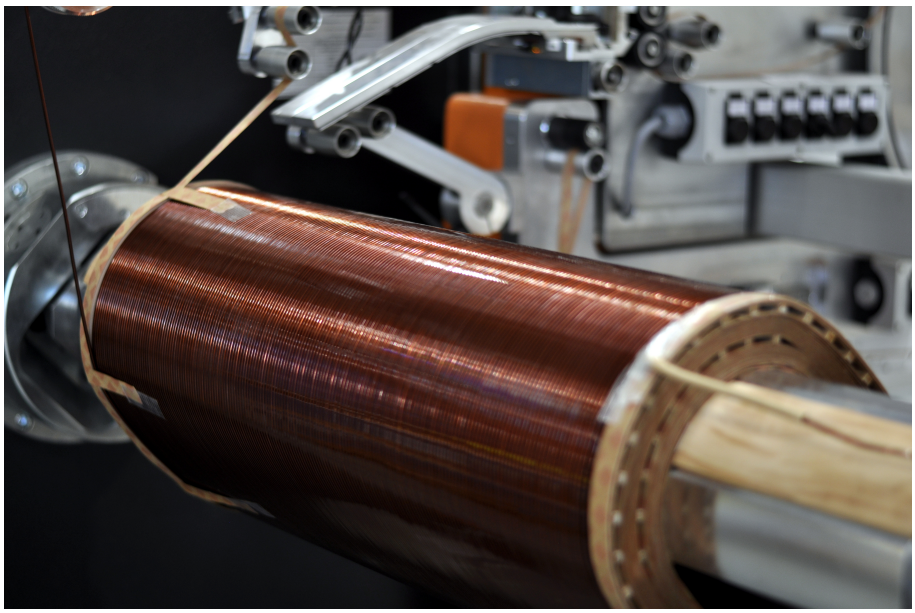
The outer winding of a typical distribution transformer (or medium power transformer) is generally designed for medium-voltage (MV) operation. In this case, the current is significantly lower than in the low-voltage (LV) winding, which makes wire winding technology the preferred choice. For oil-immersed transformers, the most common manufacturing method is the *multilayer winding* technique, illustrated in the manufacturing process in Fig. 2.2(b), where the conductor is made from flattened round or rectangular wire, as reported by winding machine manufacturers [25], [26]. Flattening improves the filling factor of the winding and reduces additional eddy-current losses caused by leakage magnetic field, as discussed in [9]. The wire winding is applied directly over the main cooling duct on LV winding, with diamond pattern paper (DPP) strips or wide DPP sheets inserted between layers providing inter-layer insulation and electrical separation at the winding ends.

In dry-type transformers, the outer winding is typically manufactured using the *gravity-line* process, in which the wire is pre-coiled and placed into a mold. Alternatively, multiple foil-strip coils connected in series may be used, depending on the current rating. The entire assembly is then vacuum-cast with epoxy resin and subsequently cured in an oven to achieve high mechanical and dielectric strength.

The outer winding also typically accommodates the transformer's tap connections for voltage adjustment. Positioning the taps on the outer winding simplifies their implementation due to the greater number of turns available, which enables finer voltage regulation. However, the introduction of taps slightly deforms the winding's shape; therefore, they are usually positioned near the end of the winding to minimize manufacturing complexity.

The block of the mathematical model in Fig. 2.1 responsible for design of MV winding uses inputs similar to those for the inner winding, but incorporates recalculations based on the specified voltage ratio and tap settings. The calculation yields the outer winding's geometrical dimensions, including diameters, height, and insula-

(a)



(b)

Fig. 2.2: Manufacturing process on winding machines for oil-immersed transformers: (a) LV foil winding and (b) MV multilayer wire winding, adapted from [18].

tion structure. It also determines the total conductor length, conductor and insulation mass, and the final direct current (DC) resistance of the winding. When both LV and MV windings have completed designs, this block should also include the calculation of short-circuit forces and the verification of the mechanical endurance of both windings, as they are interdependent.

Fig. 2.3: Assembled transformer core with clamping beams prepared for winding installation, adapted from [18].

2.2.6 Core Dimensions Calculation

After completing both windings designs, the next block in Fig. 2.1 is responsible for calculating the core dimensions. Based on the winding geometry and specified clearance requirements between the core and windings, the core window size and limb pitch can be determined. As confirms [27] for a three-phase transformer (both oil-immersed and dry-type), the magnetic circuit typically adopts a *core-type* configuration, where each phase is assigned its own limb, and these limbs are interconnected by upper and lower yokes, as demonstrated in Fig.2.3.

The cross-sectional dimensions of the core are subsequently estimated according to predefined design parameters, which define the limb shape and geometry. Using these calculated values, the total core weight is derived, along with the dimensions and weight of associated structural elements, such as the clamping beams.

2.2.7 No-Load Losses Calculation

The process of no-load losses calculation represented by the corresponding block in Fig. 2.1 is discussed in detail in Chapter 3, as the accurate estimation of no-load

losses is closely related to the optimization of the transformer design and constitutes one of the primary objectives of this thesis.

In brief, the calculation process involves equations that incorporate the total core weight, material properties as a function of the induced flux density, and correction factors that account for variations introduced by manufacturing processes. This block also typically includes an estimation of the no-load current magnitude as well as predictions of the sound power level and the sound pressure levels, commonly evaluated at reference distances of 1 m and 0.3 m from the transformer.

2.2.8 Load Losses Calculation

Along with no-load losses, the calculation of load losses is a key output of transformer design, and it has become even more important since the EcoDesign requirements came into force. In the block for load losses calculation shown in Fig. 2.1, DC losses are estimated based on the winding resistances and their interconnections, together with nominal currents. These losses are then recalculated to a reference temperature, which is, for example, 75 °C for typical oil-immersed transformers according to [28], or 120 °C for dry-type transformers according to [29].

Another important step is adding additional losses to the DC losses. These losses are mainly caused by eddy currents induced in the windings by the leakage magnetic field. They can be estimated approximately using analytical calculations, where the dimensions of the windings, nominal current, and number of turns are required for their estimation. However, FEM analysis provides more accurate results, as demonstrated in [30]. The most challenging case is estimating the additional load losses in foil windings, where analytical procedures fail, as reported in [10]. Therefore, FEM analysis is currently the only viable method for estimating losses in this winding type, as shown, for example, in [31]. Additional components of load losses also arise in construction steel parts, such as clamping beams, and a significant portion is generated in the transformer tank, as stated in [20], and in the cover, particularly near bushings, as analyzed in [32].

2.2.9 Short-Circuit Impedance Calculation

The block for short-circuit impedance calculation in Fig. 2.1 primarily involves the calculation of the leakage reactance produced by the interaction of both windings. The leakage reactance constitutes the dominant portion of the short-circuit impedance, and its estimation requires the dimensions of the windings, nominal current, and number of turns. The active component of the short-circuit impedance can then be easily derived from the load losses, although its influence on the overall value is minor.

Table 2.3: Transformer Design Variables of the Oil Tank Grouped by Similarity of Effect on the Final Shape

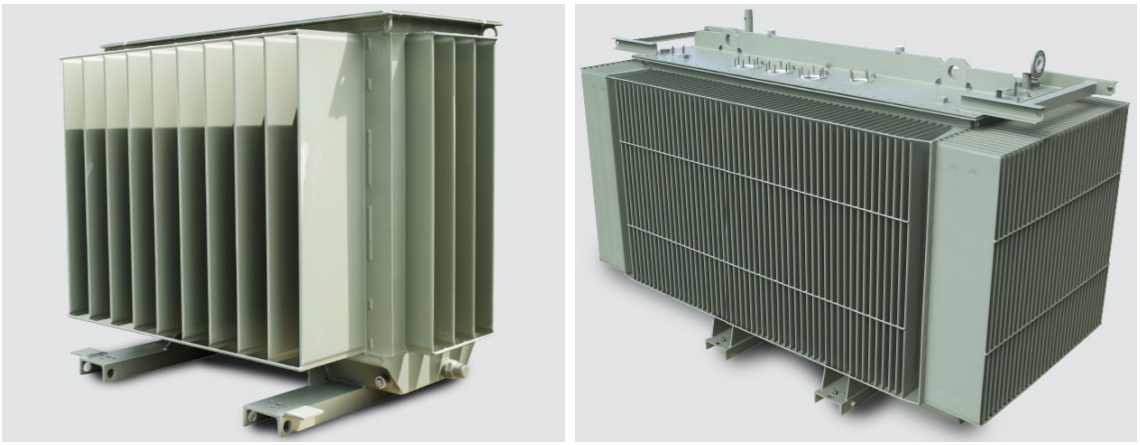
Group	Variable	Note / options
Boolean settings	Symmetry of the tank	Yes or no, same fins on all sides
	Additional fins	Yes or no, additional fins in corners
	Bottom narrowing	Yes or no, narrowing under windings
Depth of fins	Fins depth on LV side	Fins of corrugated walls may be manufactured in range of 50 mm to 400 mm with step of 10 mm; deeper fins with embossing and spot welding
	Fins depth on MV side	
	Fins depth on the left side	
	Fins depth on the right side	
Pitch of fins	Fins pitch on LV side	Fins of corrugated walls may be manufactured with pitch from 45 mm to 90 mm with step of 5 mm between
	Fins pitch on MV side	
	Fins pitch on left side	
	Fins pitch on right side	
Distance adaption	Increase in tank length	Inner dimensions of the tank may be increased from minimal distances to the active part due to voltage safety
	Increase in tank width	
	Increase in tank height	

The accuracy of short-circuit impedance calculation is highly dependent on the type of transformer. For instance, the calculation differs significantly for zig-zag winding configurations, as described in [33], or for split-winding transformers, as analyzed in [11], compared to a conventional two-winding transformer. Nevertheless, analytical calculations are in most cases sufficiently precise.

2.2.10 Design Variables of the Tank

The design variables of the hermetically sealed oil tank (without conservator) can be considered separately from those of the active part, as illustrated in Fig. 2.1, since they are independent of each other. In the case of dry-type transformers, this block does not participate in the design process, except for the calculation of temperature rise of the core and windings in air. Only the calculation of tank dimensions and the estimation of oil temperature rise is omitted.

The fundamental variables that determine the final shape and dimensions of the oil tank, including its cover, are listed in Table 2.3, along with a brief explanation of their meaning and the common options typically encountered in practical manufacturing.



(a) (b)

Fig. 2.4: Tanks for oil-immersed three-phase transformers: (a) smaller distribution with a typical structure of corrugated walls, featuring shorter fins and larger pitch, (b) medium-power transformer with a specialized structure of corrugated walls, where fins in the corners rest on fins from the adjacent wall and deeper fins have the lowest possible pitch, adapted from [34].

2.2.11 Complete Tank Design

When variables for the oil tank are selected, detailed tank dimensions are calculated in a separate block in Fig. 2.1, enabling determination of the cooling surface, which differs for radiation and airflow, along with the inner and outer dimensions. The inner dimensions depend on the active part dimensions, to which the insulation distance with a safety factor is added to account for inaccurate insertion of the active part into the tank or unpredictably enlarged dimensions of the active part. The inner dimensions can also be intentionally increased for higher insulation safety or to accommodate installation of atypical components or more complex connections inside the transformer. For this purpose, the variables in the group *Distance adaption* in Table 2.3 are used.

Other variables listed in the table determine the placement and depth of the fins on corrugated walls, with the option to select different values for each wall. In practice, symmetry is most frequently applied, where all walls have fins with the same parameters or at least three identical walls, to simplify and accelerate the manufacturing process. A typical oil tank for three-phase oil-immersed transformer is shown in Fig. 2.4(a). For larger transformers requiring additional cooling surface, special tank designs may be needed, where fins are also added to outer surfaces of adjacent walls to increase cooling without enlarging the overall length or width of the tank, as shown in Fig. 2.4(b).

2.2.12 Oil Temperature Rise Calculation

The oil temperature rise is another step in the design process, performed when the detailed dimensions of the tank, no-load losses, and load losses are known. The block in Fig. 2.1 incorporates empirical expressions to determine heat transfer coefficients, for example from [35], which are needed to estimate the average temperature rise of the oil in the tank. These expressions provide sufficient precision for distribution transformers, while for larger transformers correction factors must be applied based on experience from temperature rise tests. Using empirical coefficients and tank size, the top-oil temperature rise can be estimated. At this point, the oil temperature is highest, and for a typical transformer filled with naphthenic oil, it should not exceed 100 °C, which means a temperature rise of 60 K for the standard maximum ambient temperature of 40 °C, according to [36].

For a more precise temperature rise calculation, the computational fluid dynamics (CFD) method can be used for airflow analysis and heat transfer coefficient estimation, as presented in [37]. This approach requires extensive parameter tuning based on measured data to obtain valid results.

2.2.13 Windings Temperature Rise Calculation

When the oil temperature rise in the tank is calculated, the design process continues with the block in Fig. 2.1, which covers the calculation of the temperature rise of the windings. If these windings include cooling channels, the temperature rise must be calculated separately for each section divided by a channel. For this calculation, the load losses generated in the windings—ideally with a precise estimation of additional eddy losses—as well as the surface areas of the windings and the thicknesses of the insulation layers within the windings must be known.

An important part of this calculation is the accurate estimation of heat transfer coefficients. Since these coefficients strongly depend on the temperature difference between the windings and the oil, the calculation must be performed iteratively until the error is within an acceptable range. According to [36], the average winding temperature cannot exceed 105 °C, while the hot-spot temperature must not exceed 118 °C. This corresponds to a temperature rise of 65 K and 78 K, respectively, when considering the standard maximum ambient temperature of 40 °C.

Similarly to the oil temperature rise calculation, the analytical approach provides only limited accuracy, though it is sufficient for distribution and medium-power transformers. For a more precise estimation of temperature rise, CFD methods—sometimes combined with FEM analysis—are necessary, as demonstrated in [38].

Table 2.4: Key Output Parameters of Oil-Immersed Transformer Grouped by Linked Calculations

Group	Variable	Note
Core	Flux density in core	Recommended max 1.7 T
	No-load losses	No tolerance to EU limits
	No-load current	Usually very low for EcoDesign
	Sound power level	Or sound pressure levels
Windings	Load losses in LV w.	For loss distribution comparison
	Load losses in MV w.	
	Total load losses	No tolerance to EU limits
	Short-circuit impedance	Tolerance ±10%
Overall	Peak efficiency index (PEI)	In EU for powers >3150 kVA
Temperature rise	Top-oil temperature rise	For ambient 40 °C, limit 60 K
	LV w. average temp. rise	For ambient 40 °C, limit 65 K
	LV w. hot-spot temp. rise	For ambient 40 °C, limit 78 K
	MV w. average temp. rise	For ambient 40 °C, limit 65 K
	MV w. hot-spot temp. rise	For ambient 40 °C, limit 78 K
Outer dimensions	Total transformer length	Including fixing bars on fins
	Total transformer width	Including fixing bars on fins
	Total transformer height	Including rollers and accessories
Costs	Total material costs	Of both active part and oil tank

2.2.14 Output Parameters and Total Costs Calculations

When all previous calculation steps are completed, the main outputs can be extracted, including the calculation of total costs, in the separate block shown in Fig. 2.1. an example summary of these key parameters for an oil-immersed transformer is provided in Table 2.4. These parameters are further evaluated in another block to verify whether they meet the requirements and remain within specified limits, as described in Section 2.2.15.

Typically, the most important output of the design process is the total material cost of the transformer, since processing and other costs are less significant, as explained in Section 2.1. The total cost is calculated based on the masses of primary materials, such as electrical steel, conductors (foil and wire), insulation papers, steel for clamping beams, steel for the tank and cover, and oil, multiplied by their respective unit prices, including material processing costs provided by suppliers. In addition, accessories such as bushings, tap-changer, thermometer, protective elements, conservator, and chassis with rollers should also be considered.

2.2.15 Evaluation with Design Constrains and Machine Limits

The final two blocks in Fig. 2.1 serve for the evaluation and verification of the results obtained according to the selected input variables with respect to the design constraints and machine limits. The design constraints and machine limits are essentially the same as those described in Section 2.2.2, but for clarity and to distinguish the different phase of their application, they are presented separately. Their roles are as follows:

- *Design constraints*—constants defining acceptable tolerances, including limits for losses, efficiency, temperature rises, deviations from the voltage ratio, dimensional constraints, and maximum tank pressure.
- *Machine limits*—production-related constraints reflecting the capabilities of available equipment. These include, for instance, maximum lamination size for cutting lines, weight restrictions on core assembly tables, permissible winding diameters on winding machines, and maximum transformer dimensions allowed in the oil-filling chamber.

Based on the evaluation, the designer should be able to definitively decide whether the engineered design is valid and, considering the final cost, determine how to proceed with the next selection of variables or whether a probable optimum has been achieved in terms of minimal cost or other requirements.

2.3 Summary of Chapter 2

This chapter has provided a comprehensive overview of the transformer design process, highlighting the interaction between analytical methods, empirical corrections, and modern simulation tools. The design workflow integrates human expertise with structured computational models to ensure that the final transformer satisfies technical specifications, machine constraints, while maintaining the cost-effective requirement. Current industrial practices were reviewed, showing that manual calculations, empirical adjustments, and semi-automated tools dominate, with cost optimization guiding design decisions. Analytical models remain standard for distribution and medium-power units, while finite element method (FEM) is primarily used for large or custom transformers to perform detailed thermal and mechanical analyses. The chapter emphasizes the reliance on engineer expertise and simplified models due to time and cost limitations, underscoring the need for automation and advanced optimization in modern transformer design.

The process begins with the definition of input parameters derived from customer specifications, including rated power, voltage levels, insulation class, cooling method, and mechanical constraints. These inputs guide the selection of material properties, component databases, and predefined parts, forming the foundation of the mathematical model. Design variables of the active part, such as core geometry, conductor type, number of turns, layer configuration, and cooling channels, are iteratively optimized to meet electrical, thermal, and mechanical objectives. The low-voltage (LV) winding is generally implemented using foil technology for high-current handling, whereas the medium-voltage (MV) winding typically uses multilayer wire or foil in the case of oil-immersed transformers, which serves as the main object of research.

The core dimensions are calculated based on the winding geometry and specified clearances, ensuring that the magnetic circuit meets flux density requirements while maintaining mechanical stability. No-load and load losses are estimated using a combination of analytical equations, empirical correction factors, and FEM simulations where necessary. Short-circuit impedance is similarly evaluated, with special attention to winding configuration and interactions. For oil-immersed transformers, tank design—including dimensions, fin placement, and cooling surface—ensures sufficient heat dissipation, while dry-type units rely exclusively on air cooling.

Thermal performance is assessed through iterative calculation of oil and winding temperature rise. Empirical methods provide adequate precision for distribution transformers, while computational fluid dynamics (CFD) and FEM simulations may be used to capture detailed fluid dynamics, heat transfer, and hot-spot effects for more accurate estimation. For oil-immersed transformers, maximum allowable temperature rises, such as top-oil (60 K) and winding hot-spot (78 K) for standard ambient conditions (40 °C), are enforced to ensure long-term reliability.

The final outputs of the design process include all electrical, thermal, and mechanical parameters, along with total material costs. These outputs are evaluated against design constraints and machine limits, which reflect both engineering tolerances and manufacturing capabilities. The designer remains central to the process, making iterative decisions to achieve an optimal balance between performance, manufacturability, and cost efficiency.

In summary, transformer design represents a complex, multi-objective optimization problem that blends classical analytical techniques, empirical knowledge, and advanced simulation tools. A structured workflow, as presented in this thesis, enables systematic exploration of design alternatives, ensuring that the engineered solution meets performance requirements, complies with standards, and remains economically viable for industrial production.

3 State-of-the-art Transformer No-Load Losses Calculation

As stated in [20], the transformer's magnetic circuit, also known as the core, is a critical component whose geometry and material properties strongly influence every other part of the transformer. Consequently, when designing a transformer with the goal of optimization, the process should begin at the transformer's center—the core itself. Among the core's various characteristics, one of the most significant is the losses generated within it during operation. These losses are manifested as no-load losses, where they constitute the dominant share, and are typically assessed during a standard no-load test.

Accurate estimation of no-load losses is essential for effective transformer optimization. Therefore, this chapter provides a theoretical overview of the mechanisms causing these losses, describes their calculation using conventional analytical methods, and examines the factors contributing to discrepancies between calculated and measured values, while also considering whether such deviations can be reliably predicted.

A solid understanding of no-load loss behavior is closely linked to transformer core manufacturing technologies, which are only briefly introduced in this thesis. A comprehensive discussion of manufacturing options, core processing techniques, and the overall production workflow is presented in the author's master's thesis [12].

3.1 Theory of Transformer No-Load Losses

No-load losses are measured by the open-circuit test, i.e., when the transformer secondary is unloaded. In this case, the secondary current I_2 is zero, and the input current I_{10} is approximately equal to the magnetizing current I_μ as stated in [20], [28]. Under these conditions, the measured input power consists primarily of core (steel) losses together with minor conductor losses as Joule heat caused by the no-load current in the energized winding. Since both I_μ and the primary resistance R_1 are small, the no-load Joule loss $P_{\mathrm{j}10}$ is negligible, and the no-load losses P_{10} can be approximated by the core (steel) losses P_{Fe} as described in [20], [39]. This relationship can be expressed as:

$$P_{10} = P_{\mathrm{j}10} + P_{\mathrm{Fe}} \cong P_{\mathrm{Fe}}. \tag{3.1}$$

Core losses are generally decomposed into hysteresis, classical eddy current, and anomalous components according to [40], [41]. Hysteresis losses are proportional to the area of the magnetic hysteresis loop of the steel as illustrated later in Fig. 6.6(b) and are strongly affected by the chemical composition, thermal treatment (e.g., annealing), and mechanical processing (rolling direction and texture). Impurities such

as carbon, sulfur, oxygen, or nitrogen widen the loop and increase hysteresis losses. In contrast, silicon increases electrical resistivity and inhibits carbon in solid solution, thereby reducing eddy current losses and partially lowering hysteresis losses [41]–[43]. For soft magnetic steels at power frequency, hysteresis losses are approximately proportional to frequency f, whereas eddy current losses scale with f^2. Anomalous losses represent microscopic processes in the domain wall not captured by the classical components [40], [41]. For engineering practice, a simplified Steinmetz-type expression is often applied, obtained by combining (3.3), (3.4), and (3.5), in accordance with [20], [40], [44]:

$$P_0 = P_\mathrm{h} + P_\mathrm{e} + P_\mathrm{a}. \tag{3.2}$$

The individual loss components are given as:

$$P_\mathrm{h} = k_\mathrm{h} \cdot f \cdot B^\alpha, \tag{3.3}$$

$$P_\mathrm{e} = k_\mathrm{e} \cdot (f \cdot B)^2, \tag{3.4}$$

$$P_\mathrm{a} = k_\mathrm{a} \cdot (f \cdot B)^{1.5}, \tag{3.5}$$

where P_h denotes hysteresis losses, P_e eddy current losses, and P_a anomalous (excess) losses. The constants k_h, k_e, and k_a depend on the grade, volume, lamination thickness, and processing of the core material. The symbol B represents the peak flux density. For soft magnetic steels, the exponent α typically ranges between 1.6 and 2.2, depending on the hysteresis loop shape, as reported by [40], [41].

At low frequencies, core hysteresis losses dominate, whereas at higher frequencies, eddy-current and anomalous losses in the core, become significant [20], [41]. Standardized test procedures, as defined in IEC 60076-1 [28], specify that no-load losses P_{10} are measured at rated voltage and frequency under defined ambient conditions, ensuring comparability across different designs and materials as stated in [39]. The no-load test procedure is described in detail in Section 6.2.1.

3.2 Analytical Calculation of Core Losses

According to well-established transformer theory, the no-load losses are almost entirely attributable to the losses occurring within the core, which is typically constructed from high permeability (HiB) cold-rolled grain-oriented (CRGO) steel laminations. The core losses predominantly arise from magnetic flux traversing the core limbs and yokes, producing mentioned hysteresis, eddy current, and anomalous losses as the primary contributors. Understanding these losses is essential for accurate transformer design, efficiency assessment, and thermal management.

From a practical standpoint, the simplest yet acceptably accurate expression for estimating the nominal core losses is provided according to [20]:

$$P_{\mathrm{Fe,n}} = G_{\mathrm{Fe}} \cdot p_{\mathrm{Fe}}. \tag{3.6}$$

where P_{Fe} denotes the total core loss, G_{Fe} is the core mass (calculated from its volume and the lamination material density), and p_{Fe} is the specific loss of a single core lamination sheet determined through Epstein's test. The value of p_{Fe} can be obtained from the manufacturer's loss curve at the peak operating flux density B. Examples of such curves are shown in Fig. 3.1 for frequencies of 50 Hz and 60 Hz, corresponding to the steel grade most commonly applied in transformers under EcoDesign requirements. This grade exhibits specific losses of 0.7 W/kg at 1.7 T—the standard reference point for HiB steel—with a sheet thickness of 0.23 mm. The curves were measured by the single sheet test (SST), an alternative method for steel sheet characterization. According to IEC 60404-8-7 [45], SST results should be multiplied by a correction factor of 0.925 to be comparable with Epstein test results (the reference method); however, this correction has already been incorporated into the presented curves. The curves are typically provided by suppliers in the form of plots or tables, requiring interpolation of the data points. A piecewise linear interpolation is generally sufficient within the operating interval, whereas cubic spline interpolation may be employed for higher accuracy.

The flux density B for the in the core cross-section A can be calculation by well-known expression as:

$$B = \frac{\sqrt{2} \cdot U_{\mathrm{rms}}}{2 \cdot \pi \cdot f \cdot N \cdot A}, \tag{3.7}$$

where U_{rms} is the root-mean-square phase voltage and N is the number of turns of the winding whose voltage is used for the calculation.

However, the nominal core losses $P_{\mathrm{Fe,n}}$ do not account for the additional losses arising in the assembled core. Therefore, (3.6) is extended for calculation of total core losses P_{Fe} as:

$$P_{\mathrm{Fe}} = P_{\mathrm{Fe,n}} \cdot k_{\mathrm{BF}} = G_{\mathrm{Fe}} \cdot p_{\mathrm{Fe}} \cdot k_{\mathrm{BF}}, \tag{3.8}$$

where the variable k_{BF} is known as the *building factor* and serves as a correction coefficient accounting for the increase in losses of the assembled core compared to the nominal core losses $P_{\mathrm{Fe,n}}$. This equation forms the basis of the conventional approach for calculating total core losses, which is commonly used by most transformer manufacturers during the design process.

The building factor in (3.8) reflects the effect of additional losses, and their interdependence can be expressed as:

$$k_{\mathrm{BF}} = \frac{100\% + p_{\mathrm{ANLL}}}{100\%}, \tag{3.9}$$

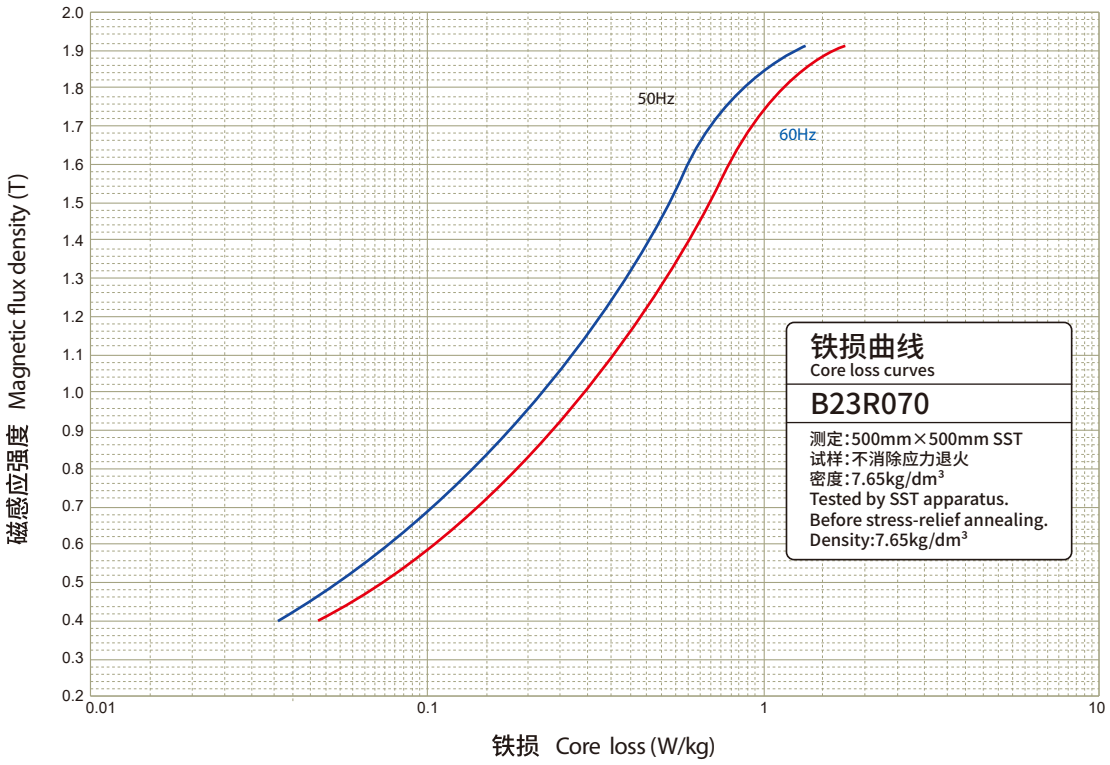


Fig. 3.1: Specific core loss curves for frequencies of 50 Hz and 60 Hz for electrical steel of grade 0.7 W/kg at flux density 1.7 T with sheet thickness 0.23 mm, adapted from [46].

where p_{ANLL} denotes the percentage value of *additional no-load losses* (ANLL). These losses arise from numerous factors, described later in Section 3.3, some of which can be estimated in advance, whereas others remain difficult to predict.

Currently, most transformer manufacturers estimate ANLL primarily from experience, often relying only on the transformer rated power as the main input variable, as noted in [47]. An example of such ANLL values, expressed as a function of rated power, is provided by the transformer manufacturer Elpro-Energo Transformers s.r.o. [18] in Table 3.1, which also lists the corresponding building factors. These tabulated values are routinely applied during the transformer design process to estimate total core losses with a suitable design margin. Later in this thesis, these reference values are used to compare measured no-load losses with two estimation approaches: a) conventional analytical calculations supplemented by corrections from the table, and b) a novel data-driven machine learning method proposed in this work, as illustrated in Figs. 6.5 and 6.15.

The precise additional losses p_{ANLL}, whose accurate prediction represents one of the two principal objectives of this thesis, can be calculated directly from measured total core losses P_{Fe} (obtained from no-load tests) and theoretical nominal core losses

Table 3.1: Building Factors and Relevant ANNL Depending on Transformer Rated Power from Experience by Elpro-Energo Transformers s.r.o. [18]

Rated power (kVA)	**Building factor $k_{\mathbf{BF}}$ (-)**	**ANLL $p_{\mathbf{ANLL}}$ (%)**
≤ 160	1.4	40
≤ 250	1.35	35
$\leq 3,150$	1.3	30
$\leq 10,000$	1.25	25

$P_{\mathrm{Fe,n}}$ by rearranging (3.8) and (3.9) as:

$$p_{\mathrm{ANLL}} = \frac{P_{\mathrm{Fe}} - P_{\mathrm{Fe,n}}}{P_{\mathrm{Fe,n}}} \cdot 100\%. \tag{3.10}$$

This equation forms the basis for processing the measured data analyzed in Chapter 6.

3.3 Causes of Additional No-Load Losses

The values of ANLL depend on several influencing factors. Some of these are generally predictable and can be categorized as follows:

- nonuniform flux density distribution,
- lamination joints,
- stacking holes, and
- rotational losses.

However, accurately accounting for these aspects during core design and no-load loss calculations requires a detailed analysis of flux density distribution within the magnetic circuit. This level of detail can currently be achieved only through finite element method (FEM) simulations. These simulations are computationally demanding and, to achieve sufficient accuracy, typically require full three-dimensional FEM analysis, which further increases modeling complexity and computation time.

In addition, several other factors contribute to ANLL, but are more variable and difficult to predict using traditional analytical methods. These include:

- mechanical processing of steel sheets,
- steel grade tolerances, and
- accuracy of core assembly.

All of these factors are further discussed in detail in the following subsections.

It is important to note that the magnitude of ANLL may vary even between cores of identical design due to differences in lamination batches, assembly precision, and local flux distribution irregularities. Consequently, accurate modeling and prediction of ANLL require consideration of both deterministic and stochastic elements of core

construction. This motivates the adoption of advanced computational techniques, such as data-driven or machine learning approaches, for precise estimation of core losses, which forms one of the primary objectives of this thesis.

3.3.1 Nonuniform Flux Density Distribution

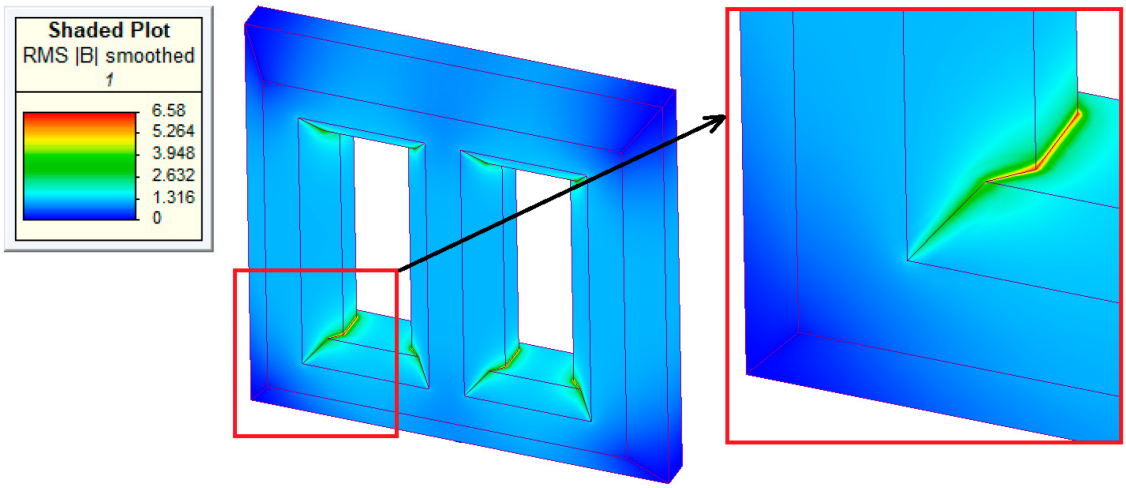
The magnetic flux density is rarely uniform across the lamination sheets. Pronounced irregularities are typically observed near the corners of the magnetic circuit, as illustrated in Fig. 3.2a, where localized regions of higher saturation occur at the interfaces between the core limbs and the yokes, particularly near the edges of the window as confirms [48]. These localized increases in flux density lead to elevated losses, since the specific core loss strongly depends on the flux density, as demonstrated by the curves in Fig. 3.1.

Furthermore, the flux path does not always align precisely with the rolling direction of the laminations, as illustrated in Fig. 3.2(b). In the case of CRGO steel sheets, magnetic properties are optimal exclusively along the rolling direction; therefore, specific parameters such as the core loss are defined, measured, and guaranteed only in this orientation. Any misalignment between the flux path and the rolling direction leads to deterioration of the magnetic performance and introduces additional supplementary losses, as confirmed in [49].

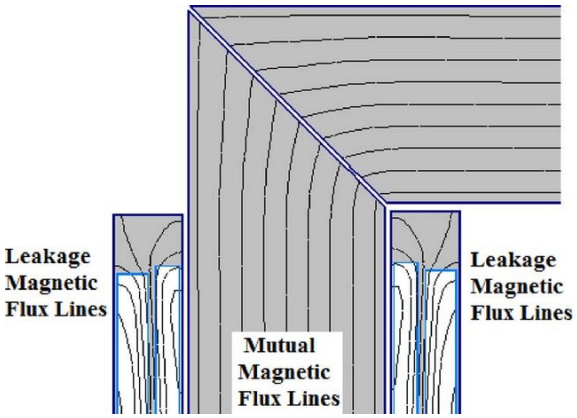
3.3.2 Lamination Joints

Another phenomenon increasing losses in the core occurs due to the reluctance at lamination joints, where air gaps exist between connecting laminations. Near the joints, the magnetic flux is redistributed into adjacent laminations of neighboring layers, as illustrated in Fig. 3.3(a), in order to circumvent the high reluctance of the air gaps, as stated in [52]. The additional losses are concentrated in areas next to the joints, as shown in Fig. 3.3(b). These losses are mainly associated with step-lap technology for core manufacture because of the large number of joints. Their significance is higher in the case of single step-lap (only one or two steps of laminations for the overlapping cycle) than in multi step-lap, which was primarily developed to mitigate this phenomenon, as reported in [53]. These losses also occur in wound-type cores, as analyzed in [54], although this technology contains fewer joints than the step-lap design.

According to [55], the increase in additional losses depends on the number of steps in one overlapping packet, the length of the overlap, and the number of lamination sheets in each step. In addition, the thickness of the laminations further affects



(a)



(b)

Fig. 3.2: (a) Flux-density distribution in the transformer magnetic circuit, including limbs and yokes, adapted from [50]. (b) Distribution of magnetic-potential lines in the transformer magnetic-circuit corner, showing the mutual and leakage magnetic flux, adapted from [51].

the magnitude of joint losses, as reported in [47]. The redistribution of flux into adjacent laminations substantially diverts the flux direction from the rolling direction, which also causes additional losses, similar to the effects discussed in Section 3.3.1.

3.3.3 Stacking Holes

Stacking (stamping) holes in core laminations are primarily introduced to facilitate automated core stacking during the cutting process. During production, newly cut laminations fall onto guiding pins that pass through these holes, thereby forming separated parts of the transformer magnetic circuit. Typically, two holes are created in each limb and yoke, for example, 10.5 mm in diameter for 10 mm pins, as displayed

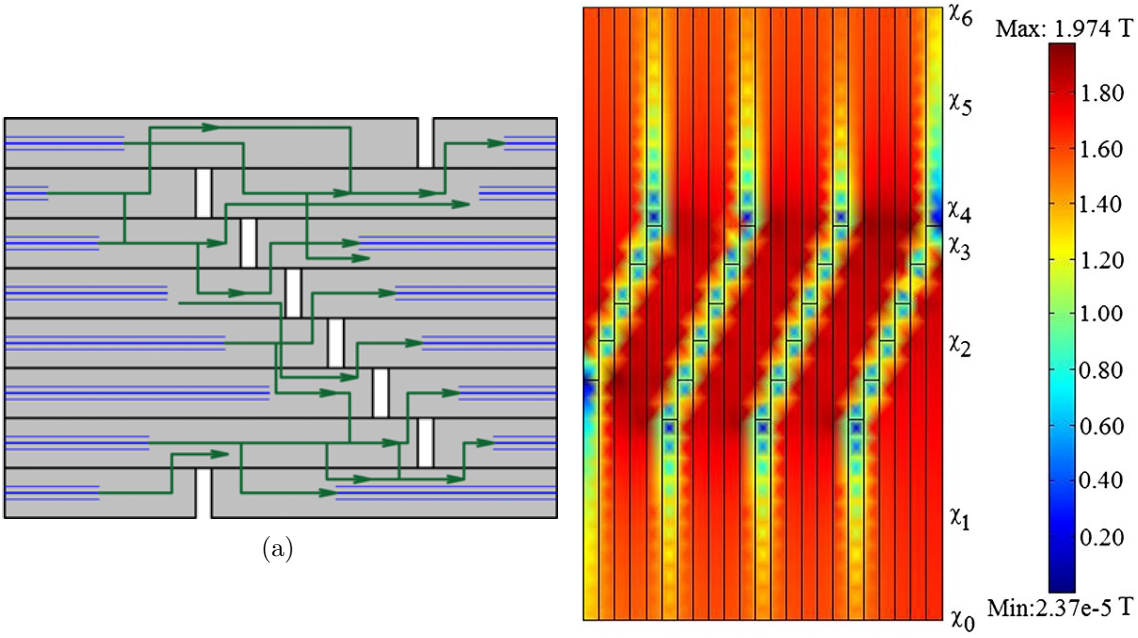


Fig. 3.3: (a) Redistribution of magnetic flux near joints of core laminations, adapted from [18]. (b) Increase of magnetic flux density near joints, adapted from [54].

in Fig. 3.4(a). In addition to their mechanical role, these holes are essential during core assembly to ensure precise step-lap joint alignment, as described in [56].

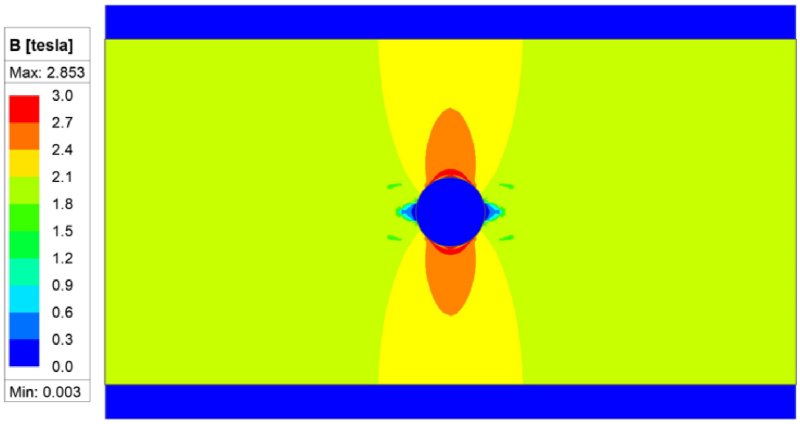
However, according to [57], such perforations cause local flux crowding, which leads to increased core losses that scale with the hole diameter, geometry, and relative positioning, as illustrated in Fig. 3.4(b). The impact is particularly significant in smaller cores, where the ratio between the lamination width and the hole diameter is higher, resulting in a greater distortion of the magnetic flux. Conversely, in larger cores (typically in high-power transformers), the relative influence of stacking holes becomes negligible, as confirmed in [58].

It is important to note that in wound-type cores, stacking holes are not present, and thus no additional losses from this effect occur. Nevertheless, for laminated step-lap cores, careful consideration of stacking-hole diameter and location remains an important design factor to limit localized flux disturbances and avoid unnecessary increases in iron losses.

3.3.4 Rotational Losses

In common three-phase core-type transformers, a significant source of additional core losses is observed in the T-joints, located at the junctions between the central limb and the yokes, as documented in [59], [60]. Within these regions, the magnetic

(a)



(b)

Fig. 3.4: (a) Finished limbs and yokes of the transformer core on the pins from the cutting line, adapted from [18]. (b) Magnetic flux density distribution around a stacking hole, adapted from [58].

flux vector undergoes cyclic variations not only in its magnitude but also in its spatial orientation throughout each excitation period, leading to rotational magnetization effects. Such changes in flux direction, clearly illustrated in Fig. 3.5, where the critical regions are highlighted with blue circles, cause localized increases in eddy currents and hysteresis losses that are not captured by simplified one-dimensional core loss models. The strong anisotropic magnetic properties of CRGO electrical steel make these losses particularly pronounced, as optimal magnetic performance is only achieved along the rolling direction. Consequently, the flux misalignment at T-joints introduces significant localized heating and a non-negligible contribution to the overall no-load losses, making this phenomenon an important consideration in accurate transformer loss modeling and core design optimization.

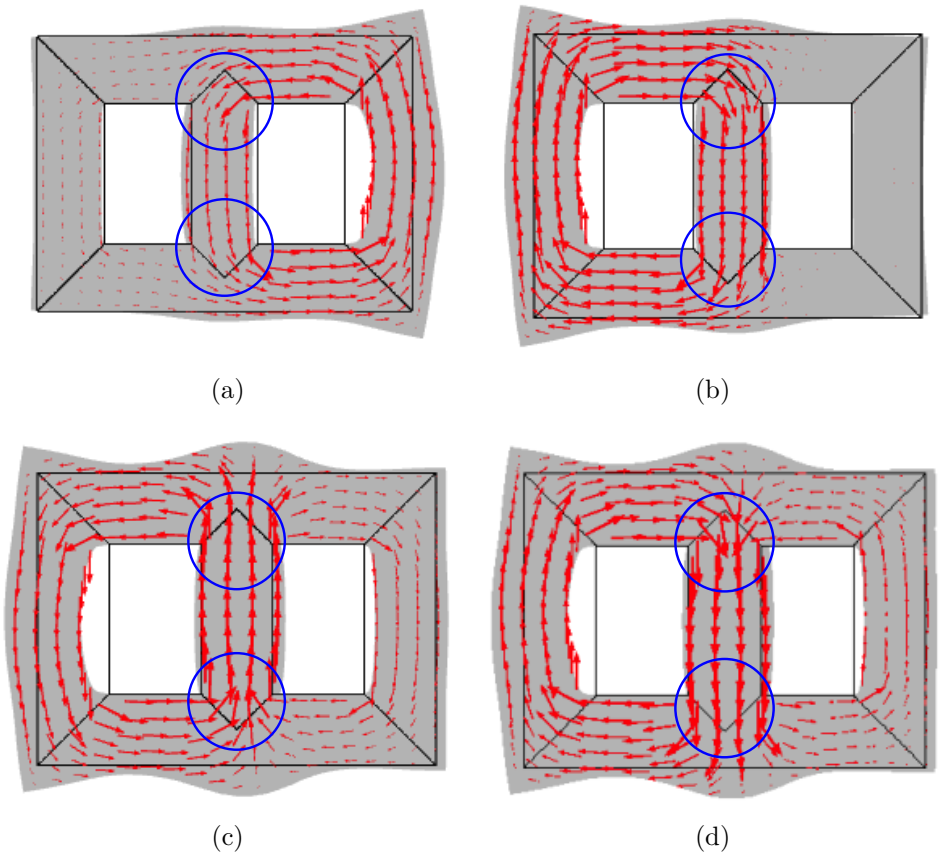


Fig. 3.5: Magnetic flux vector field with marked T-joints (blue circles) in the core during one period at times: (a) 4 ms, (b) 8 ms, (c) 12 ms, (d) 16 ms, adapted from [61].

3.3.5 Mechanical Processing of Steel Sheets

CRGO electrical steel is typically supplied in wide coils, which are then slit into narrower coils prior to lamination cutting, as illustrated in Fig. 3.6. This slitting process is carried out on dedicated slitting lines designed to achieve precise widths for transformer core construction. However, the mechanical stresses introduced during slitting and subsequent cutting can cause localized edge damage, reducing the magnetically effective width of each lamination sheet. This degradation in usable magnetic material contributes to increased core losses, as reported in [62]. Furthermore, cutting operations may produce burrs along the edges of laminations, leading to potentially conductive bridges between adjacent sheets. Such unintended electrical contacts compromise lamination insulation, exacerbate eddy current formation, and introduce unpredictable additional losses that are difficult to model accurately. These manufacturing-induced imperfections highlight the importance of precision cutting, burr minimization, and quality control in transformer core production to achieve consistent no-load loss performance.

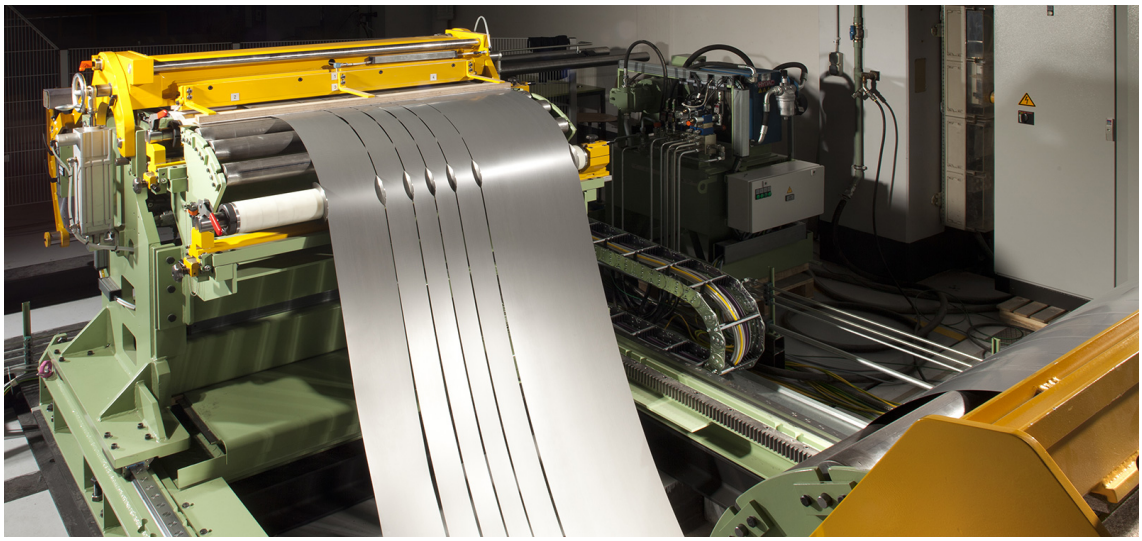


Fig. 3.6: Slitting machine for electrical steel, adapted from [63].

3.3.6 Steel Grade Tolerance

Although manufacturers guarantee that the magnetic properties of electrical steel remain within specified tolerances, practical material quality often exhibits measurable variability from coil to coil as confirms [64]. This variation arises due to factors such as localized inconsistencies in steel processing, annealing, and coating application during manufacturing. As continuous quality control along the entire length of a coil is not always feasible, specific core loss values may occasionally exceed expected limits, even when the material formally meets specification criteria. Consequently, transformers produced from nominally identical designs and material grades may demonstrate slight but noticeable deviations in no-load loss performance. Such variability underscores the importance of careful material selection, coil traceability, and statistical evaluation of steel quality to ensure consistent transformer efficiency.

3.3.7 Accuracy of Core Assembly

Human factors during core assembly, particularly in the precise placement and alignment of yoke packets in step-lap cores, represent a significant source of variability in measured no-load losses as reported in [65]. In the case of step-lap cores, these are usually assembled on special tables, one of which is shown in Fig. 3.7, where the limbs are already attached in their final form and the yokes are inserted in smaller groups of laminations. If these tables are not available or if the core is too large for vertical assembly, the magnetic circuit is built in a horizontal position, with the laminations of limbs and yokes stacked gradually, step by step.

Fig. 3.7: Core building performed on a special rotating hydraulic table for vertical assembly, adapted from [18].

Even minor deviations in the positioning of laminations or misalignment of step-lap joints can alter the intended magnetic flux paths, leading to localized increases in flux density and elevated losses. Additionally, factors such as edge quality, sheet cleanliness, and precision of insertion during stacking can influence interlaminar insulation integrity and mechanical stress distribution within the core. These subtle assembly-related discrepancies make it challenging to achieve perfectly consistent core performance across transformers of identical design, further contributing to uncertainty in predicted no-load loss values.

3.4 Summary of Chapter 3

This chapter provided a detailed technical review of transformer no-load losses, emphasizing their fundamental dependence on the magnetic core design, material characteristics, and manufacturing processes. The discussion began with a systematic breakdown of the core loss components, namely hysteresis, eddy-current, and anomalous losses, and explored their variation with frequency, flux density, and material anisotropy. These physical mechanisms form the basis for core loss evaluation and serve as a foundation for practical design calculations.

Industry-standard analytical approaches for estimating no-load losses were then introduced, focusing on methods that use the core mass, manufacturer-provided specific loss data, and empirically derived correction factors to estimate the final

assembled-core performance. This approach is widely applied in transformer design due to its simplicity and reasonable accuracy for preliminary calculations. However, its reliance on fixed correction factors and engineering experience limits its ability to capture subtle variations introduced by manufacturing tolerances and assembly techniques.

A comprehensive investigation of additional no-load losses was presented, identifying predictable effects such as flux density irregularities, lamination joint losses, stacking hole effects, and rotational magnetization at T-joints. These predictable mechanisms were contrasted with factors of greater variability, including steel processing damage, material property tolerances, burr formation, and precision of core assembly. The chapter showed how these influences interact in complex ways, often resulting in loss deviations between nominal design values and measured performance, even for transformers of identical design.

This review highlighted the practical limitations of traditional deterministic approaches, which remain dominant in industry but lack the granularity to model all sources of variation accurately. The findings support the use of advanced numerical modeling techniques, such as finite-element simulations, as well as emerging data-driven and machine learning methods, to improve predictive accuracy. This chapter thus establishes a technical baseline for transformer optimization studies, illustrating how a deep understanding of loss mechanisms can inform improved design methodologies and drive future advancements in high-efficiency transformer development.

4 Optimization and Machine Learning Methods

This chapter outlines the theoretical background and mathematical foundations of the optimization algorithms applied in this thesis to optimize transformer designs. The primary goal of the optimization is to identify configurations that minimize production costs while satisfying all constraints related to required electrical parameters and manufacturing feasibility, without any human interaction during the process, thereby enabling full automation.

Additionally, the chapter introduces the essential concepts of machine learning, with a particular focus on surrogate modeling utilizing Gaussian Process Regression (GPR). GPR is later employed in the thesis as a core technique for more accurate estimation of transformer no-load losses, offering a significant improvement over conventional methods, such as widely used analytical approaches.

4.1 Mathematical Optimization

Optimization problems in the electrical engineering domain typically involve multiple constraints that must be satisfied while minimizing an objective function. These problems are often characterized by intricate dependencies between variables and a solution space that may be highly non-convex or even discontinuous. In the case of distribution transformers, design optimization presents an even greater challenge due to the highly nonlinear nature of the underlying mathematical model. Furthermore, the design space consists entirely of discrete or discretized variables, reflecting the fact that real transformer components—such as core sheet dimensions, winding conductors, or insulation thicknesses—are selected from standardized manufacturing options rather than continuous ranges.

The constraints in the investigated optimization problem are diverse and domain-specific. They may include electromagnetic requirements (such as no-load losses, load losses and magnetic flux limits), material availability (linked to procurement or manufacturing limitations), spatial constraints (reflecting physical layout and tank dimensions), thermal performance limits (to ensure safe operating temperatures), and overall cost constraints (covering materials, labor, and manufacturing processes). Meeting all these constraints simultaneously while optimizing performance or cost metrics forms a complex and high-dimensional design space that cannot be navigated through brute-force search or manual tuning.

4.1.1 Constrained Optimization

Constrained optimization (CO) offers a structured mathematical framework to address these challenges. It allows the design space to be systematically explored while ensuring that all predefined physical and practical requirements are satisfied.

Formally, a constrained optimization problem can be defined as the task of minimizing an *objective function* $f(\mathbf{x})$ of one or more design variables $\mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{x} \in \mathcal{C}, \tag{4.1}$$

where $\mathbf{x}^* \in \mathbb{R}^n$ is the *optimal solution*, and $\mathcal{C}$ denotes the set of all feasible solutions defined by the system of design *constraints*—including both equality and inequality conditions.

4.1.2 Optimization Methods

Approaches to solving CO problems are generally classified into two categories: exact methods and heuristic methods. Exact methods, such as integer nonlinear programming (INLP) or mixed-integer programming (MIP), provide mathematically rigorous solutions by exhaustively searching or pruning the feasible region. While these methods are well-suited to small-scale or convex problems, they become computationally infeasible in the context of large, discrete, and highly nonlinear systems like transformer design. The complexity of the model, the combinatorial explosion of variable combinations, and the presence of multiple local minima severely limit the applicability of exact methods in this domain.

Heuristic approaches, on the other hand, offer greater adaptability and scalability. Algorithms such as genetic algorithms, particle swarm optimization, and tabu search, described in [66], rely on problem-specific knowledge, randomness, or memory-based exploration to efficiently search large solution spaces. However, these methods often require fine-tuning of numerous hyperparameters (e.g., mutation rates, population size, search depth), and their performance can vary significantly depending on the initialization and configuration.

4.1.3 Simulated Annealing

Simulated annealing (SA), introduced in [67], is a probabilistic metaheuristic algorithm inspired by the physical process of annealing in metallurgy. During annealing, a material is heated to a high temperature and then cooled gradually under controlled conditions. This slow cooling allows atoms within the material to explore various configurations and settle into a stable, low-energy crystalline state. SA transfers this physical principle to mathematical optimization by allowing the algorithm

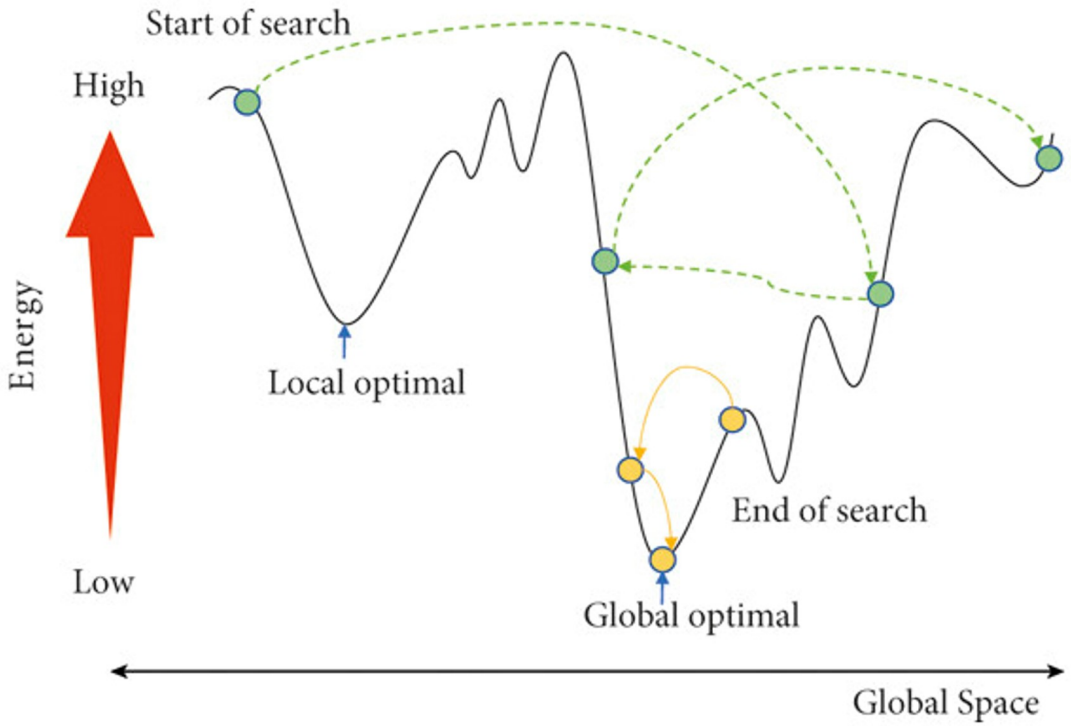


Fig. 4.1: Principle of the simulated annealing algorithm, adapted from [68].

to probabilistically accept worse solutions at the beginning of the process—when the temperature is high—and become increasingly conservative as the temperature decreases, as illustrated in Fig. 4.1.

This probabilistic acceptance of inferior solutions is key to avoiding local optima and exploring the solution space more thoroughly, especially in complex optimization landscapes. As such, SA is well-suited for challenging engineering problems involving nonconvex, nonlinear, or discrete solution spaces. These characteristics align closely with the transformer design optimization task, which features numerous interdependent parameters and hard constraints related to electromagnetic behavior, material use, physical dimensions, thermal limits, and economic viability.

Formally, SA proceeds in discrete time steps, where at each iteration t, a new candidate solution $\mathbf{x}'$ is evaluated relative to the current solution $\mathbf{x}_t$. The probability P_t of accepting this new candidate is defined as:

$$P_t = \begin{cases} 1 & \text{if } f(\mathbf{x}') < f(\mathbf{x}_t), \\ e^{-\frac{f(\mathbf{x}')-f(\mathbf{x}_t)}{T_t}} & \text{otherwise,} \end{cases} \tag{4.2}$$

where T_t is the temperature parameter at iteration t. As the algorithm progresses, T_t gradually decreases according to a predefined cooling schedule (e.g., exponential

decay), reducing the likelihood of accepting poorer solutions and thereby refining the search around promising regions. The initial solution $\mathbf{x}_0$ is typically selected at random or based on an educated guess from domain knowledge.

For transformer design, applying SA requires careful integration of constraint handling, since many feasible solutions must satisfy strict regulatory or manufacturing requirements. To address this, a penalty-based approach is employed, where constraint violations are embedded into the objective function through a weighted penalty term. Each constraint k is associated with a nonnegative penalty function $c_k(\mathbf{x})$, which quantifies the severity of its violation. The corresponding weight w_k determines the relative importance of the constraint and normalizes differences in scale. This results in the modified objective function:

$$f'(\mathbf{x}) = f(\mathbf{x}) + \sum_k w_k \cdot c_k(\mathbf{x}), \tag{4.3}$$

where $f'(\mathbf{x})$ is the penalized function used by SA during optimization. If a solution satisfies all constraints, its penalty terms vanish and the objective reduces to the true cost function. If any constraint is violated, a penalty is applied proportional to the violation magnitude and its associated weight.

This approach effectively transforms the original constrained optimization problem into an unconstrained one, simplifying the application of SA while maintaining feasibility compliance. The penalty functions can be linear, quadratic, or even exponential, depending on how strictly violations need to be discouraged.

In the context of transformer design, the SA framework can be tailored to prioritize different subsystems of the transformer—such as core dimensions, winding geometries, or tank structure—at various stages of optimization. This staged optimization strategy enhances convergence and reflects the practical dependencies of real-world transformer design processes.

Additionally, SA is highly compatible with parallel computing. Multiple independent optimization runs, each with different random initializations, can be performed simultaneously to explore the solution space more broadly. This parallelization increases the chance of discovering globally optimal or near-optimal designs, particularly when high-performance computational resources are available.

Compared to traditional manual design methods, SA offers a systematic, repeatable, and fully automatable approach that reduces human error and design cycle time. As demonstrated in this work, its flexibility and robustness make it a strong candidate for integration into the transformer design process.

4.1.4 Integration with Surrogate Modeling

An accurate, fast, and computationally efficient estimation of no-load losses plays a crucial role in the transformer design process, particularly when integrated into optimization algorithms such as Simulated Annealing (SA). The significance of computational cost becomes even more pronounced in such iterative optimization loops, where thousands of evaluations may be required. However, none of the known conventional methods, as described in Chapter 3, fully satisfy all these requirements simultaneously—either lacking in precision, speed, or general applicability to the wide range of transformer designs.

Consequently, a new approach was necessary to meet the demands of both accuracy and computational efficiency. This led to the development of a novel method based on machine learning, specifically leveraging surrogate modeling to estimate no-load losses. Since surrogate models inherently approximate complex functions with reduced computational effort, they are well-aligned with the nature of mathematical optimization problems. The application of machine learning in this context offers a practical and innovative solution, bridging the gap between detailed physical modeling and high-speed evaluation. This method is further elaborated in the following section.

4.2 Surrogate Modeling

For the development of a surrogate model, it is essential to have both a set of training data (also referred to as observations) and an appropriate machine learning technique. The training data can be obtained either by conducting numerical simulations or through real-world measurements and experiments. The latter approach, while often more accurate, tends to be both restrictive and costly. Consequently, when only a limited amount of data is available, the choice of the most suitable machine learning method becomes particularly critical.

Machine learning techniques differ significantly in their ability to handle varying quantities of data, their sensitivity to the distribution of data samples, and their adaptability to different data characteristics (e.g., linear versus nonlinear behavior). According to the classification presented in [69], [70], the following techniques are commonly considered for surrogate modeling under such constraints:

- artificial neural networks (ANN),
- decision tree regression (DTR),
- support vector regression (SVR),
- random forest regression (RFR),
- multilayer perceptron (MLP),

- k-nearest neighbors algorithm (KNN),
- extreme gradient boosting (XGBoost), and
- Gaussian process regression (GPR).

Among these, GPR was identified in this thesis as the most suitable approach, primarily due to its flexibility in modeling nonlinear relationships and its effectiveness when dealing with sparse datasets. A more detailed explanation of the GPR method and its application in the context of this work is provided in the following section. However, a comparison with the other machine learning techniques mentioned above, applied to the investigated problems in this thesis, is also presented later in Section 6.6.

4.2.1 Gaussian Process Regression

Gaussian process regression (GPR), also known as Kriging, is a non-parametric, supervised machine learning method widely used for surrogate modeling [71], [72]. It is grounded in the theory of *Gaussian Processes*, which are collections of random variables, any finite subset of which follows a joint Gaussian distribution. A Gaussian Process is fully defined by its mean and covariance functions [73]:

$$f(x) \sim \mathcal{N}(\mu, \sigma) \sim \mathcal{N}(\mu, K), \tag{4.4}$$

where $f(x)$ denotes the function to be modeled, $\mathcal{N}$ represents the Gaussian (normal) distribution, and μ and σ refer to the mean and covariance functions, respectively. The covariance function is commonly referred to as the *kernel*, denoted as K.

4.2.2 Kernels

Various kernel functions can be employed in GPR, each affecting the model's flexibility and ability to capture specific data characteristics. The most commonly used kernels as describes [74] include:

- Matérn kernel,
- radial basis function (RBF) kernel,
- rational quadratic (RQ) kernel,
- linear kernel,
- polynomial kernel,
- piecewise polynomial kernel,
- periodic kernel, and
- Hamming distance kernel.

Among them, the *Matérn kernel* [75], [76] is generally recommended as a robust first choice, primarily due to its ability to model non-smooth functions and to interpolate real-world phenomena effectively. As demonstrated later in Section 6.3.2, this kernel proves to be the most suitable for modeling the investigated problems in this thesis. The Matérn kernel is characterized by a smoothness parameter ν, with typical recommended values being $\nu = 1/2$, $3/2$, $5/2$, and ∞. The general form of the Matérn kernel is expressed as:

$$K(x_i, x_j) = \sigma_{\mathrm{f}}^2 \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\frac{\sqrt{2\nu}}{l} d(x_i, x_j) \right)^{\nu} \cdot K_{\nu} \left(\frac{\sqrt{2\nu}}{l} d(x_i, x_j) \right), \tag{4.5}$$

where $d(x_i, x_j)$ denotes the distance between two input samples x_i and x_j, σ_{f} represents the scaling factor for the covariance, l is the lengthscale hyperparameter, and the previously mentioned ν determines the smoothness of the resulting function ($\nu > 0$).

The analyzed problem involves mixed-input variables, including both continuous and categorical components. While the Matérn kernel is well-suited for modeling continuous inputs, it does not effectively capture relationships between categorical variables. To address this limitation, the *Hamming distance kernel* was employed for categorical features such as material types or inputs defined by discrete categories.

The Hamming distance kernel provides a simple similarity measure for categorical inputs by evaluating how many features exactly match between two input samples. The kernel is defined according to [77] as:

$$K(x_i, x_j) = \sum_{k=1}^{d} \delta(x_i^{(k)}, x_j^{(k)}), \tag{4.6}$$

where x_i and x_j are two input vectors of dimension d, and $\delta(x_i^{(k)}, x_j^{(k)})$ is the Kronecker delta function given by

$$\delta(x_i^{(k)}, x_j^{(k)}) = \begin{cases} 1, & \text{if } x_i^{(k)} = x_j^{(k)}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.7}$$

where the superscript (k) in $x_i^{(k)}$ and $x_j^{(k)}$ denotes the k-th feature (or dimension) of the input vectors. This kernel counts the number of matching categorical values across all input dimensions, effectively quantifying similarity via inverse Hamming distance. It is positive semi-definite and commonly used in Gaussian process regression for problems involving discrete or categorical input spaces.

4.2.3 Bayesian Inference and Posterior Prediction

The choice of kernel corresponds to specifying a prior distribution over the space of functions—the initial belief about the behavior of the system being modeled.

Once a prior is established, the likelihood function is defined to relate observed data to the underlying function. In most cases, a Gaussian likelihood function is assumed according to [78]:

$$p(y|f(x)) \sim \mathcal{N}(f(x), \sigma_n^2 I), \tag{4.8}$$

where $p(y|f(x))$ is the likelihood of observing y given $f(x)$, σ_n^2 is the noise variance, and I is the identity matrix.

By incorporating observed data, the prior is updated to form the posterior distribution, which provides improved function estimates [73]. The posterior mean and covariance are:

$$\begin{aligned} \mu_* =& \mu(x_*) + K_*^T (K + \sigma_n^2 I)^{-1} (y - \mu(x)), \\ \sigma_* =& K_{**} - K_*^T (K + \sigma_n^2 I)^{-1} K_*, \end{aligned} \tag{4.9}$$

where μ_* and σ_* denote the posterior mean and covariance; x_* is a new input; K_* is the covariance between training and new data; and K_{**} is the covariance among the new data points.

To make accurate predictions, the GPR model must be trained by optimizing its hyperparameters, particularly those of the kernel function. This process typically uses gradient-based optimization techniques to minimize the discrepancy between predicted and observed values, as stated in [79], [80]. An overview of this process is illustrated in Fig. 4.2.

4.2.4 Strengths and Limitations of GPR

GPR models are highly favored for surrogate modeling due to their flexibility and strong capability to approximate complex functions using limited training data [81]. A key strength of GPR is its built-in uncertainty quantification, which allows it to provide not only predictions but also confidence estimates.

GPR is also relatively lightweight in terms of parameterization: only a small number of kernel-related hyperparameters need to be optimized. This contrasts with models like Artificial Neural Networks, which require tuning many more parameters and are often more computationally expensive to train.

Models trained using GPR are widely appreciated for their adaptability, capacity to approximate a broad spectrum of functions with limited data, and their inherent ability to quantify prediction uncertainty [81]. In contrast to artificial neural networks, GPR models are considered as non-parametric, implying that only a small number of hyperparameters are tuned during training. A key limitation of these models lies in their scalability with respect to both dataset size and input dimensionality. As

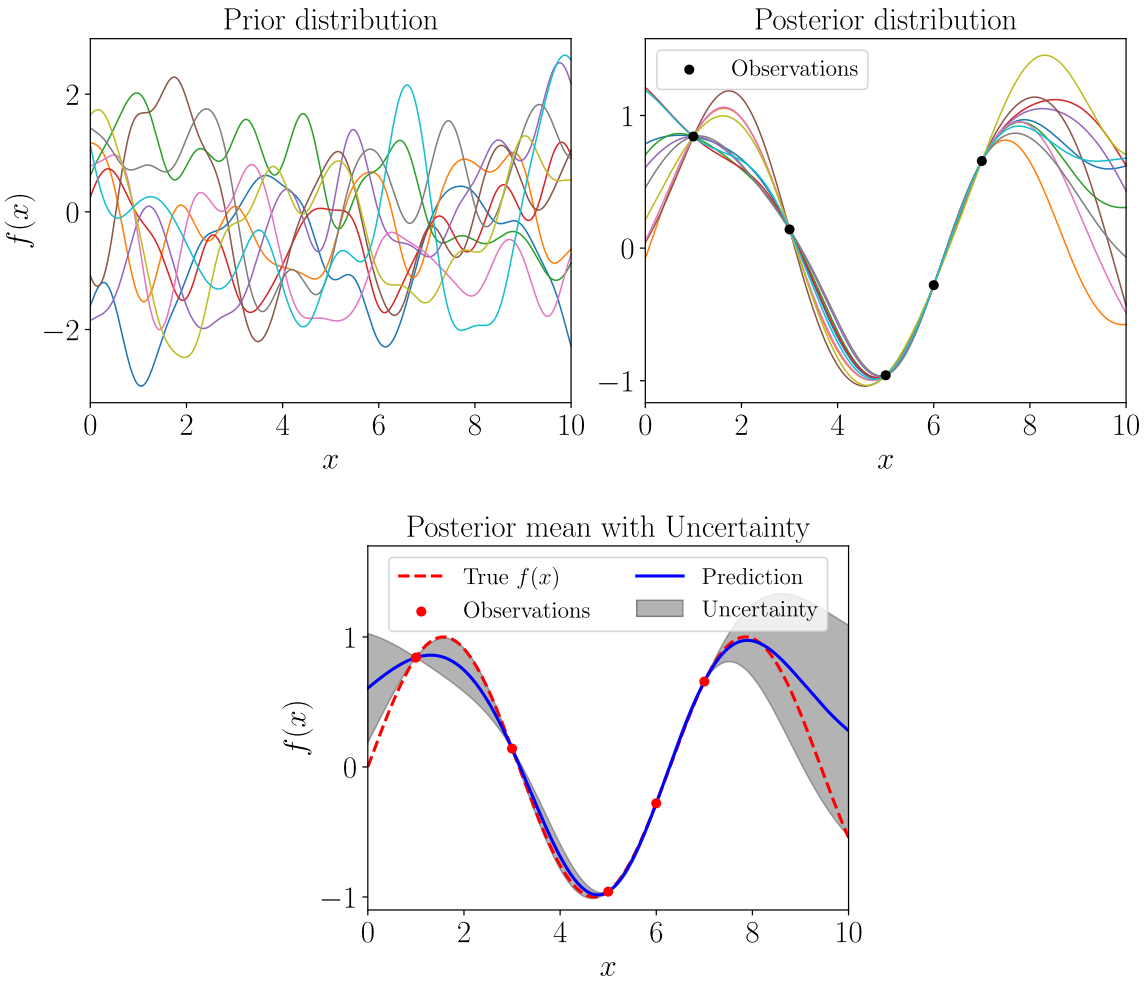


Fig. 4.2: Example of Gaussian process regression modeling of a simple 1D function shown in three steps, adapted from [8].

highlighted in [74], GPR methods typically scale well only up to about 2,000 training instances due to the high computational cost associated with inverting the covariance matrix σ, which requires $O(n^3)$ operations and $O(n^2)$ memory, where n is the number of training points, as discussed in [82]. Furthermore, they are generally best suited for problems with a moderate number of input features—typically no more than 20 as reported by [83].

4.2.5 Effect of Training Data Distribution

The accuracy of GPR models strongly depends on both the distribution and quantity of training data. Even a well-tuned model will yield poor predictions if the training samples are sparse or unevenly distributed. Fig. 4.3 demonstrates this effect with three 2D examples showing different training sample scenarios.

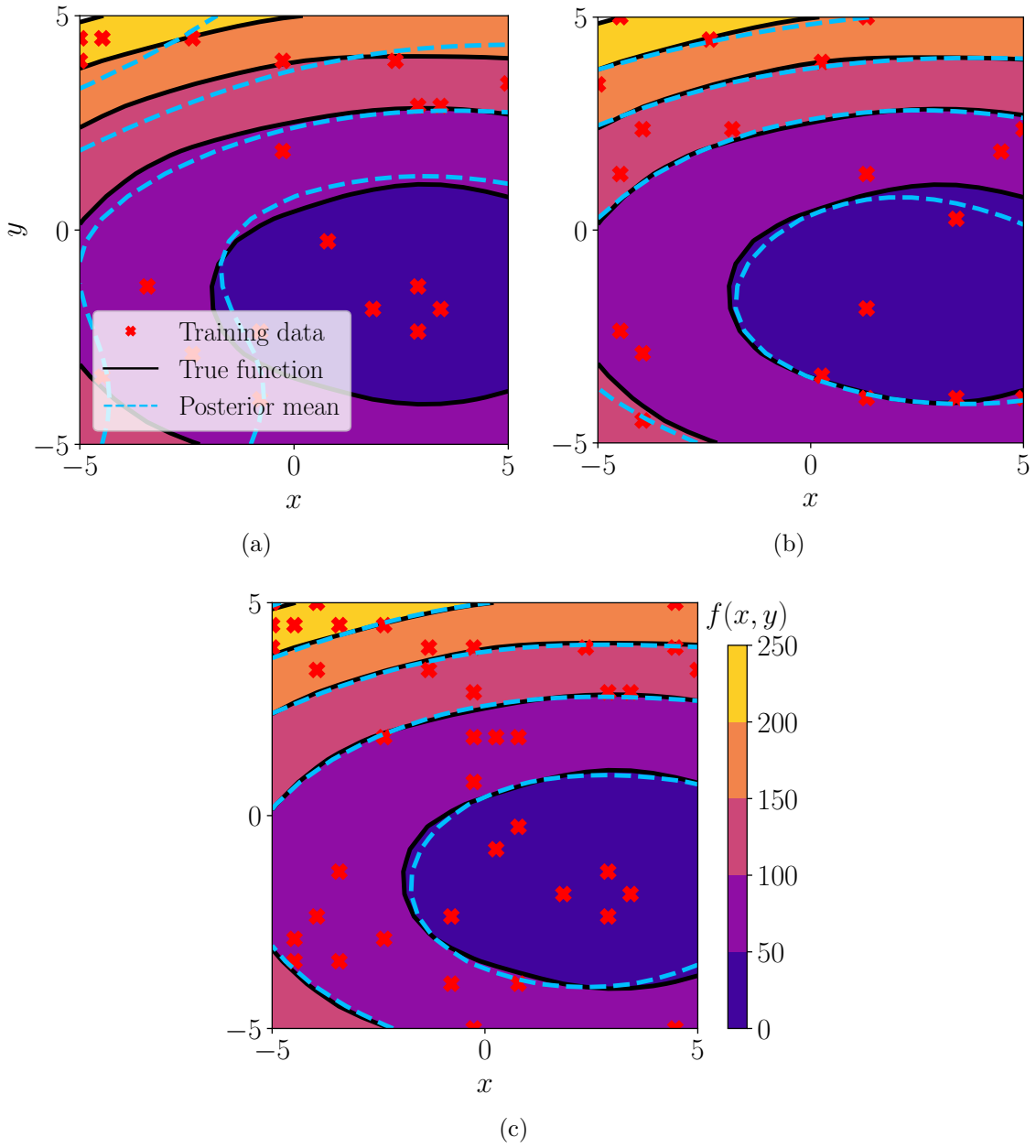


Fig. 4.3: Example of 2D Gaussian process regression modeling for (a) inappropriate distribution of training data samples, (b) more evenly distributed training data samples, and (c) suitable training data distribution with doubled samples count.

4.3 Summary of Chapter 4

This chapter presented the mathematical and theoretical underpinnings of the optimization and machine learning techniques employed in this thesis to automate and enhance transformer design. The optimization component was centered on constrained optimization, with a focus on Simulated Annealing (SA) as the primary

algorithm. SA was selected for its robustness, adaptability to discrete and nonconvex design spaces, and suitability for integration with domain-specific constraints through a penalty-based formulation. Its flexibility, combined with the ability to perform parallel evaluations, makes it particularly effective for complex engineering tasks such as transformer design, where numerous interdependent constraints and variables must be handled concurrently.

To address the computational demands of evaluating critical performance metrics—most notably transformer no-load losses—this chapter introduced surrogate modeling based on machine learning. Among various techniques evaluated, Gaussian Process Regression (GPR) was identified as the most effective due to its strong performance with sparse data, built-in uncertainty quantification, and minimal hyperparameter tuning requirements. The use of the Matérn kernel and Bayesian inference techniques enables GPR to produce accurate and interpretable predictions, which are crucial for guiding the optimization process efficiently.

Together, these methods form a cohesive framework that supports automated and data-driven transformer design. Optimization and surrogate modeling need not be treated as isolated tools; instead, they can function as tightly integrated components, with GPR models that enable faster and more accurate evaluations within the SA algorithm. This synergy significantly reduces design time, eliminates the need for manual tuning, and results in more cost-effective and technically compliant transformer configurations. The next chapter demonstrates how these methods are implemented and validated through experimental results.

5 Transformer Design Process Automatization and Optimization

As presented in the introduction of the thesis, according to [5], recent regulations have significantly impacted transformer prices in the European market, as reducing losses typically requires more aluminum, copper, and transformer steel—materials that directly affect cost. Beyond increased prices, stricter efficiency requirements have also led to larger and heavier transformer designs, sometimes complicating replacements in existing installations built for smaller units.

To comply with new standards, transformer manufacturers throughout the EU had to recalculate their entire product range, many of which had been validated through years of practical application. This process often involved hundreds of recalculations in a short period, placing a heavy burden on manufacturers—particularly given the shortage of experienced engineering staff. With at least one more recalculation round expected, the need for more efficient design workflows is increasingly evident.

Material price volatility, especially for critical resources like electrical steel [84], further emphasizes the benefit of regularly recalculating transformer designs to maintain cost-effectiveness while meeting Ecodesign standards.

However, most manufacturers still rely on manual or semi-automated tools, such as [19], where the outcome depends largely on designer experience and available time. To address these limitations, this chapter introduces a fully automated transformer design method using the Simulated Annealing algorithm (see Section 4.1.3), which is particularly effective for solving the discrete, nonlinear optimization problems that characterize transformer design in a manufacturing context.

5.1 Incorporating the SA Algorithm into the Transformer Design Process

Replacement of the human designer in the transformer design workflow by an automated SA optimization algorithm represents a significant shift in how modern distribution transformers can be engineered. This replacement is schematically illustrated in Fig. 5.1, where the upper loop, highlighted in the yellow rectangle and labeled *Optimization algorithm*, encapsulates the automation of design iteration and decision-making. In contrast to conventional methods that depend on the experience and intuition of an engineer, the SA-based framework allows for systematic and objective evaluation of transformer configurations.

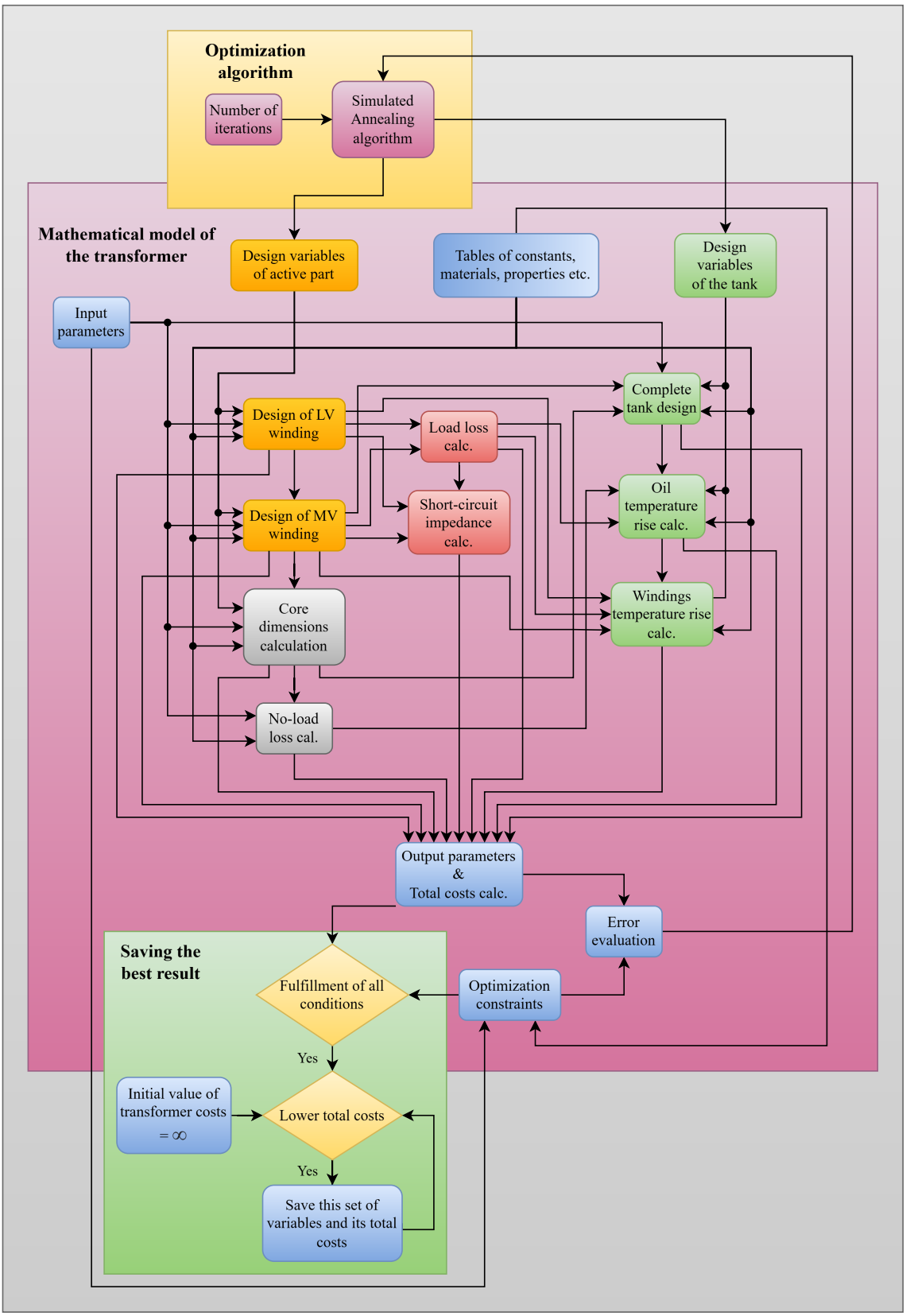


Fig. 5.1: Block diagram of the Simulated Annealing optimization algorithm for transformer design.

The rectangle labeled *Mathematical model of the transformer* in the figure is practically based on Fig. 2.1, whose function is fully described in Chapter 2, with only one modified block, *Error evaluation*, whose purpose is explained later in Section 5.1.2.

5.1.1 Algorithm Inputs and Evaluation Framework

In addition to standard input parameters—such as rated power, voltage levels, connection group, short-circuit impedance or load and no-load losses—the optimization algorithm also requires an additional control input: the total number of iterations to be executed. This parameter serves as a balance between computation time and solution quality. A higher number of iterations typically increases the likelihood of discovering a global optimal or near-optimal design, but naturally comes at the cost of greater computational resources and time. Conversely, a lower number of iterations shortens the runtime but may compromise the optimality or feasibility of the final solution.

As the algorithm progresses, each candidate design generated by the SA mechanism is subjected to a sequence of evaluation steps. These steps calculate the behavior of a physical transformer and yield relevant electrical, thermal, and geometrical output data. Once these outputs are obtained, the algorithm computes the total material cost associated with the candidate design, typically including the cost of copper and aluminum in windings, electrical steel, insulation paper, oil, tank, and other structural components.

Following cost evaluation, the resulting outputs are tested against a set of predefined design constraints. These constraints may reflect both regulatory requirements (e.g., loss and peak efficiency index according to standards), physical limitations (e.g., maximum dimensions), and manufacturing capabilities (e.g., material availability, limits of manufacturing machines). If a candidate design satisfies all constraints and achieves a material cost less than the best solution found in previous iterations, it is accepted as the new incumbent. This decision process is managed by a dedicated internal loop, illustrated in Fig. 5.1 inside the green rectangle labeled *Saving the best result*, which continuously tracks and updates the current optimal solution along with the corresponding design parameters.

5.1.2 Constraint Penalty and Search Dynamics

A fundamental part of this decision-making process is the application of a cost penalty mechanism, in the Fig. 5.1 referred to as block Error evaluation. The basic principle of this penalty-based approach is presented in Section 4.1.3. When a candidate design violates one or more constraints, its base material cost is augmented by an error value. This value reflects the magnitude of deviation from the required specification,

thereby discouraging the algorithm from accepting non-compliant designs. The impact of these penalties can be fine-tuned through the use of recalibration factors, which allow the designer to assign relative weights to different constraints. For instance, a violation of thermal limits may be penalized more heavily than a slight excess in geometric dimensions if the former has more severe implications for long-term performance or safety.

Regardless of the outcome—whether the design is feasible, partially feasible, or infeasible—the final adjusted cost (i.e., base cost plus penalty) is fed back into the SA algorithm. This feedback loop enables the algorithm to probabilistically explore the design space in a manner that balances exploitation of known good areas with exploration of new, potentially better regions. The selection of new candidate solutions is governed by the Metropolis criterion according to [85], which permits occasional acceptance of worse solutions to escape local minima—a key feature that makes SA particularly well-suited for complex, nonlinear, and discontinuous search spaces like that of transformer design.

5.1.3 Termination and Strategy Overview

This iterative process continues until the specified number of iterations is completed (convergence criteria—minimal cost improvement over several iterations—cannot be applied in this case, as any expected value does not exist). Upon termination, the algorithm returns the best candidate solution identified throughout the search, including all associated variable values and output results. However, if no valid solution is found due to infeasible input parameters or overly restrictive constraints, the algorithm terminates with an error message indicating the absence of a feasible region. In this case, a soft cap on iterations can be applied: the process halts if a feasible solution is not found within a predefined subset of the total iteration budget, thus avoiding excessive computation on unpromising configurations.

In essence, this optimization strategy automates the entire transformer design process by replacing heuristic and experience-driven decisions with a structured, mathematically grounded search. It enhances repeatability, scalability, and speed—while ensuring compliance with modern design constraints and material cost considerations. Moreover, by embedding evaluation, penalization, and memory in the algorithm internal mechanics, it provides a robust platform for cost-effective high-performance transformer designs in both small-scale recalculations and large-scale industrial redesign initiatives. The effectiveness of this algorithm, compared to conventional manual calculations, is evaluated in the following section.

Table 5.1: Technical Specification of Selected Transformers for Optimization Comparison

Rated power	**kVA**	**50**	**630**	**1,600**
Voltage ratio	V	22,000 / 400		
Insulation levels (U_m; AC; BIL)	kV	25; 50; 150 / 1.1; 3; -		
Connection group	-	Dyn1		
Short-circuit impedance	%	4	4	6
Primary winding taps	%	$\pm 2 \times 2.5$		
Maximum no-load loss	W	81	540	1,080
Maximum load loss	W	750	4,600	12,000
Frequency	Hz	50		
Cooling method	-	ONAN		
Temperature rise (oil / windings)	K	60 / 65		
Maximum ambient temperature	°C	40		
Installation altitude	m.a.s.l.	1000		

5.2 Comparison between Manual and SA Algorithmic Optimization

To comprehensively assess the effectiveness of the proposed Simulated Annealing (SA) optimization algorithm—particularly in terms of its ability to reduce material costs, shorten design time, and enhance overall design quality—a comparative analysis was conducted using three representative distribution transformers. These units, selected for their widespread use in the Czech 22kV distribution network, offer a practical basis for evaluating real-world performance improvements. The key technical specifications and operating parameters of the selected transformers are summarized in Table 5.1, providing the context for the subsequent optimization and comparison process.

5.2.1 Manual vs. Algorithmic Design Process

The design process for the selected transformers was carefully monitored and documented over time to enable a clear comparison between conventional manual calculations and the results produced by the proposed optimization algorithm. The evolution of the lowest material costs achieved by both approaches is illustrated in Fig. 5.2, which captures the trajectory of cost reduction as each method progresses through successive design iterations.

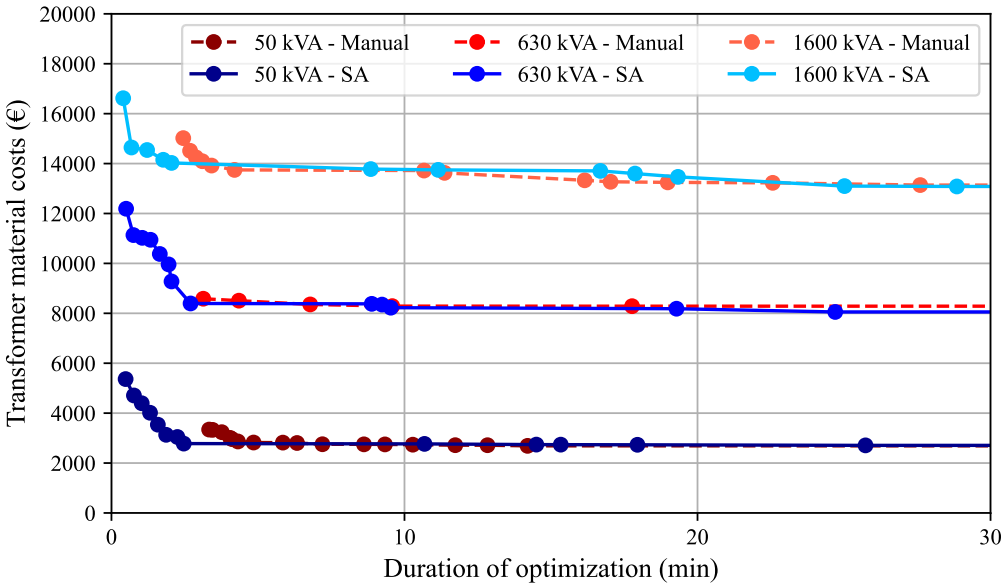


Fig. 5.2: Comparison of the lowest transformer costs over time between manual calculations and SA algorithm-based optimization for three design cases.

The manual design process was carried out by a professional transformer designer with over five years of practical experience at Elpro-Energo Transformers s.r.o. [18], utilizing proprietary design software based on the mathematical model that includes the data tables outlined in Chapter 2. Although this software automates many of the underlying calculations, the designer was still required to manually select and adjust design variables based on experience, engineering judgment, and iterative trial and error.

In contrast, automated design optimization was performed using the SA algorithm, running on a dedicated high performance computing server equipped with a 32-core Intel(R) Xeon(R) E5-2650 v2 processor operating at 2.60 GHz. This environment ensured sufficient computational resources for intensive iterative evaluations and allowed the algorithm to explore a significantly broader design space than would be practical using manual methods alone. The calculations were executed on all of the processor's cores simultaneously.

5.2.2 Performance Comparison and Convergence Behavior

The cost development trends observed during the evaluation clearly demonstrate the superior efficiency of the SA optimization algorithm compared to conventional manual calculation methods. One of the most notable advantages of the SA approach is its ability to rapidly generate an initial valid transformer design, often within

a fraction of the time required by an experienced designer using conventional techniques. This early-stage efficiency is important, particularly when time constraints are a factor in the design workflow to check if the requested design is even possible to create within technological limits.

Moreover, the convergence behavior of the SA algorithm is markedly steep in the initial phase of the optimization process. For all transformer designs evaluated in this thesis, the SA algorithm was able to approach near-optimal cost values before a complete valid solution was produced through manual calculation. This indicates a significant advantage not only in terms of speed but also in the quality of early solutions. However, as optimization progresses and the solution approaches the global minimum cost, the convergence rate of the SA algorithm naturally slows, reflecting the increasing difficulty in improving an already optimized solution.

5.2.3 Final Results and Observations

Interestingly, at this stage, manual calculations may catch up in terms of cost performance, and in the evaluated cases, they occasionally achieved marginally lower final costs during observation. Nevertheless, these improvements come at a higher time and labor cost. Around the 20-minute mark, manual calculations showed signs of stagnation, with no further meaningful reduction in cost despite continued effort. In contrast, the SA algorithm continued to explore the design space and identified new local minima even after manual optimization plateaued. Although the magnitude of these additional improvements decreased with time, the algorithm retained the potential to discover slightly better solutions if the optimization were allowed to proceed beyond the 30-minute limit.

At the end of the 30-minute comparison period, the maximum cost savings observed in favor of the SA optimization method were €56 for the 1,600 kVA transformer and € 231 for the 630 kVA transformer. For the 50 kVA unit, both the manual and automated methods ultimately reached the same final result, highlighting the fact that simpler designs with fewer variables offer limited opportunities for further optimization. These final material cost results for each transformer type examined obtained by both methods are listed in Table 5.2. The results suggest that the benefits of the SA algorithm scale with the complexity of the transformer design. Specifically, for larger-rated power transformers or custom designs that involve complex winding arrangements, auxiliary components, or strict thermal constraints, the optimization algorithm offers a substantial advantage over manual methods, which can become increasingly time-consuming and prone to human error in such circumstances.

Table 5.2: Comparison of Final Material Costs of Transformer Designs Obtained by Manual and SA Optimization

Rated power	**kVA**	**50**	**630**	**1,600**
Final manual best result	€	2,688.9	8,285.7	13,138.7
Final SA best result	€	2,688.9	8,054.4	13,082.7

5.2.4 Influence of Market Prices and Material Availability

It is important to emphasize that the optimal transformer designs identified through the optimization process are inherently sensitive to the current prices of raw materials and components. In particular, the cost of conductors, such as copper or aluminum wire and foil in windings, plays a dominant role, and their market prices are subject to continuous fluctuation. These fluctuations are primarily driven by trading activity in global commodity markets, most notably the London Metal Exchange (LME), which directly influences the daily price of these materials according to [86].

Equally significant is the cost of electrical steel, which serves as a critical component in the transformer core and greatly affects both performance and overall material expenditure. In addition to conductors and electrical steel, the prices of transformer oil and construction steel used for tank fabrication also contribute substantially to the total manufacturing cost. Variations in any of these inputs can lead to notable changes in the cost-optimal configuration produced by the optimization algorithm.

5.2.5 Volume-Based Pricing and Supplier Dynamics

Beyond the gross prices of key raw materials listed on global markets such as the LME, the desirability and cost-effectiveness of a transformer design are also significantly influenced by the specific characteristics and availability of high-turnover materials. In the case of round wire, this refers primarily to its diameter, whereas for foil or rectangular conductors, both the width and thickness play critical roles in determining production feasibility and cost. Although the LME provides a baseline for the valuation of raw materials, a considerable portion of the final price of the material is derived from the processing costs. These costs are often volume-dependent and influenced by the complexity of handling and shaping the material to meet specific design requirements.

As such, the anticipated number of transformer units associated with a given design should be considered during the optimization process. This can be incorporated as an additional input parameter that reflects the production volume expectations.

One practical implementation is to link the quantity of units with the total conductor mass required per transformer, enabling dynamic pricing models that account for bulk ordering. Material suppliers frequently offer tiered pricing structures based on the ordered quantity, making it possible to include such logic in the optimization model to reflect real-world cost variations.

However, this approach is not without complications. For many conductor types, suppliers enforce a minimum order quantity (MOQ), which can present challenges when the expected consumption of a given conductor dimension falls well below that threshold, as mentioned [87]. In such cases, any excess material remains unused, occupying valuable storage space and tying up capital in non-liquid inventory. This scenario is particularly undesirable for low-volume or custom transformer designs, where material leftovers may not be reused efficiently in future projects.

Another complexity arises from the fact that most transformer manufacturers engage with multiple suppliers for the same material types to enhance procurement flexibility and negotiate better prices or delivery terms. Unfortunately, each supplier may employ a distinct pricing structure, volume discount logic, and delivery policy, making it difficult to accurately reflect all variables in a unified cost estimation model. A feasible workaround is to apply averaged pricing data based on historical procurement or to follow the pricing logic of the most commonly used supplier to standardize cost projections.

5.2.6 Practical Implications and Conclusion

In practice, for small distribution transformers and low-volume production runs, the optimization process will likely favor high-turnover materials or dimensions already in the manufacturer's inventory, rather than introducing new conductor specifications that could lead to excessive costs or logistical inefficiencies. Conversely, for larger power transformers—where even a single unit requires significant conductor quantities—or for smaller units produced in high volumes, introducing new material dimensions may prove advantageous. In such scenarios, cost benefits from volume-based discounts and material optimization outweigh initial procurement challenges, making custom conductor dimensions a viable and often beneficial choice.

In conclusion, the results support the assertion that the SA-based optimization approach is highly effective for complex and high-power transformer designs. Its ability to explore a large and multidimensional design space systematically, combined with its independence from designer expertise, makes it a valuable tool in modern transformer engineering. Especially in industrial environments where rapid design cycles and material cost optimization are crucial, such algorithms can contribute significantly to improving both productivity and competitiveness.

5.3 Summary of Chapter 5

This chapter has introduced and demonstrated an innovative fully automated transformer design workflow based on the Simulated Annealing (SA) optimization algorithm. The motivation for this work arises from the increasing complexity of transformer design processes, which are influenced by evolving regulatory standards, rising material costs, and labor shortages in the transformer manufacturing industry.

The proposed SA-based optimization approach replaces conventional trial-and-error design methods driven by experience with an iterative algorithm capable of systematically exploring the solution space. By incorporating mechanisms for constraint checking, cost evaluation, and probabilistic exploration, the algorithm is well-suited to tackle the nonlinear, discrete, and constrained nature of transformers.

A detailed implementation of the SA into transformer calculation was presented, highlighting its modular structure: input parameters, design variables, iterative design generation, evaluation of the mathematical model, output parameters, constraint penalization, and final result selection. Special attention was paid to the penalty mechanism, which allows the algorithm to account for constraint violations in a flexible way, promoting the selection of feasible and cost-effective solutions, where the design meets the required specifications and manufacturing limits.

The performance of the SA algorithm was compared against conventional manual design processes across three representative transformer ratings (50 kVA, 630 kVA, and 1,600 kVA). The results demonstrate the high efficiency of the algorithm in quickly identifying feasible designs and its ability to reduce the total cost of the material, particularly for larger and more complex transformer types. For example, cost savings of up to € 231 were achieved for the 630 kVA unit within a 30-minute optimization window. Furthermore, the SA optimizer consistently outperformed manual methods in the speed and robustness of early convergence, even when manual designs eventually reached similar cost levels.

The practical implications of the optimization framework were also discussed, including the influence of the dynamics of the raw material market, supplier pricing structures, and minimum order quantities. These real-world considerations suggest that the effectiveness of the SA-based design process is also suitable for integrating volume-based pricing logic and supplier constraints into the model, especially based on high-volume production or size of the transformer designs.

Overall, this chapter demonstrates that the SA algorithm offers a scalable and repeatable solution for transformer design in real-world manufacturing conditions. Its independence from designer intuition, combined with the ability to incorporate complex constraints and cost models, makes it a strong candidate for widespread adoption in modern transformer engineering environments.

6 Transformer No-Load Losses Estimation Utilizing Machine Learning Solution

With respect to no-load losses, it is essential to develop a novel computational method that enables accurate estimation of these losses while maintaining sufficient speed for practical use in mass production environments. This thesis presents an innovative approach for no-load losses calculation that accounts for all the previously discussed factors contributing to additional no-load losses (ANLL).

The proposed methodology is based on machine learning techniques to create a surrogate model, utilizing a comprehensive dataset composed of no-load test results from thousands of transformer units, accompanied by their respective design parameters. The core objective of the machine learning model is to predict the percentage value of ANLL, thereby improving the precision of core losses estimation under industrial constraints.

6.1 Design of the Case Study Transformer Core

For the accurate creation of a surrogate model aimed at estimating transformer no-load losses, it is essential to clearly define the type of transformer cores used as a case study, the manufacturing technologies applied, the construction methods, and, perhaps most importantly, the geometry of the design. Each of these factors directly affects the final no-load losses measured after the completion of the manufacturing process. Therefore, this section describes the origin and characteristics of the dataset used in the investigation.

It should be emphasized that the approach introduced in this chapter is general and can be applied to all known core construction methods, provided that the input variables are appropriately chosen. Nevertheless, the data used for the modeling process must originate from cores to employ the same construction technology to ensure consistency and reliability.

The analyzed cores in this thesis were built using the step-lap construction method and in a core-type shape, as described in [88]. For three-phase transformers, the steel laminations are cut separately along the cutting line into five segments: three legs and two yokes that connect them. Each segment is manufactured with two stacking holes that are used to secure the desired core shape but simultaneously contribute to ANLL. In the step-lap method, the joints of these five parts overlap within every lamination layer, enabling the magnetic flux to bypass the small air gaps between adjacent laminations. The overlapping is divided into multiple steps, and the sequence is repeated in all subsequent lamination packets. As outlined in [89],

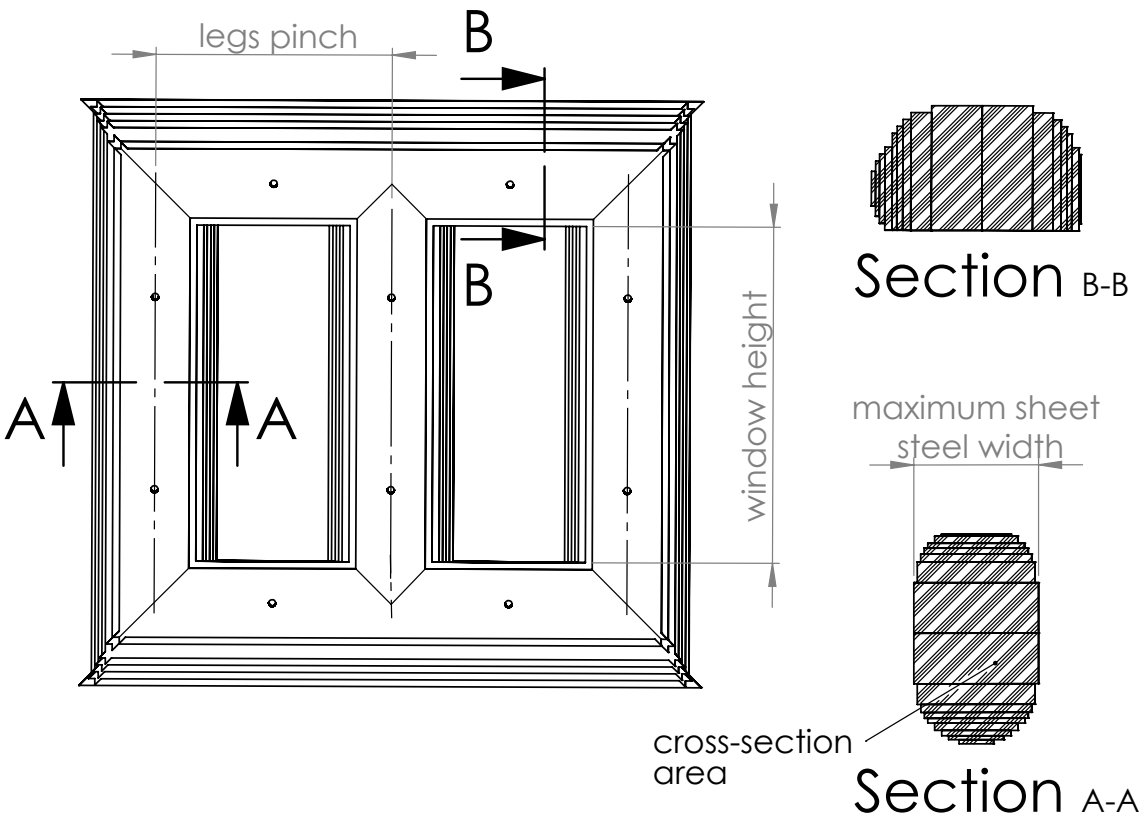


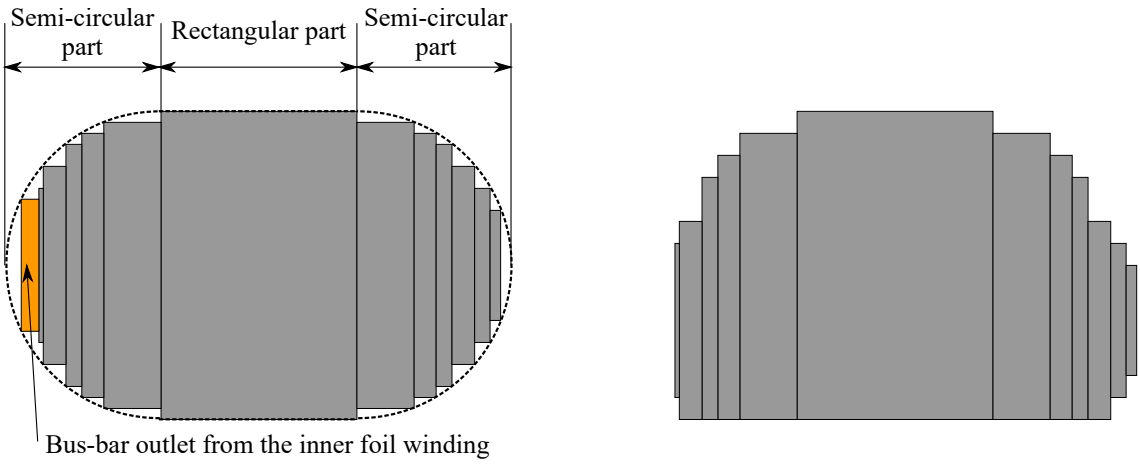
Fig. 6.1: Transformer core drawing with variables marked for surrogate modeling.

this multi-step arrangement improves the redistribution of magnetic flux and reduces the core assembly time.

All cores examined in this thesis have six step-lap steps with a standard overlap of 3.6,mm, except for laminations with a width of 50,mm, where the overlap is reduced to 2.5,mm due to cutting-line constraints. Each step consists of two identical lamination layers, forming a single packet—that is, one complete sequence of all step-lap steps—with a total of 12 laminations. The dataset includes cores with both round and oval cross-section limbs, as illustrated in Section A–A of Fig. 6.1.

The shape of the cross-section generally depends on the rated power of the transformer: oval cores are predominantly used for lower-power transformers, typically up to approximately 3150 kVA, whereas round cores are preferred for higher power ratings. The shape of the core is always carefully adapted to the design of the winding, which is determined by the forces of short circuit, the dielectric requirements, and the limitations of the winding machines.

Another notable feature of the studied cores is the use of flat yokes, where the lamination steps are stacked in alignment with the core window, and the edge of each step is defined by the widest steel sheet. This arrangement, shown in Section B–B of Fig. 6.1, helps reduce the total core weight because the laminations are slightly



(a) Limb

(b) Upper yoke

Fig. 6.2: Detailed view of (a) the core limb cross-section with geometric divisions and (b) the upper core yoke cross-section in a flat-yoke arrangement.

shorter compared to those in a conventional stepped-yoke configuration. The cores analyzed here are typically constructed using a combination of lamination steps with widths ranging from 3 to 10, allowing for a precise balance between efficient use of core space by increasing the filling factor of transformer steel and cutting time and difficulty.

A more detailed view of the cross-sections of the core limbs and yokes is presented in Fig.6.2. In many cases, the cores are not perfectly symmetric when viewed from the front and back sides. This asymmetry arises primarily when the inner winding is manufactured from foil, as the bus-bar outlets and the initial bus-bar protrude toward the inner perimeter of the winding, creating the need for additional space, as illustrated in Fig.6.2(a). Depending on the dimensions of the bus-bars, certain lamination layers—or, in extreme cases, entire lamination steps—must be removed to accommodate these components without compromising the winding's mechanical integrity.

This issue is not present in transformers where the inner winding is constructed as a multi-layer wire winding, which is typically used when the secondary nominal voltage exceeds approximately 10 kV. Wire windings allow for a more compact arrangement, eliminating the need for lamination removals and enabling more symmetrical limb cross-sections. Fig. 6.2(a) also depicts the geometric definition of oval cores. These cores can be conceptualized as a combination of three geometric elements: a semi-circular arc on one side, a central rectangular segment, and another semi-circular arc on the opposite side. In other words, the oval cross-section can be considered an extension of a round core with an inserted straight middle section, designed

to more effectively utilize the space within the active part of the transformer while maintaining an efficient magnetic flux path.

The yokes are equally influenced by the lamination configuration of the limbs. As shown in Fig. 6.2(b), any removal of lamination steps in the limbs must be mirrored in the yoke, ensuring that the stacking sequence remains consistent and magnetic continuity is preserved across the joint. Furthermore, the figure highlights that some yoke steps are intentionally misaligned with the edges of the core window. This design choice helps to achieve a more uniform distribution of mechanical forces during the disassembly or withdrawal of the supporting steel structure, thereby reducing the risk of deformation or lamination displacement.

6.2 Measured data

A critical prerequisite for the proposed machine learning-based estimation of transformer no-load losses is the availability of high-quality measured data. These measurements, which serve as training and testing samples for the surrogate model, must be accurate and representative of real-world transformer performance. In this thesis, the dataset was provided by the company *Elpro-Energo Transformers s.r.o.* [18], a manufacturer specializing in medium-voltage oil-immersed and dry-type distribution and power transformers, with ratings up to approximately 10 MVA.

The factory currently performs nearly all stages of the oil-immersed transformer manufacturing process in-house. This includes lamination cutting, core stacking and assembly (as illustrated in Fig. 6.3), foil and wire winding, oven hardening, active part assembly, winding connections, drying procedures, and final oil filling. The only exception at present is the production of transformer tanks and covers, which are supplied externally. However, plans are in place for a future expansion of the facility to internalize these remaining processes as well.

In the case of dry-type transformer manufacturing, the company operates entirely independently. The factory is equipped with a gravity winding line for medium-voltage coils, as well as a casting machine for resin insulation, integrated with hardening ovens. This setup enables the complete in-house production of dry transformers, including both winding and encapsulation processes.

The company generously provided no-load loss test results from their in-house production of three-phase oil-immersed transformers. These measurements were obtained at the manufacturer's facility, which includes a dedicated transformer testing laboratory—shown in Fig. 6.4—equipped to perform standardized routine tests and selected type tests. The data reflect real, final-stage test results of fully assembled and operational units, thus offering a reliable basis for model development.

Fig. 6.3: Storage area with assembled transformer cores awaiting winding installation at Elpro-Energo Transformers s.r.o. [18].

Fig. 6.4: Completed transformer positioned in the testing laboratory at Elpro-Energo Transformers s.r.o. [18], where the no-load test is conducted.

The dataset was delivered in two distinct stages. The first batch was used to construct and validate the initial surrogate models, enabling the identification of relevant input features and assessing the overall feasibility of applying machine

learning to this problem. Subsequently, the dataset was expanded with additional measurements, allowing for the refinement and enhancement of the modeling approach to develop a more accurate and robust final model. The structure and content of the measured data, along with the key modeling milestones associated with each dataset delivery, are described in the following sections.

As illustrated in Fig. 6.5, the complete set of measured no-load losses exhibits significant dispersion compared with the values calculated using the conventional method described in Chapter 3 by (3.8) and the building factor from Table 3.1. The dataset was evaluated using (3.6) and (3.10) to calculate the nominal core losses based on transformer and core parameters, as well as to determine ANLL from measured data in a retrospective manner. Although the data follow a general trend, the measured values clearly depend on many more variables than only the rated power, which is used as the single input for analytical calculations. This single-variable approach is insufficient for achieving precise no-load loss estimation.

Consequently, a novel method based on machine learning is particularly suitable for this task, as it can account for a wide range of input variables with diverse characteristics and nonlinear interactions. By leveraging this multivariate information, the proposed approach can significantly improve the accuracy and reliability of no-load loss predictions, surpassing the limitations of traditional analytical methods.

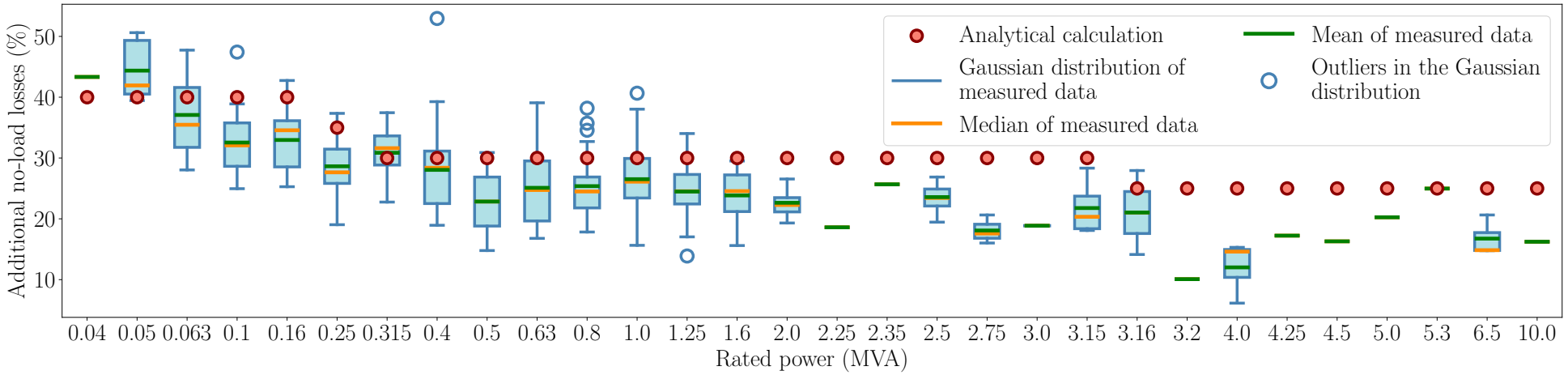


Fig. 6.5: Additional no-load losses as a function of transformer rated power: comparison of analytically estimated values (points) and measured values (box plots), adapted from [6].

6.2.1 Transformer No-Load Test

No-load loss measurement in power transformers is a routine standardized test procedure, performed as the final stage of transformer manufacturing. It is carried out in accordance with standards published by the International Electrotechnical Commission (IEC), specifically IEC 60076-1:2011, Clause 11.1: "General requirements for routine, type and special tests," Clause 11.5: "Measurement of no-load losses and current"[28], and IEC 60076-8:1997, Clause 10: "Guide to the measurement of losses in power transformers" [90].

A power transformer no-load test, also known as an open-circuit test, is a diagnostic procedure used to assess the transformer's performance when no load is connected to the secondary winding. The test involves applying rated voltage and frequency to one winding (typically the low-voltage side, to minimize the burden on the power source and simplify the testing procedure), while keeping the other winding(s) open-circuited. The primary objective is to measure the no-load losses (primarily iron losses in the transformer core) and the no-load current. These are critical for verifying compliance with EcoDesign requirements, evaluating the magnetic core behavior, and assessing overall transformer efficiency.

Additionally, the test serves as a diagnostic tool for identifying potential issues, such as improperly connected windings, incorrect magnetic circuit assembly, or lamination defects. During the test, the input active power, excitation current, and voltage on each phase are recorded, from which the no-load losses and current are calculated.

6.2.2 No-Load Loss Measurement and Voltage Distortion

If the voltage is applied to a winding with tap changers, the nominal tap should be used. Alternatively, the appropriate voltage corresponding to the selected tap should be applied if the test is performed at a different tap position. In the case of a prototype unit, as part of type testing, the no-load loss measurement should also be carried out at 90% and 110% of the rated voltage (or of the corresponding tap voltage). During the measurement, the transformer should be at approximately the ambient temperature of the testing laboratory.

The testing voltage must be measured separately on each phase, as illustrated in the circuit diagrams in Fig. 6.7. The test voltage is set using voltmeters that measure the mean value of the waveform but are scaled to display the corresponding root-mean-square (RMS) value of a sinusoidal waveform having the same mean value. This voltmeter reading is denoted as $\bar{U}$.

A second set of voltmeters must be connected in parallel with the first set. These voltmeters measure the true RMS values of the test voltage and are denoted

as U. The waveform of the test voltage is considered acceptable if the difference between the two readings, $|\bar{U} - U|$, does not exceed 3% of $\bar{U}$.

If this difference exceeds 3%, the validity of the test must be confirmed by agreement with the customer. A higher deviation may be tolerated for test voltages above the rated voltage, provided that the measurement is not part of a warranty evaluation.

6.2.3 No-Load Current Measurement

Due to the nonlinear magnetization characteristic of the transformer core, shown in Fig. 6.6(a), and the corresponding iron losses represented by the area enclosed in the magnetic hysteresis loop, see Fig. 6.6(b), the resulting no-load current becomes inherently distorted when a sinusoidal voltage is applied, as illustrated in Fig. 6.6(c).

For smaller power transformers, the no-load current typically ranges from 1% to 5% of the rated current, while for larger power transformers it is usually between 0.1% and 0.3%. The no-load current is defined as the RMS value of the current measured during the no-load loss test and is usually expressed as a percentage of the rated current of the winding to which the test voltage is applied.

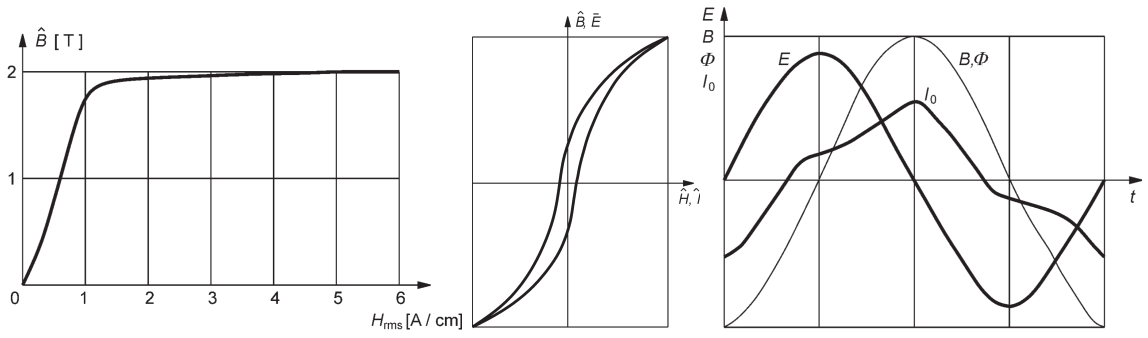
In the case of three-phase transformers, the no-load current is calculated as the average of the three phases. For typical core-type magnetic circuits, the current in the phase connected to the middle limb is lower than that in the outer phases, due to the approximately halved magnetic reluctance of the center limb's path. As shown in [91], the current ratio across the three phases can be approximated as $1 : 0.718 : 1$.

The total current drawn by the transformer during the no-load test comprises two components: the magnetizing current required to establish the magnetic flux in the core, and a capacitive current resulting from the inter-winding capacitance. At low excitation levels, especially in high-voltage transformers, the capacitive current may dominate. This causes the no-load current to initially decrease with increasing voltage until it reaches a minimum value. Beyond this point, as the voltage continues to rise, the magnetizing component becomes dominant, and the no-load current increases accordingly.

6.2.4 Measurement Circuit Configuration

The connection of the measuring circuit depends on the transformer vector groups, as described in [92]:

- The circuit in Fig. 6.7(a) is suitable for transformers with vector groups Yd, YNd, or Dd. The voltage transformers and wattmeters are connected in star.



(a) (b) (c)

Fig. 6.6: (a) Magnetization characteristic of the transformer core, where $\hat{B}$ denotes the peak magnetic flux density and H_{rms} represents the RMS value of the magnetic field intensity; (b) Hysteresis loop of the iron core under dynamic excitation from an alternating voltage source, where $\hat{E}$ is the peak applied voltage, $\hat{H}$ the peak magnetic field intensity, and $\hat{I}$ the peak magnetizing current; (c) Oscillographic representation showing the time evolution of the induced voltage E, magnetic flux density B, total magnetic flux ϕ, and no-load current I_0 during a no-load test. Adapted from [92].

The voltmeters are connected between two phases and measure the line-to-line voltages, which correspond to the phase voltages of a delta-connected transformer under test.

- Fig. 6.7(b) shows the connection for transformers with vector group Dyn. The voltage transformers and wattmeters are connected in star. The voltmeters measure the line-to-line voltages. This configuration is acceptable despite the star connection because the delta connection on the high-voltage side allows third-order harmonic currents to circulate, preventing voltage distortions on the low-voltage winding.
- In Fig. 6.7(c), for vector groups Yyn, YNyn, Yzn, or YNzn, the voltmeters must measure the phase voltages.
- For the circuit in Fig. 6.7(d), applicable to vector groups Yy, YNy, and ZNy, the voltage transformers are delta-connected on the secondary side. The voltmeters measure the line-to-line voltages.

6.2.5 Measuring Procedure

Prior to testing, built-in current transformers must be short-circuited, and condenser bushing taps earthed. Neutral points of windings without full insulation must also be grounded. The voltage ratio shall be verified before initiating the no-load loss test.

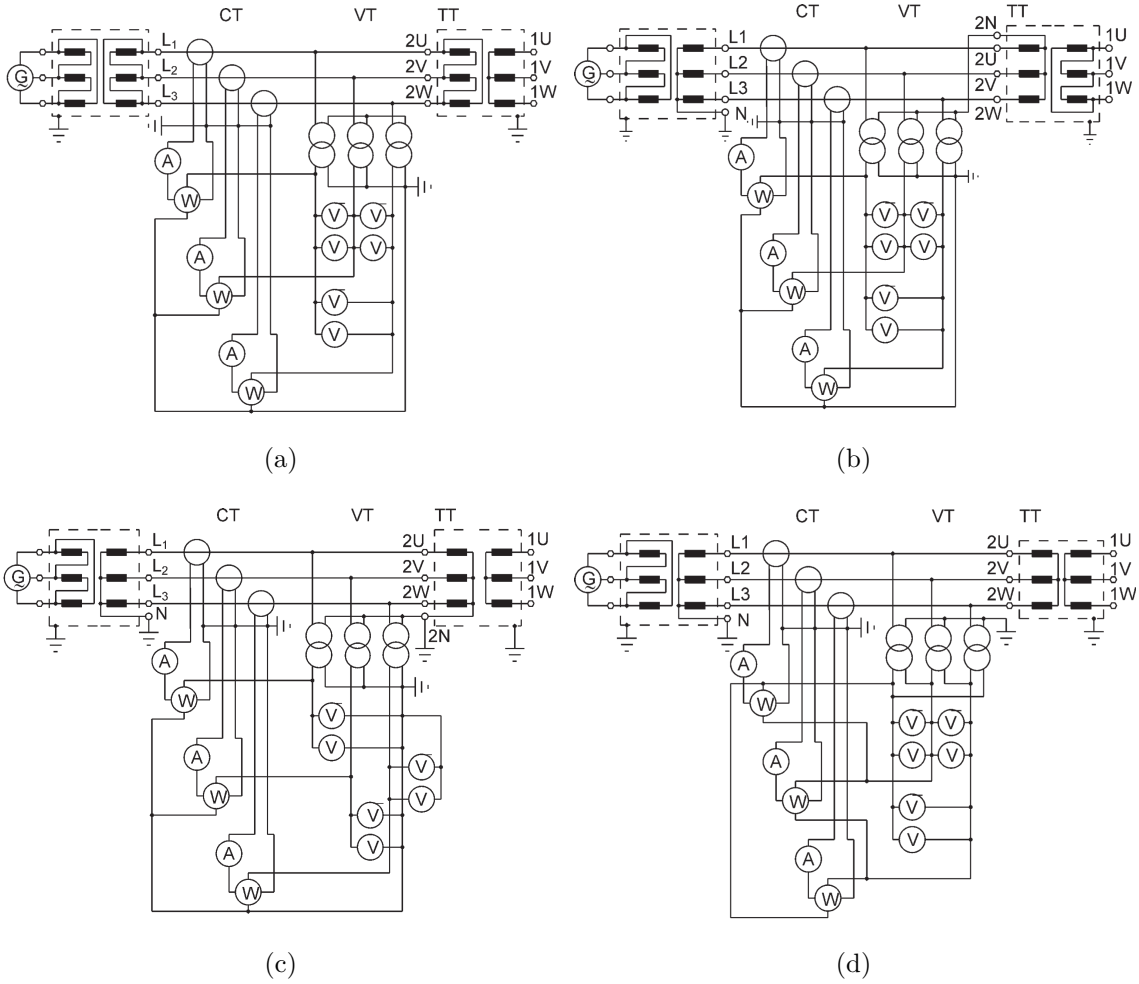


Fig. 6.7: Three-phase transformer wiring diagrams for the no-load test of a transformer with the following connection groups: (a) Yd, YNd, or Dd; (b) Dyn; (c) Yyn, YNyn, Yzn, or YNzn; and (d) Yy, YNy, or ZNy. The circuits also include current transformers (CT) and voltage transformers (VT) for connecting ammeters, voltmeters, and wattmeters required for the measurement, adapted from [92].

For oil-immersed transformers, bushing and Buchholz relays must be vented, and oil levels in both the transformer and on-load tap changer (if present) must be checked.

To mitigate remanent magnetization effects caused by previous direct current (DC) excitation (e.g., resistance measurement or switching transients), the transformer should be energized at 1.10–1.15 per unit of the rated voltage. Several voltage cycles are required to stabilize the magnetization curve. During this period, readings on ammeters and wattmeters typically decrease. Once steady-state conditions are achieved, measurements may commence.

Measurements are usually recorded at 110%, 100%, and 90% of the rated voltage. Voltage is regulated using average-reading voltmeters. For three-phase transformers,

the arithmetic mean of the three measured voltages is used. If the voltage cannot be set within ±0.1% of the target value, loss values are interpolated.

In large three-phase transformers, wattmeter readings may vary across phases; one reading may even become negative. Total input power is calculated as the sum of the three wattmeter values. This is attributed to magnetic asymmetry in the core, which induces unbalanced no-load currents. At certain flux densities, phase displacement between current and voltage may exceed 90°, resulting in negative power as discussed in [93].

6.2.6 Postprocessing of Measurement Results

Depending on the type of voltage distortion, the no-load losses under distorted excitation can be either lower or higher than that under sinusoidal excitation. If correction is required, the measured no-load losses P_{m} should be adjusted to sinusoidal waveform equivalent no-load losses P_0 as

$$P_0 = P_{\mathrm{m}}(1 + d), \tag{6.1}$$

where d is the relative difference between the measured voltages $\bar{U}$ and U, defined as

$$d = \frac{\bar{U} - U}{\bar{U}}, \tag{6.2}$$

with d typically being negative.

According to IEC standard [28], the measured no-load losses are not corrected to a reference temperature relative to the ambient temperature. If the measurement follows the IEEE standard [94], an empirical correction formula is applied when the average oil temperature deviates by more than ±10 °C.

6.3 Surrogate Model Construction

For the construction of all surrogate models presented in this thesis, the Python library `GPyTorch` was utilized, as it provides state-of-the-art algorithms and highly efficient implementations for Gaussian Process Regression (GPR) modeling [95]. The available dataset was always divided into training and testing sets with a 90:10 ratio throughout this work.

Given that the input space of the analyzed problem includes both continuous and categorical input variables, a single-task exact GP model designed for mixed search spaces from the `BoTorch` library—a high-level extension of the GPyTorch ecosystem—was employed. This model architecture is specifically capable of handling such hybrid input domains. It leverages a composite kernel, which combines a categorical

kernel (based on Hamming distances), defined in (4.6), with a continuous kernel. The resulting kernel formulation is expressed as:

$$\begin{aligned} K((x_1, c_1), (x_2, c_2)) = K_{\text{Cont1}}(x_1, x_2) + K_{\text{Cat1}}(c_1, c_2) \\ + K_{\text{Cont2}}(x_1, x_2) \cdot K_{\text{Cat2}}(c_1, c_2), \end{aligned} \tag{6.3}$$

where $K\text{Cont}$ and K_{Cat} denote the kernel functions for continuous (x) and categorical (c) input components, respectively.

All GPR models were trained using a Gaussian likelihood function as defined in (4.8), and the model hyperparameters were optimized using the gradient-based Limited-memory Broyden–Fletcher–Goldfarb–Shanno (*L-BFGS*) algorithm, as introduced in [96].

6.3.1 Model Accuracy Evaluation

To assess the performance and reliability of the surrogate models, three primary evaluation metrics have been employed:

1. The first metric involves a scatter plot comparing the predicted and measured values for both training and testing samples based on [97]. In the ideal case, all sample points align along the diagonal, indicating perfect model performance with zero error. To further quantify the fit, the coefficient of determination (R^2) was calculated, reflecting how well the predicted values approximate the true measurements [98]. The R^2 score ranges from 0 to 1, where a value closer to 1 indicates a higher proportion of the variance in the observed data is explained by the model. In other words, R^2 quantifies the precision of the model's predictions relative to the actual outcomes.
2. The second metric presents the distribution of the relative prediction error for the testing dataset known as the residual error plot in line with [99]. This visualization provides insight into the error spread and helps identify potential outliers or systematic biases in model predictions.
3. The third and final metric is a comparison based on a statistical analysis of the predicted testing data. This analysis includes the most commonly used and informative statistical coefficients, which are widely applied to assess the accuracy of surrogate models, as presented in [100], [101].

While not all of these metrics are strictly necessary in every modeling scenario, the first metric—the comparison plot of predicted versus actual values along with the R^2 score—is typically the most informative and interpretable for evaluating model accuracy. It offers a quick visual reference for identifying both overall performance trends and individual discrepancies.

Table 6.1: Comparison of the Coefficient of Determination (R^2) Scores for Different Kernels Used in the Final GPR Model with Data Filtration

Kernel type	$\boldsymbol{R^2}$**—Training data**	$\boldsymbol{R^2}$**—Testing data**
Matérn with $\nu = 3/2$	1.000000	0.739369
Matérn with $\nu = 5/2$	1.000000	0.788061
Radial basis function (RBF)	1.000000	0.406682
Rational quadratic (RQ)	1.000000	0.764397
First-Deg. Piecewise Polynomial	0.999531	0.735224
Second-Deg. Piecewise Polynomial	1.000000	0.694530

6.3.2 Continuous Kernel Selection

To properly select the most suitable kernel function for the continuous input variables in the model for transformer no-load losses estimation, a comparison of several commonly used kernel types was conducted. This comparison was initially performed using the first surrogate model set-up, introduced in Section 6.4, where only the continuous kernel from (6.3) was varied, while the categorical kernel remained unchanged. For validation purposes, the same comparison was later repeated for the final model with data filtration described in Section 6.5. As the results from both the initial and final models were highly similar, only the outcomes of the final GPR model for no-load loss estimation are presented and analyzed in detail below.

The kernel comparison included six distinct types of kernel functions, previously described in Section 4.2.2. The results of this evaluation are summarized in Fig. 6.8, which displays plots of predicted versus measured data for both training and testing datasets—the first metric for surrogate models from previous Section 6.3.1. All other modeling parameters and settings were held constant, ensuring that only the kernel function varied across the evaluations. The coefficient of determination (R^2) is also reported for each case in Table 6.1, providing a quantitative measure of model performance. The variation in sample distribution observed in each plot of Fig. 6.8 results from differences in the number of samples retained after filtering process those with high deviation, which varies for each kernel.

Based on both the visual assessment of prediction accuracy and the reported R^2 scores in Table 6.1, the Matérn kernel with smoothness parameter $\nu = 5/2$ shown in Fig. 6.8(b) consistently delivered performance with the highest accuracy. Although the rational quadratic (RQ) and first-degree piecewise polynomial kernels, depicted in Figs. 6.8(d) and (e), respectively, also yielded competitive results, the Matérn kernel remained the preferred option. This preference is due to its robustness, general

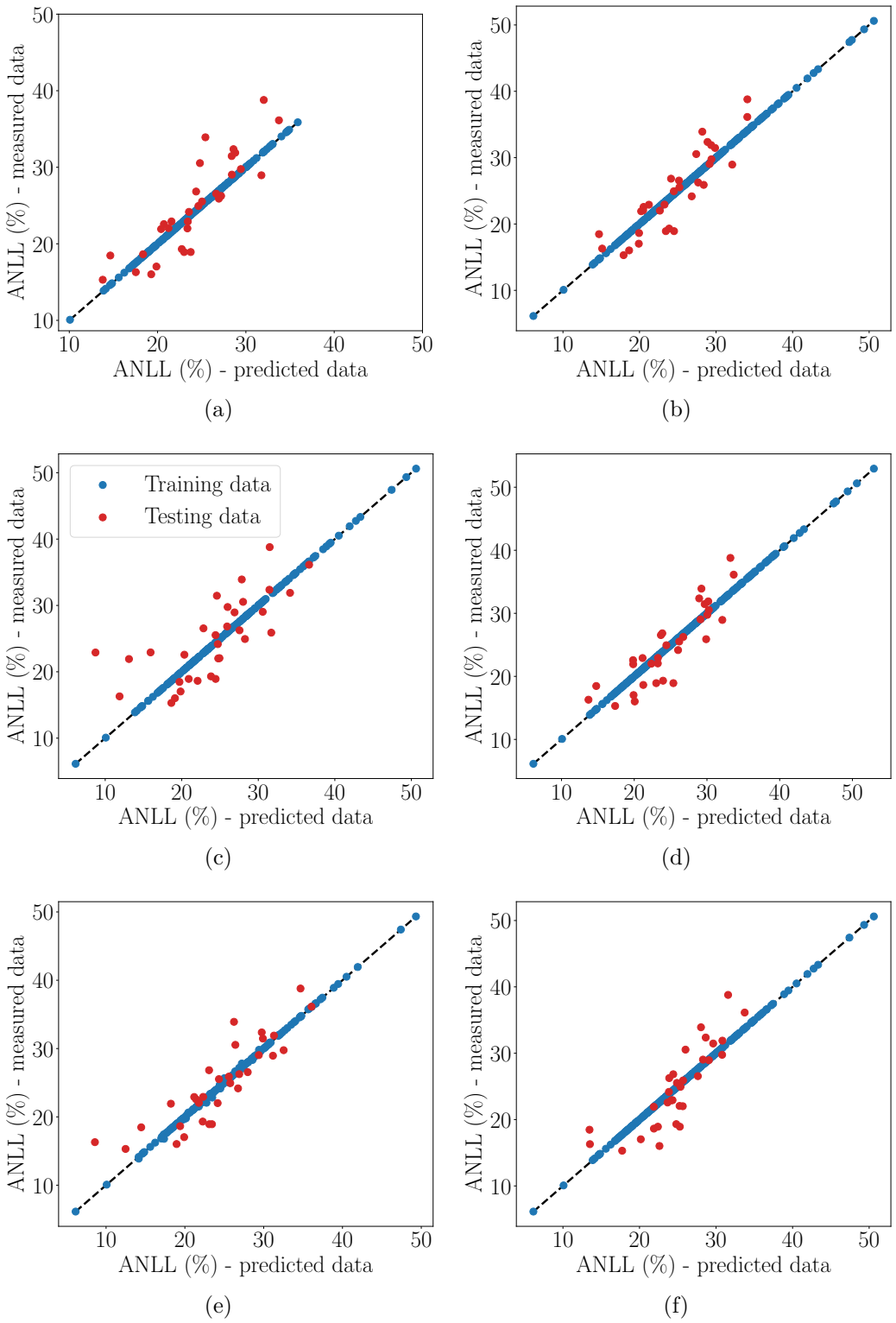


Fig. 6.8: Comparison of the final GPR model with different kernels: Matérn with (a) $\nu = 3/2$ and (b) $\nu = 5/2$, (c) radial basis function (RBF), (d) rational quadratic (RQ), and (e) first- and (f) second-degree piecewise polynomial.

applicability to real-world problems, and widespread use in the literature for modeling smooth yet flexible functions. It is important to note that the results for all examined kernels show almost perfect precision in modeling of learning data, which confirms the choice of GPR as a well-suited machine learning technique.

Simultaneously, a comparison of the performance of two different smoothness parameters was carried out, and the results are shown in Figs. 6.8(b) and 6.8(a). Both options demonstrated high accuracy, but the result for smoothness $\nu = 3/2$ in Table 6.1 shows worse R^2 than the result for smoothness $\nu = 5/2$.

The findings of this kernel comparison align with previous research on GPR modeling for problems with similar characteristics. For example, in [102], the Matérn kernel with a smoothness parameter of $\nu = 5/2$ was also shown to outperform the variant with $\nu = 3/2$. Similar conclusions were also drawn in [103], which compared several kernel functions and highlighted the advantages of the Matérn kernel in terms of prediction accuracy and model interpretability.

Based on these results, the value $\nu = 5/2$ was adopted for all model configurations employing the Matérn kernel in this work. This specific setting provides a suitable compromise between smoothness and modeling capacity, enabling the GPR model to accurately capture complex relationships in the data while avoiding overfitting. By substituting $\nu = 5/2$ into the general Matérn kernel formulation (4.5), the resulting kernel function is expressed as:

$$K(x_i, x_j) = \sigma_{\mathrm{f}}^2 \left(1 + \frac{\sqrt{5}}{l} d(x_i, x_j) + \frac{5}{3l^2} d(x_i, x_j)^2\right) \cdot \exp\left(-\frac{\sqrt{5}}{l} d(x_i, x_j)\right), \quad (6.4)$$

where $d(x_i, x_j)$ denotes the Euclidean distance between samples x_i and x_j, σ_{f} is the covariance scaling factor, and l is the lengthscale hyperparameter.

6.4 Initial Surrogate Models for No-Load Losses Estimation

In this section, the development of three initial surrogate models is described, each representing a successive step toward achieving higher prediction accuracy. These models were trained and validated using the results of no-load loss measurements from a total of 3,889 three-phase oil-immersed transformers. These transformers comprised 249 distinct designs, with each design serving as a single sample for model training. The production volume per design varied, ranging from a single manufactured unit to several hundred.

Overall, the provided dataset spans a broad spectrum of transformer configurations, exhibiting significant variability in core dimensions and electromagnetic properties. Table 6.2 summarizes the most relevant parameters of the cores—some

Table 6.2: Parameter Ranges of Transformer Designs and Variable Division by Its Application in the Surrogate Model Type for the Initial Models

Application	Variable	Unit	Range or options
Informative	Rated power	kVA	50—5,300
	Primary rated voltage	kV	6—35
	Secondary rated voltage	V	400—1,275
	Winding material	-	Al / Cu
Basic	Core leg pitch	mm	225—596
	Core window height	mm	209—1,205
	Core cross-section area	mm^2	12,460—63,360
	Nominal core loss	W	55.4—2,229.6
Additional	Steel grade (M0xx-23L)	-	70 / 73 / 75 / 80 / 85
	Maximum steel sheet width	mm	100—300
	Calculated flux density	T	0.8—1.74
Output	ANLL mean value	%	10.6—52.9

of which are discussed in Section 6.1—including their minimum and maximum values or the range of material options observed in the initial dataset batch.

For the purpose of creating and training the initial surrogate model for estimating transformer no-load losses, a carefully selected subset of core design parameters was chosen as modeling input variables based on their theoretical relevance and expected influence on core loss behavior. These parameters were systematically categorized into four distinct groups:

- *Informative*—variables that are not directly used for model training but are retained in the dataset to provide additional context. These variables are useful for post-analysis tasks such as identifying outliers, tracing errors, and conducting visual inspections;
- *Basic*—a set of four core parameters considered the most influential in modeling ANLL. These variables were used exclusively in the development of the very first surrogate model;
- *Additional*—three supplementary variables introduced in the subsequent two models to improve overall model robustness, predictive accuracy, and generalization capability;
- *Output*—the target variable is not directly represented by the total no-load losses. Instead, better results were obtained by predicting the percentage value of ANLL, which can be readily used for total no-load loss calculation, as defined in Section 3.2 using (3.10).

Another significant parameter that could potentially influence the model is the lamination stacking factor. However, since all available data correspond to the same lamination type and identical insulation layer thickness, the stacking factor remains constant across the entire dataset. As a result, this parameter does not vary and therefore cannot influence the model outcomes or affect the resulting predictions.

This structured approach allowed for iterative model refinement and better understanding of the impact of individual input variables on the surrogate model's performance.

6.4.1 Set-Up of Three Initial Surrogate Models

At this stage, three initial models have been introduced:

1. The very first set of input variables used for model creation involved four core parameters, theoretically among the most important in terms of their effect on no-load losses—a group named as *Basic* in Tab. 6.2: core leg pitch, core window height, core cross-section area, and nominal core loss. All these variables were modeled using a purely continuous Matérn kernel in this model.
2. Secondly, to the first model another group of core parameters was added—named as *Basic+Additional* in Tab. 6.2: steel grade, maximum steel sheet width, and flux density from calculation. This model is thus represented by seven input variables, all of which are processed using continuous Matérn kernel.
3. The third model uses exactly the same input variables as the second model. However, due to the categorical nature of the input variable *steel grade*, a separate surrogate model was developed that incorporated a mixed input variable space within the GPR framework. In this model, a categorical kernel was employed for the steel grade, while all other variables continued to use the continuous Matérn kernel, as described previously.

6.4.2 Results Evaluation with the First Metric

Utilizing the first verification metric described in Section 6.3.1, Fig. 6.9 presents the predicted versus measured plots of ANLL for all three developed surrogate models, along with the analytically calculated ANLL, allowing for a direct visual comparison of accuracy and generalization performance. The analytical reference is derived from the manufacturer's long-term empirical experience, where the value of ANLL is assigned solely based on the transformer's rated power. According to this rule, ANLL decrease gradually with increasing rated power, ranging from approximately 40% down to 25%, as illustrated in Fig. 6.9 (a).

Table 6.3: Comparison of the Coefficient of Determination (R^2) Scores for Analytical Calculation and Initial GPR Models in No-Load Loss Estimation

Model type	$\boldsymbol{R^2}$**—Training data**	$\boldsymbol{R^2}$**—Testing data**
Analytical calculation	−0.208114	−0.640080
GPR with 4 inputs—continuous	1.000000	0.240776
GPR with 7 inputs—continuous	1.000000	0.223799
GPR with 7 inputs—categorical	1.000000	0.636264

For the training data, all models demonstrate very low prediction errors, indicating proper model construction and successful hyperparameter optimization. However, noticeable performance differences arise when evaluating the testing dataset. The models trained with four input variables, as shown in Fig. 6.9 (b), and with seven purely continuous input variables, shown in Fig. 6.9 (b), yield similar training accuracy levels. Nevertheless, the latter model, which incorporates three additional continuous inputs, exhibits reduced generalization capability, evident from the broader dispersion of predictions on unseen testing samples. Both models display substantial scattering in test data predictions, indicating weaker extrapolation performance. These trends are corroborated by the coefficient of determination (R^2) values presented in Table 6.3.

In contrast, the third model, illustrated in Fig. 6.9 (d), applies a mixed-kernel approach incorporating a categorical kernel for the steel grade variable and achieves significantly better prediction performance on the testing dataset. The predicted samples for this model align more closely along the diagonal, indicating higher accuracy and reduced bias. Furthermore, its R^2 score is notably higher than those of the other two models, highlighting the benefits of appropriately handling categorical inputs within the GPR framework.

6.4.3 Results Evaluation with the Second Metric

The second evaluation metric analyzes the distribution of the relative prediction error for all three surrogate models when compared against the measured data. These results are presented in single Fig. 6.10, which visualizes the relative errors specifically for the testing dataset.

Upon closer examination, the first two models—those using purely continuous kernels with four and seven inputs, respectively—exhibit relatively similar error distributions. In particular, the model with seven inputs demonstrates a slightly narrower distribution in some regions, yielding marginally smaller relative errors

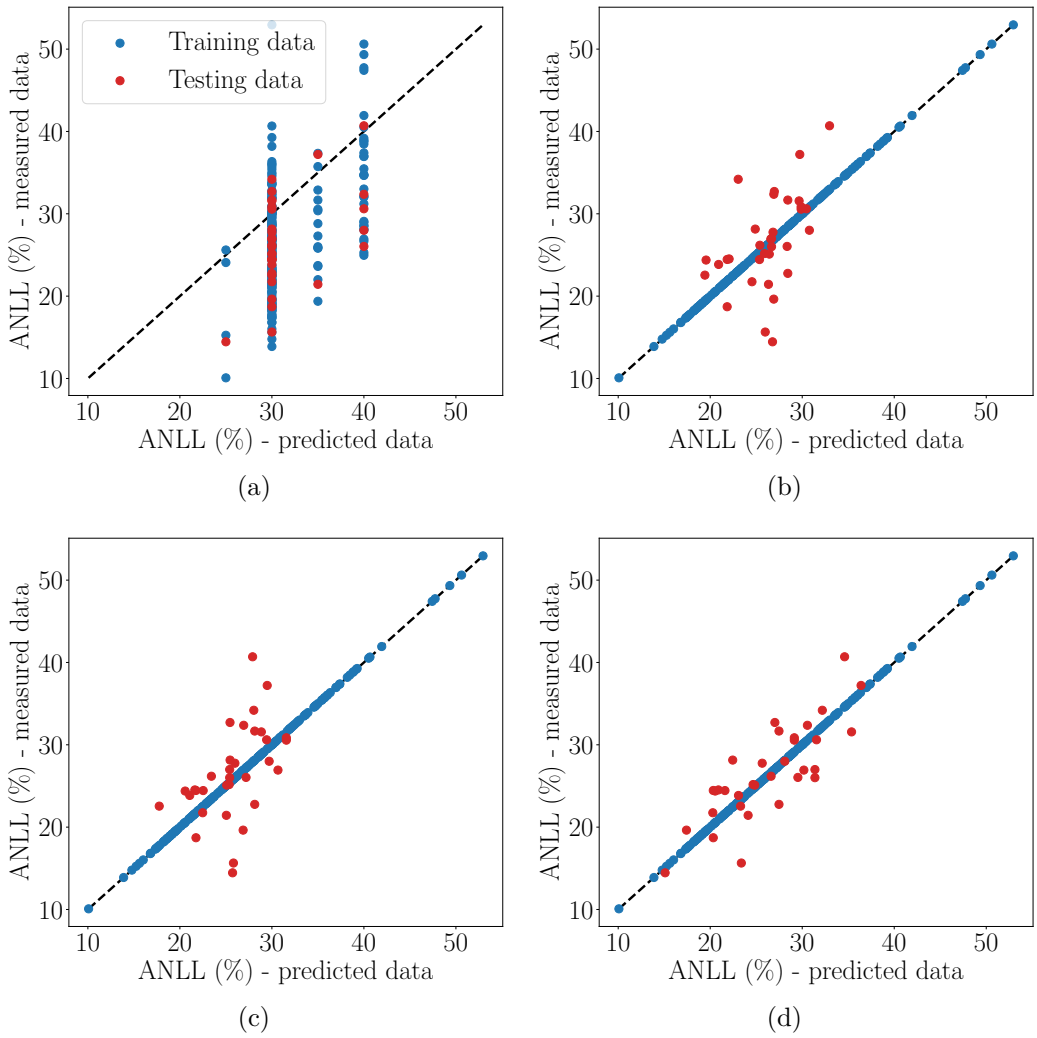


Fig. 6.9: Comparison of the evaluated analytical calculation (a) and GPR model accuracy for models with 4 (b) and 7 input variables using continuous kernels (c), and a mixed categorical kernel (d), adapted from [8].

by a few percentage points compared to the simpler four-input model. This indicates a modest improvement in predictive accuracy, although the difference is not substantial.

Most importantly, the third model, which incorporates a categorical kernel to appropriately handle the steel grade options, again shows clearly superior performance. Its relative prediction errors are consistently lower across a wide range of samples, with differences reaching up to tens of percentage points in certain cases. This reinforces the importance of using an appropriate kernel structure when modeling mixed input types in GPR.

It is also important to highlight that all three GPR-based models significantly outperform the analytical estimation method, particularly in terms of predictive

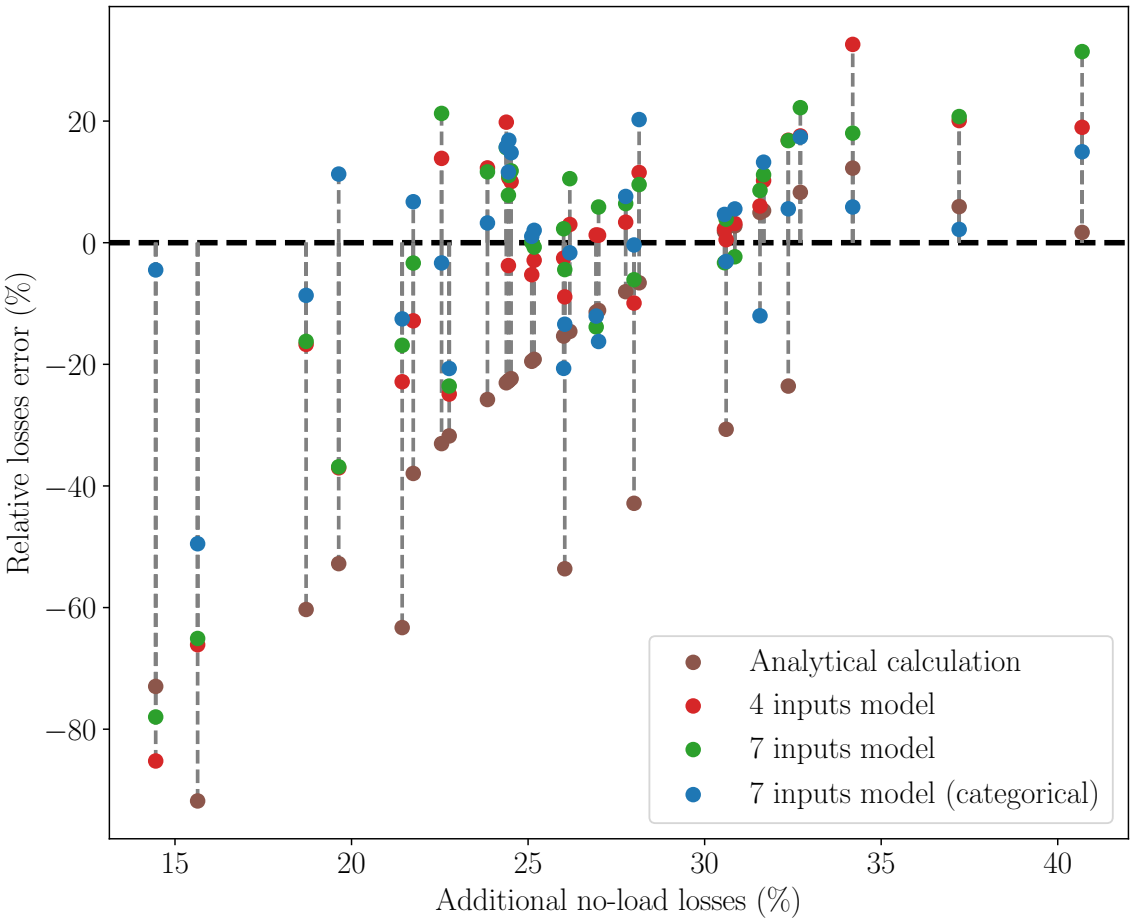


Fig. 6.10: Comparison of analytical calculation and GPR models with four and seven (continuous and categorical) input variales, using evaluated relative errors for testing data. The plot highlights differences in prediction errors across the range of no-load losses, adapted from [8].

accuracy across the testing data. This further confirms the advantage of data-driven surrogate modeling techniques for transformer no-load loss prediction over traditional rule-based approximations.

6.4.4 Results Evaluation with the Third Metric

The final stage of model verification involved a comprehensive comparison of the three surrogate models based on a statistical analysis of the prediction errors on the testing dataset. The results of this evaluation are summarized in Table 6.4, which reports multiple key error metrics, including the maximum, minimum, mean, and median of the relative errors, as well as higher-order statistics such as variance, standard deviation, and mean squared error.

As observed in the table, the first two models—those using four and seven contin-

Table 6.4: Comparison of Key Statistical Error Metrics on Testing Data for Analytical (Analyt.) Calculation and GPR Models Using Four and Seven Input Variables. The Last Two Models Were Evaluated Chronologically with Continuous (Cont.) and Mixed Continuous and Categorical (Cat.) Kernels.

Error type	**Unit**	**Anal.**	**4 inputs**	**7 inputs**	
				Cont.	**Cat.**
Maximal relative error	%	12.3	32.6	31.4	20.3
Minimal relative error	%	−91.8	−82.2	−78.0	−49.5
Mean relative error	%	23.5	−2.5	−0.72	0.06
Median relative error	%	−22.35	2.2	5.87	2.2
Variance of relative error	$\%^2$	617.6	554.5	521.5	202.4
Standard deviation (STD)	%	24.9	23.5	22.8	14.2
Mean squared error (MSE)	$\%^2$	51.5	23.8	24.4	11.4
Root mean squared error (RMSE)	%	7.2	4.88	4.94	3.38
Mean absolute error (MAE)	%	6.0	3.7	3.8	2.8

uous input variables—again produce relatively similar performance. The first model yields slightly better results in terms of central tendency, with lower mean and median relative errors. Conversely, the second model with seven inputs demonstrates reduced extremes (i.e., smaller maximum and minimum relative errors), indicating a lower variance and a slightly more stable error distribution.

The third model, which incorporates a categorical kernel to handle the steel grade variable, again outperforms both continuous models across nearly all statistical metrics. It delivers the lowest variance, standard deviation, and overall error magnitudes, confirming its superior generalization capability. This model achieves notably improved accuracy without compromising stability, making it the most robust choice for predicting transformer no-load losses.

Furthermore, all three GPR models substantially outperform the analytical calculation approach, which shows considerably higher errors across all listed metrics. These results reinforce the effectiveness of GPR models—particularly those incorporating mixed input types—for surrogate modeling in transformer design optimization tasks.

6.4.5 Final Discussion for Initial Models

The results have shown that the machine learning technique GPR is a highly effective method for predicting transformer no-load losses when a sufficient number of learning samples is available. Three surrogate models are presented, all with one output:

the percentage value of ANLL relative to nominal no-load losses. All introduced models indicate a significant improvement in the accuracy of loss predictions compared to analytical calculations. The first model incorporates four input variables, while the second model expands to include three additional inputs. Both models utilize variables with purely continuous kernels. Because the second model with seven inputs predicts new data slightly less accurately, a third similar model is introduced. In this third model, one variable, steel grade, utilizes a categorical kernel due to its specific behavior. This minor adjustment significantly improves the accuracy of the prediction compared to previous models. There is a valuable lesson to be learned; consistent observation of variable characteristics is crucial to achieve optimal results.

The most significant advantage of this method is that it considers all the factors that contribute to core losses, even those that are impossible to simulate by other means. Another advantage is the high speed with which results can be obtained, which makes it easily applicable to analytic calculations during the design process with high accuracy. The disadvantage is its applicability only after a certain amount of measured data of cores built by the same technology is obtained. Furthermore, this method should not be used to predict losses in cores with parameter values far from the known intervals.

6.5 Improved Surrogate Model for Enhanced No-Load Loss Estimation

In this section, the previously introduced surrogate model is further refined through improved preprocessing of the measured data, specifically by filtering out samples with unusually high deviations. Model accuracy is additionally enhanced by an optimized selection of input variables. Furthermore, the model is extended to include multiple outputs for the estimation of extreme values observed during the measurement of multiple transformer units sharing the same design. This extension enables the prediction of both the maximum and minimum deviation values, which is of critical importance during the design stage, as it allows for better risk assessment and margin setting. In addition, a learning curve is constructed for the investigated GPR model, enabling analysis of the impact of dataset size on prediction performance. The improved surrogate model for enhanced no-load loss estimation is published in [6].

For the purpose of model improvement, the dataset was expanded to include no-load loss measurements of 4,503 oil-filled three-phase transformers manufactured by Elpro-Energo Transformers s.r.o. [18]. These transformers correspond to 322 distinct designs, each of which is again treated as a single sample for model training.

Table 6.5: Parameter ranges of transformer designs and variable division by its application in the surrogate model type for the improved model. Parameters serving as the optimized set of input variables yielding the highest accuracy are highlighted in red.

Application	Variable	Unit	Range or options
Informative	Rated power	kVA	40—10,000
	Winding material	-	Al / Cu
Original Model Inputs	Core leg pitch	mm	209—739
	Core window height	mm	295—1412
	Core cross-section area	cm^2	100.9—4,455.9
	Nominal no-load losses	W	55.4—4,752.7
	Steel grade (M0xx-23L)	-	70 / 73 / 75 / 80 / 85
	Maximum steel sheet width	mm	100—370
	Calculated flux density	T	0.8—1.74
Extended Inputs	Primary rated voltage	kV	5—35
	Secondary rated voltage	V	280—10,500
	Core shape	-	Oval / Round
	Core steps	-	3—10
Unused	Core real thickness	mm	125.6—370.5
Outputs	Mean ANLL	%	6.1—52.9
	Maximum ANLL	%	15.6—83.3
	Minimum ANLL	%	9.6—38.9

Table 6.5 summarizes the most relevant parameters of the transformer cores—some of which are discussed in Section 6.1—and introduces additional parameters compared to Table 6.2, reflecting further optimization of the input variable set. The table also lists observed value ranges of these parameters, as well as the available material options within the extended dataset.

For training purposes, these newly considered core parameters were again selected based on their theoretical importance in influencing no-load losses. All parameters listed in the table can be classified into five distinct groups according to their role in the modeling process:

- *Informative*—variables not included in the training model, serving exclusively for design identification and for providing complementary context during dataset analysis;
- *Original Model Inputs*—the seven key variables used in the previous model for ANLL estimation, as presented in Section 6.4, forming the baseline input

set for further optimization;

- *Extended Inputs*—variables newly added to the model, carefully selected through systematic input variable optimization in this thesis to maximize the precision and robustness of the estimation results;
- *Unused*—an example of a candidate input variable considered during the input set expansion process, showing potential relevance but ultimately excluded in the final combination due to its limited contribution to prediction accuracy;
- *Outputs*—the final model output variables, representing ANLL with their mean, minimum, and maximum values, enabling not only central tendency estimation but also prediction of extreme deviations important during the design stage.

As in Section 6.4, the lamination stacking factor, determined primarily by lamination thickness, is already accounted for through the steel grade specification. Nevertheless, the stacking factor may still exhibit minor variations due to manufacturing tolerances in the thickness of the insulation layer applied to the laminations by the steel supplier. Such subtle deviations constitute an additional source of variation in the measured no-load losses, the estimation of which forms part of the analysis presented in this section.

6.5.1 Data Filtration Techniques for Improved Training Results

One significant challenge associated with the measured dataset is the presence of deviations in the loss values, which can introduce substantial errors during model training. For example, two transformer designs with nearly identical configurations may exhibit markedly different measured losses, potentially causing the GPR model to overfit to these inconsistencies. This issue becomes particularly pronounced in cases where only a single unit of a given design was tested, thereby capturing all external influences—such as measurement noise, material variability, manufacturing tolerances, or human factors—within a single data point. Such deviations limit the generalization capability of the model and highlight the critical importance of preprocessing and validation strategies to mitigate their impact.

To address these limitations and to prevent the inclusion of anomalous data points in the predictive modeling process, a dedicated data-filtration procedure was implemented as part of this research. In the first phase of applying this technique, only the mean value of the no-load losses was considered.

For both data filtering and model performance evaluation, the K-fold cross-validation technique was applied [104]. In each filtration iteration, the dataset is divided into K equal parts (with $K = 15$ in this thesis). During every iteration, one

K-fold is used as the testing dataset while the remaining K-folds serve as the training dataset for the GPR model. For each validation dataset, the mean squared error (MSE) and the standardized residual (SR) are calculated and recorded for subsequent analysis.

After completing validation across all K-folds, the average MSE of the testing datasets is computed and stored. During the initial filtration pass, the 1% of samples exhibiting the highest SR values are removed from the dataset. In subsequent iterations, the newly obtained average testing dataset MSE is compared to the value from the preceding pass. If the relative improvement exceeds a predefined threshold (1% in this thesis) and the dataset still contains more than the specified minimum number of samples, the procedure again removes the top 1% of samples with the largest SR values. The filtering loop continues until one of the following stopping criteria is met:

- The improvement in MSE falls below the predefined threshold.
- The size of the filtered dataset reaches the lower limit.
- The maximum number of filtration iterations is reached.

It is worth noting that increasing the percentage of removed samples per iteration can accelerate the filtering process; however, this carries the risk of discarding relevant and potentially informative data, which may adversely affect model generalization. Therefore, careful consideration is required when setting this parameter to balance filtering efficiency with the preservation of high-quality training data.

This iterative filtration process ensures systematic refinement of the dataset by progressively removing high-deviation samples, thereby enhancing overall model performance. Simultaneously, it safeguards against excessive data reduction, maintaining a sufficient volume of reliable training data for robust surrogate modeling. The procedure is designed to achieve an effective balance between improving predictive accuracy and avoiding over-filtering, which could compromise the model's generalization capability. The entire workflow is summarized in the block diagram shown in Fig. 6.11.

6.5.2 Maximizing Model Performance with Optimization of Input Variable Combination

A wide range of parameters—particularly geometric ones—distinguishes one transformer core design from another. For effective machine learning-based prediction of no-load losses, it is essential to include only those input variables that significantly influence prediction accuracy. Initially, these parameters were selected manually based on engineering judgment and theoretical understanding of their impact on core

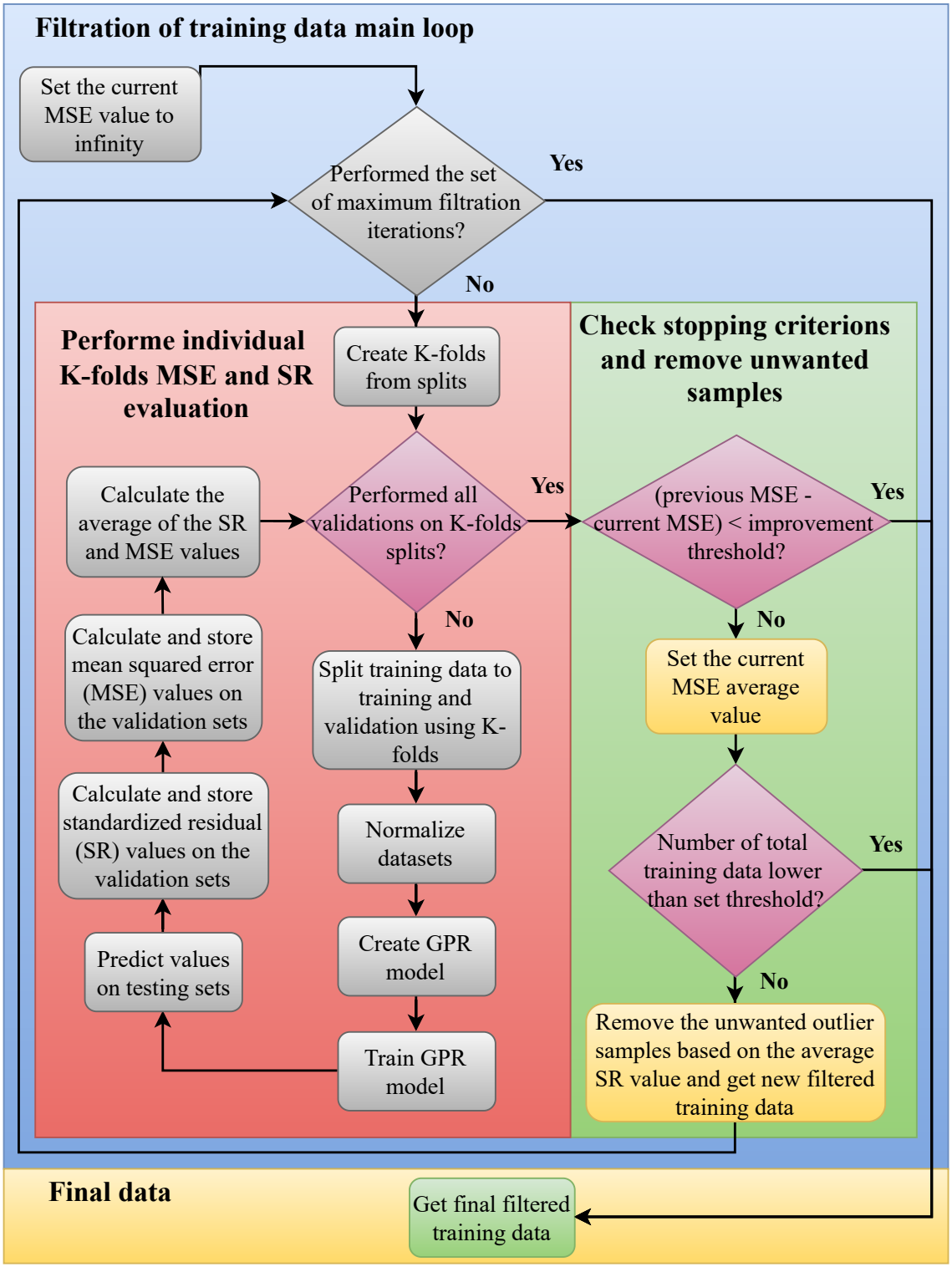


Fig. 6.11: Block diagram of the training data filtration algorithm, adapted from [6].

losses, as described in Section 6.4. To further enhance model performance and reduce reliance on subjective selection, an input variable optimization procedure was subsequently introduced. This automated approach employs an optimization

algorithm to identify the most influential features contributing to improved predictive accuracy. To facilitate this process, the set of potentially relevant input variables affecting no-load losses was expanded to include additional variables, as detailed in Table 6.5.

The primary tool used for this optimization task was the Python library `Mango`, developed by ARM Software. Mango is specifically designed for hyperparameter tuning in machine learning workflows and relies on Bayesian optimization [105] at its core. This probabilistic method efficiently explores the search space by modeling the objective function, requiring fewer evaluations to identify near-optimal solutions. Additionally, Mango supports parallel execution and accommodates search spaces with continuous, discrete, and categorical variables, making it particularly well-suited for the heterogeneous input types considered in this thesis.

The objective of the optimization was to determine the combination of input variables that minimizes the MSE of the resulting GPR model. This evaluation was based on the average MSE computed over 10 independent runs, each using a randomized split of 90% training and 10% testing data, drawn from the filtered dataset presented in Section 6.5.1. For the MSE evaluation, only a single output—the mean value of the no-load losses—was considered, excluding the minimum and maximum values. These two additional outputs were used later in this work. While a greater number of iterations would improve statistical robustness, a compromise was made to accommodate computational limitations. A graphical overview of the entire optimization process is presented in Fig. 6.12.

6.5.3 Prediction of Maximum and Minimum Loss Outcomes

A significant challenge in machine learning based on empirical datasets is the inherent variability in measured no-load losses among transformers that share the same design and have identical input variables. This variability is primarily attributed to uncontrollable deviation factors during manufacturing or testing. From the manufacturer's perspective, the worst-case scenario—i.e., the maximum observed loss—is of particular importance, as industry standards typically impose strict upper limits on no-load losses, with no tolerance for exceeding the specified threshold. Consequently, the ability to accurately predict the maximum loss value is essential for ensuring compliance.

In addition, estimating the minimum loss value can also be advantageous. It allows for determining the expected range of losses under ideal or optimal testing conditions, thereby offering insights into measurement consistency and potential manufacturing improvements.

The construction of the surrogate model for predicting loss extremes follows the same methodology used for mean no-load loss prediction, as described in detail

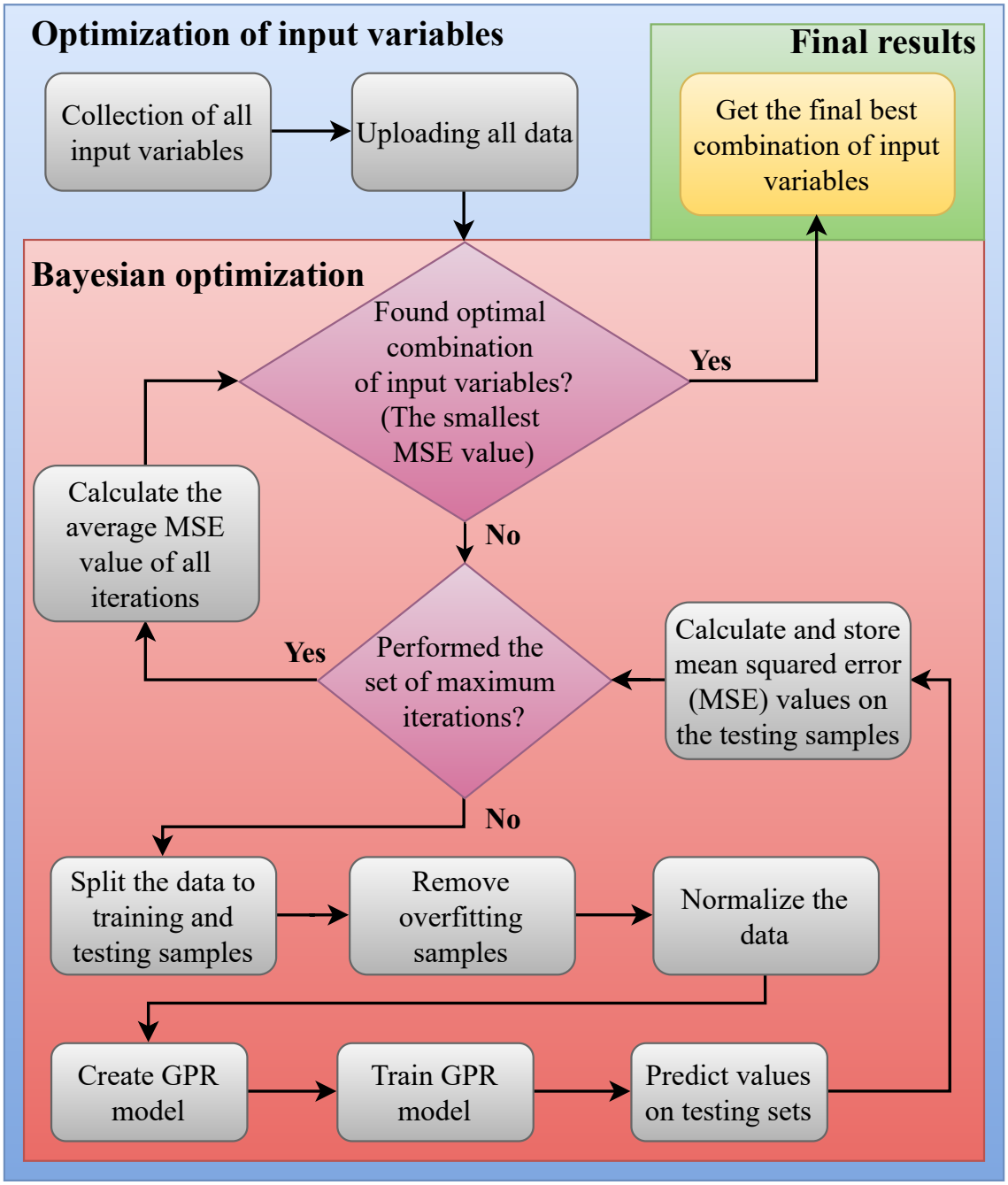


Fig. 6.12: Block diagram of the optimization algorithm for selecting the best-performing input variable combination, adapted from [6].

throughout this chapter. However, the model is extended to support multiple output targets. Specifically, it is configured to predict both the maximum and minimum losses associated with a given transformer design. Optionally, a third output—the mean loss—can be included, enabling the model to simultaneously learn and infer all three outputs within a unified framework.

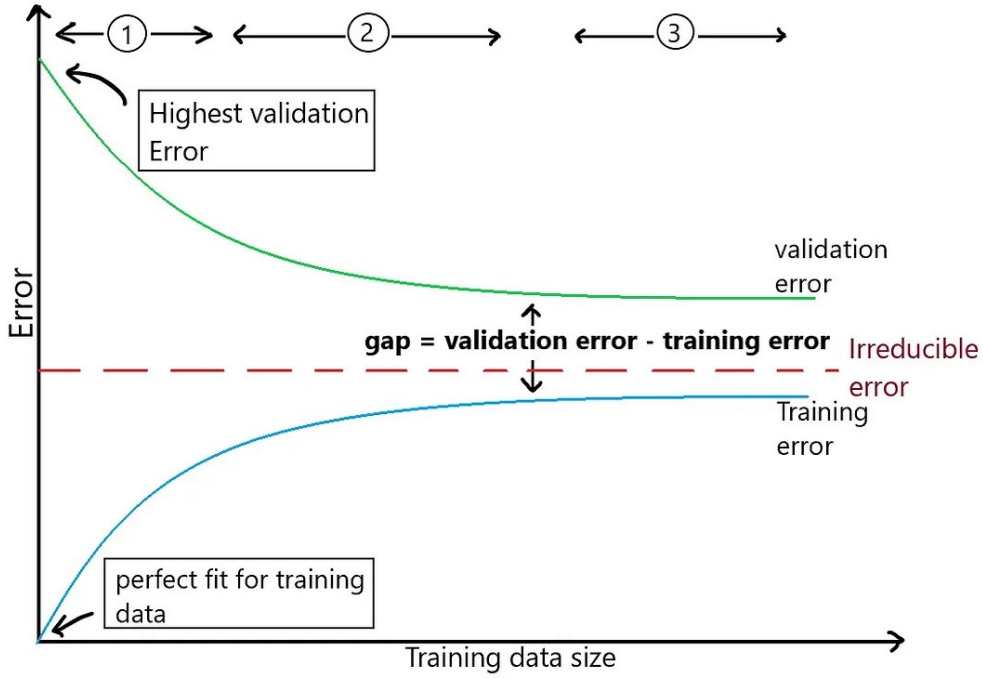


Fig. 6.13: Illustration of a typical shape of learning curves for training and testing data, adapted from [107].

6.5.4 Learning Curve

To evaluate the performance of the GPR model, a learning curve was constructed following the methodology outlined in [106]. A learning curve visualizes how model accuracy—here expressed by the MSE—varies with the size of the training dataset. Examining this curve makes it possible to determine the influence of training dataset size on predictive capacity and to decide whether further data collection or model refinement is warranted. The learning curve is plotted in two parts:

- *Training curve*—quantifies how closely the model fits the training data as additional samples are introduced.
- *Validation curve*—measures how well the model generalizes to unseen (testing) data for each corresponding training dataset size.

Together, these curves provide insight into potential underfitting or overfitting and help guide decisions on further data acquisition or model adjustments. These scenarios can be interpreted based on the shape of the curves, as illustrated in Fig. 6.13. The marked sections in the figure correspond to the following stages, adapted from [107]:

1. When the model is trained on a very small dataset (e.g., a single data point), it fits the training data perfectly, resulting in zero training error. However, when

applied to unseen validation data, the model performs poorly, yielding a high validation error.

2. As the number of training instances increases, the model learns to generalize better by minimizing error across all data points. It no longer fits every point exactly but instead captures underlying patterns while reducing sensitivity to noise. Consequently, the training error increases slightly, while the validation error decreases.
3. The optimal training dataset size is typically reached when both curves level off (plateau). Beyond this point, adding more data does not significantly improve model performance and only increases the computational cost of training.

6.5.5 Results of the New Optimized Model with Filtered Dataset

This section discusses the evaluation of modifications to the final GPR model previously introduced in Section 6.4. Since the dataset for model training has been extended, the evaluation is also repeated for the analytical calculation of no-load losses and for the previous GPR model with seven mixed input variables. Furthermore, a comparison is presented for the same GPR model incorporating data filtration—as described in Section 6.5.1—to remove samples with high deviation. Finally, the results of a new model with an optimized set of input variables, achieving the highest estimation accuracy, are presented. All evaluations presented in this section were performed using the mean value of the no-load losses.

The optimization was performed with the incorporation of the dataset filtration procedure described in Section 6.5.1, running on 30 parallel branches, each executed on its own Intel(R) Xeon(R) E5-2650 v2 processor core operating at 2.60 GHz, over a continuous period of two weeks. In total, 210 unique combinations of input variables were evaluated, some of them using both continuous and categorical kernel options. The objective was to identify the combination of variables that minimized the average MSE across multiple randomized training/testing splits. The optimal configuration, achieving the lowest average MSE, is highlighted in red in Tab. 6.5. Within this set, steel grade and core shape consistently exhibited the highest predictive relevance when modeled with a categorical kernel, whereas the remaining variables were more effectively represented with continuous kernels. This outcome confirms the importance of properly distinguishing between categorical and continuous data types in kernel design to maximize the performance of GPR models.

The accuracy of the presented models is evaluated using two metrics, as described in Section 6.3.1. The first metric involves plotting predicted vs. measured data for both training and testing samples, illustrated in Fig. 6.14. Figs. 6.14(a) and 6.9(b)

Table 6.6: Comparison of the Coefficient of Determination (R^2) Scores for Analytical Calculation and Initial GPR Models in No-Load Loss Estimation

Model type	$\boldsymbol{R^2}$**—Training data**	$\boldsymbol{R^2}$**—Testing data**
Analytical calculation	−0.166805	−2.415569
GPR—7 inputs—unfiltr. (worst)	1.000000	−1.548071
GPR—7 inputs—filtr. (worst)	1.000000	0.275018
GPR—opti. inputs—filtr. (worst)	1.000000	0.358516
GPR—opti. inputs—filtr. (best)	1.000000	0.841291

show the performance of analytically calculated losses and the final model with seven input variables, originally introduced in Section 6.4. Fig. 6.14(c) shows the results after applying data filtration to the seven-input model, while Fig. 6.14(d) presents the newest model with an optimized set of input variables, also incorporating data filtration. All GPR models shown above correspond to the worst-case scenario out of 100 random splits of the available dataset into training and testing datasets. For completeness and further verification, the best-case scenario of the optimized model with data filtration is displayed in Fig. 6.14(e). The coefficient of determination (R^2) for all models is summarized in Table 6.6.

All GPR models demonstrate near-zero error for the training data ($R^2 \approx 1$), confirming effective learning. For testing data, accuracy improves significantly with data filtration and further improves with input optimization, even in the worst-case scenario, as reflected in both the R^2 values and the sample distributions. Comparing Figs. 6.14(b) and 6.14(c) highlights the impact of data filtration on model accuracy, while the comparison between Figs. 6.14(d) and 6.14(e) emphasizes the variability introduced by different dataset splits, which explains why the worst-case scenario is usually presented. These observations are fully consistent with the reported R^2 values and underscore the importance of data quality and split selection in model evaluation.

The second metric for model verification, described as the third metric in Section 6.3.1, evaluates predicted testing data using standard accuracy coefficients. As shown in Table 6.7, performance improvements are primarily due to data filtration, with models incorporating this process yielding substantially lower errors. It is important to note that all GPR models consistently outperform analytical calculations, even in their respective worst-case scenarios. Furthermore, the results demonstrate the sensitivity of the model's performance to the choice of dataset split: error values for the best-case scenarios are typically approximately half those of the worst-case, highlighting the need to report both extremes for a comprehen-

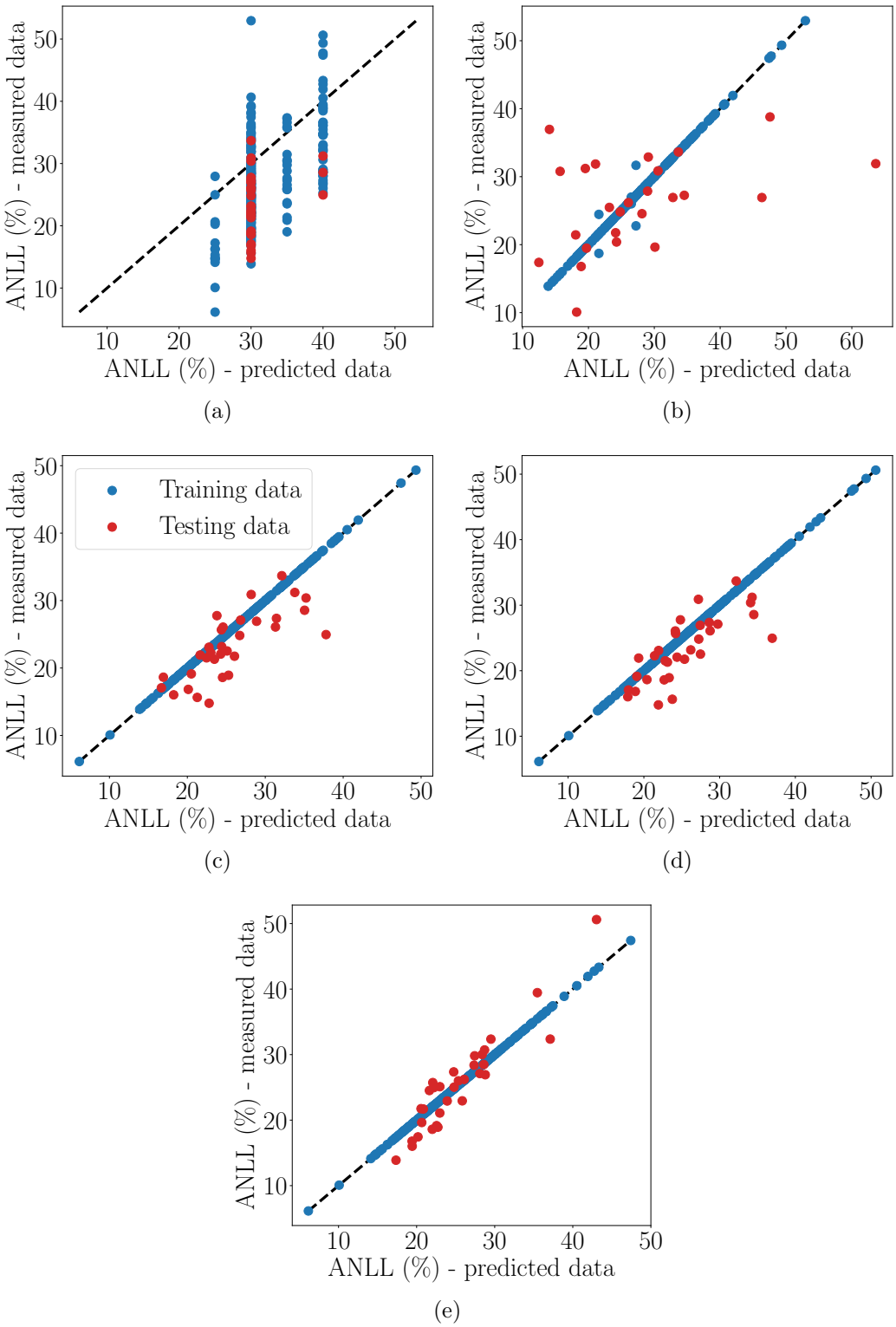


Fig. 6.14: Comparison of (a) analytical calculation and (b)–(d) GPR models with seven inputs under worst-case from 100 random splits: (b) no filtration, (c) with filtration, (d) new filtered; (e) best-case of new filtered model, adapted from [6].

Table 6.7: Comparison of Key Error Metrics for Analytical Calculation and GPR Models: Initial Seven-Input Mixed-Variable Model with Unfiltered (Unfil.) and Filtered (Fil.) Data, Optimized-Input Model with Filtered Data, Each in Worst Case, and Best Case for the Optimized Model.

Error type	**Unit**	**Analytical**	**Initial**		**Optimized**	
			Unfil.	**Fil.**	**Worst**	**Best**
Maximal relative error	%	10.9	61.9	14.3	12.0	15.0
Minimal relative error	%	−103.0	−99.4	−54.2	−51.9	−25.0
Mean relative error	%	−37.9	−7.6	−11.1	−10.2	−2.0
Median relative error	%	−35.3	−1.1	−8.0	9.8	0.7
Variance of relative error	$\%^2$	805.5	1,423.3	266.2	253.5	133.9
STD	%	28.4	37.7	16.3	15.9	11.6
MSE	$\%^2$	80.7	112.9	17.1	15.2	8.0
RMSE	%	9.0	10.6	4.1	3.9	2.8
MAE	%	7.9	7.2	3.2	3.1	2.4

sive assessment. These findings confirm that the optimized GPR model, combined with dataset filtration, provides the most reliable and accurate predictions of no-load transformer losses currently achievable.

The significant improvement of the GPR model over analytically estimated values, as well as its enhanced agreement with measured data, is clearly illustrated in Fig. 6.15. Model predictions are displayed using box plots for individual rated powers, which often encompass multiple transformer designs within the same category. Each box plot additionally marks the mean and median values for the corresponding rated power. Across the entire range of rated powers, the GPR model consistently exhibits high precision, effectively capturing both the central tendencies and the distributional spreads observed in the measurements. These findings provide strong evidence of the model's robustness and its capacity to generalize reliably across diverse transformer designs, while also highlighting the substantial performance gain compared to the analytical calculation, which is likewise presented in this figure.

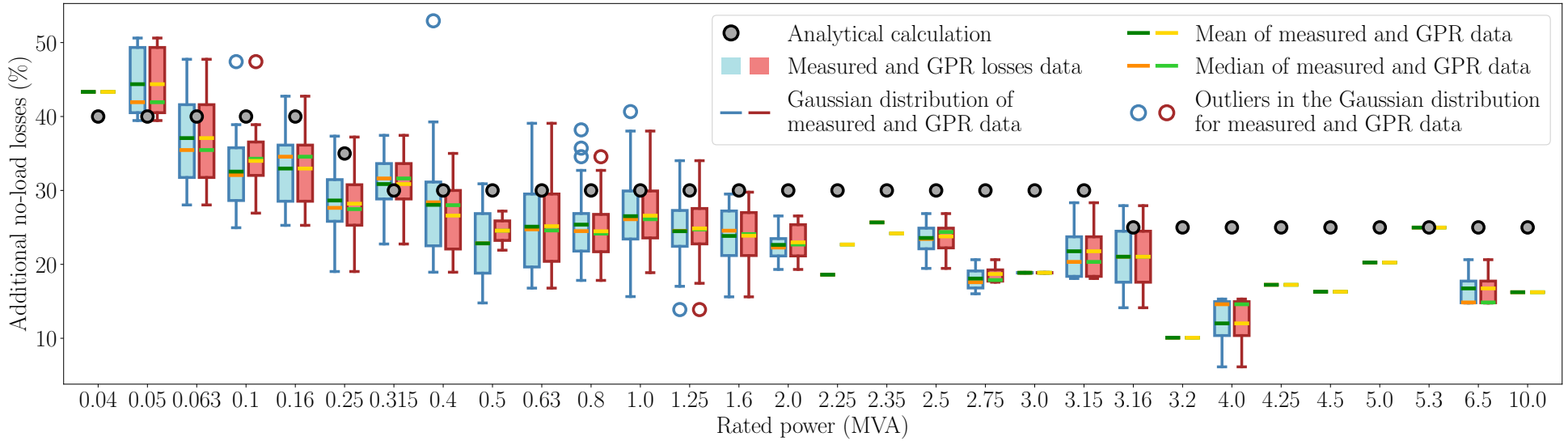


Fig. 6.15: Additional no-load losses depending on transformer rated power—comparison of analytically estimated (points), measured, and GPR model predicted values, expressed as a Gaussian distribution with box plots, adapted from [6].

Table 6.8: Comparison of Selected Key Metrics for Error Evaluation of Testing Data for Tolerances Modeling of Minimum, Maximum and Mean Values of ANNL

Error type	**Unit**	**Minimum**	**Maximum**	**Mean**
Maximal relative error	%	30.4	61.3	49.1
Minimal relative error	%	−80.9	−38.1	−22.7
Mean relative error	%	−2.9	6.7	11.5
Median relative error	%	3.2	4.8	13.4
Variance of relative error	$\%^2$	724.3	477.2	244.6
STD	%	26.9	21.8	15.6
MSE	$\%^2$	22.8	189.3	58.1
RMSE	%	4.8	13.8	7.6
MAE	%	3.8	7.1	4.9

6.5.6 Results of Maximum and Minimum Loss Modeling

The primary difference in the approach described in Section 6.5.3 lies in the requirement to filter the dataset to include only transformer designs with sufficient measured units to reliably capture extreme values. Specifically, only designs with at least two measured units in the dataset are considered, which reduces the final dataset to 184 distinct designs. Furthermore, because extremes are likely to be more pronounced as the number of measured units increases, the threshold is raised in this analysis to include at least three measured units per design. This ensures that the predicted maximum and minimum losses are both representative and practically applicable to real-world transformer evaluation.

The results of the evaluation of modeled minimum, maximum, and mean losses using the first metric described in Section 6.3.1 are presented in Fig. 6.16, along with the corresponding R^2 scores, while the third metric is summarized in Table 6.8. All results are shown for the worst-case scenario among 100 random splits of the dataset. As expected, the final prediction errors are noticeably higher than in previous models trained solely on mean losses, as illustrated in Fig. 6.14(d) and Table 6.7. This reduction in accuracy can be attributed to the substantially smaller dataset, which may still include residual variability due to the minimum sample size requirement. Additionally, it is important to recognize that the maximum possible measured no-load loss is theoretically unbounded, as various manufacturing or testing anomalies can occur. Consequently, predictions for maximum losses inherently exhibit higher errors compared with predictions for minimum and mean values, reflecting the intrinsic variability and stochastic nature of extreme outcomes.

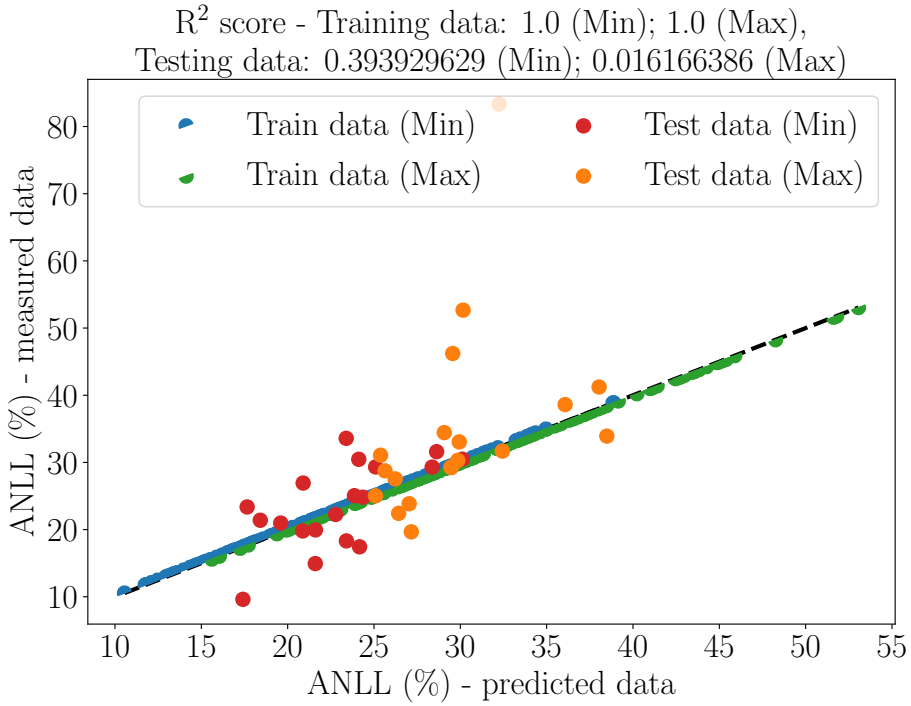


Fig. 6.16: Comparison of the GPR models accuracy in modeling extremes of additional no-load losses as minimum (Min) and maximum (Max), with the corresponding coefficients of determination (R^2) listed, adapted from [6].

6.5.7 Results for Learning Curve

The learning curve was generated through a structured procedure to assess the performance of the GPR models as a function of training dataset size. Initially, the entire dataset was filtered according to Section 6.5.1. The filtered dataset was then randomly partitioned into training and testing subsets in a 90:10 ratio, which was consistently applied across all GPR models considered in this thesis.

The starting point of the learning curve for both components was determined by training GPR models on only 10% of the training dataset, while keeping the testing dataset unchanged. Subsequently, the training set was incrementally expanded by 5% at each step, and the training and evaluation procedure was repeated until the full training dataset (100%) was utilized.

Given the sensitivity of the validation curve to specific random splits of training and testing data, the entire procedure was repeated over 100 iterations. The resulting learning curves are shown in Fig. 6.17(a). Across these iterations, the MSE values exhibited notable fluctuations as the training data size increased. Occasional rises in MSE were observed due to variability in the random splits, but the overall trend demonstrated clear convergence. Fig. 6.17(b) presents the averaged learning curve over all 100 iterations, revealing a smoother and more interpretable trend.

The analysis indicates that the presented GPR model is well-balanced and robust. The training error is effectively zero even at 10% of the training dataset, demonstrating the effectiveness of the selected surrogate model. Testing error shows consistent convergence as the training dataset grows, confirming the model's generalization capability. Moreover, the trend suggests that incorporating additional transformer designs in future datasets is likely to further reduce testing error, implying potential for even higher prediction accuracy and enhanced reliability in practical applications.

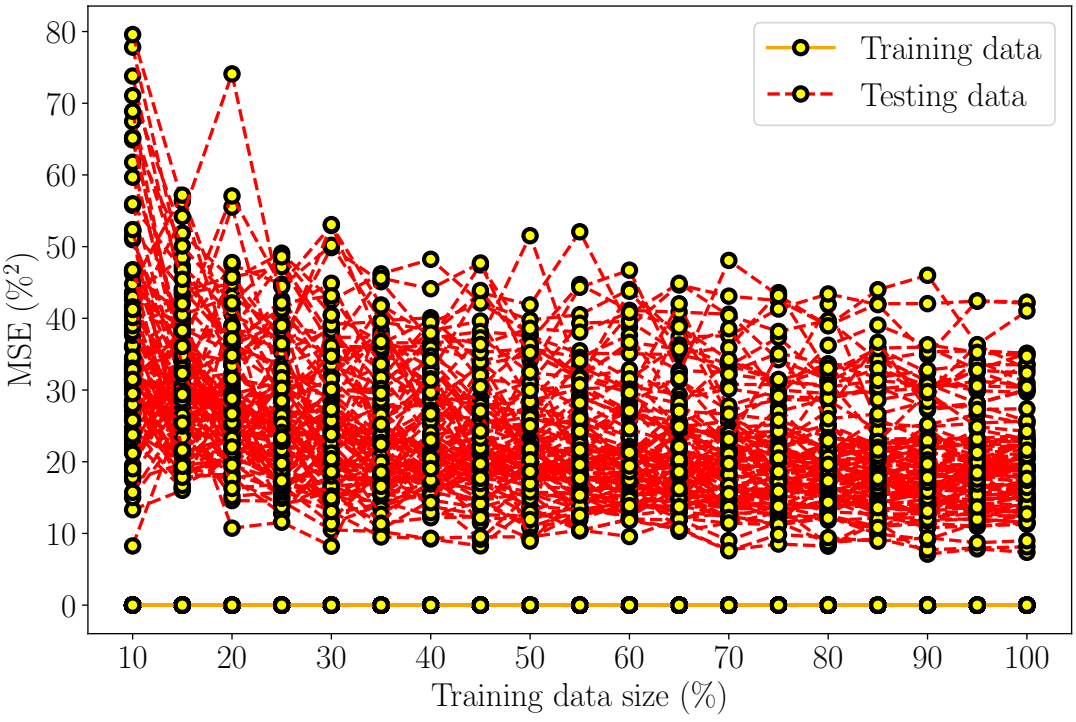
6.5.8 Final Discussion for Improved Models

The results demonstrate that the proposed modified approach is highly effective in predicting no-load losses in distribution transformers, provided a sufficient amount of training data is available. The key contribution of this work lies in the integration of well-established modeling and analytical techniques, including GPR, data filtration, and input optimization, into a domain where these methods have rarely been applied. This research introduces an extended and improved surrogate model as a follow-up to previous work. The new model optimizes the selection of input variable combinations to achieve the lowest MSE. The input variables considered for optimization included seven original variables, expanded to include five additional ones. The final model, achieving the best performance, uses nine variables and incorporates both continuous and categorical kernels.

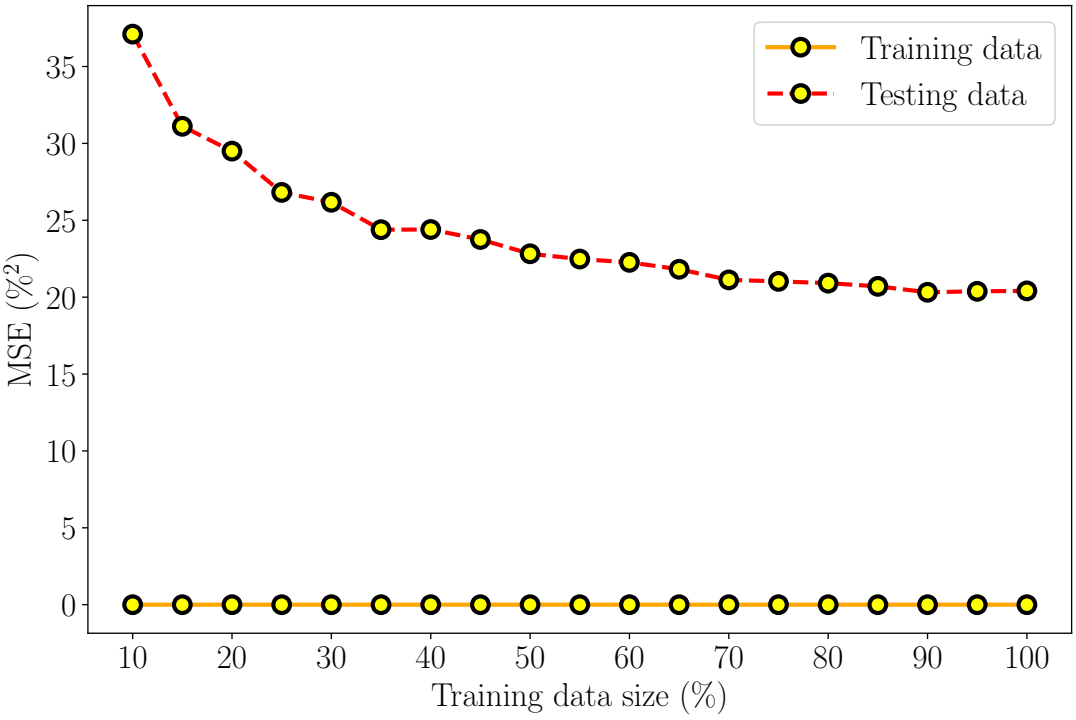
A substantial improvement in predicting no-load losses was achieved through data filtration. This procedure removes samples with high deviations, which can introduce significant errors during model training. For the original seven-variable model, data filtration reduced the MSE from $112.9\%^2$ to $17.1\%^2$ even in the worst-case scenario. For the new model with filtered data, the MSE decreased further to $15.2\%^2$ in the worst case and to $8.0\%^2$ in the best case, indicating a substantial increase in prediction accuracy. Compared with the conventional analytical method for calculating no-load losses, which exhibits an MSE of $80.7\%^2$, this represents a remarkable improvement.

In addition to predicting the mean value of losses, the method was extended to estimate the extremes of possible measured no-load losses, defining the maximum and minimum boundaries of the range. This extension was feasible only within the transformer learning and testing dataset, which included multiple measured units of the same design. Although the number of samples available for analyzing extremes is smaller, accurate prediction of maximum loss values is particularly important for mass production, often more so than estimating the mean value of losses.

The most significant advantage of the proposed method is its ability to account



(a)



(b)

Fig. 6.17: (a) Learning curves from 100 iterations of random dataset splits, (b) averaged learning curve computed from all iterations, showing a clearer trend, adapted from [6].

for all factors contributing to no-load losses, including those that cannot be simulated using conventional approaches. Additionally, the method delivers results rapidly, making it highly applicable to analytical calculations during the design process while maintaining high accuracy. Although the method is most effective when applied to data from cores built using the same technology, transformer manufacturers typically possess extensive datasets from their own production lines. This makes the developed method accessible and highly useful for most industrial applications. The model's accuracy improves as more data is incorporated and can be applied to any transformer core production technology.

6.6 Evaluation of GPR Against Other Machine Learning Techniques

To verify the suitability of GPR as the primary machine learning technique, surrogate models of average ANLL were also developed using alternative approaches, as listed in Section 4.2. These models were constructed in a manner consistent with the GPR framework, employing the same optimized set of nine input variables and the ANLL mean values as the output. Furthermore, all additional features and settings introduced in this thesis for the final GPR model were preserved, including the data filtration procedure and the 90:10 ratio used for splitting the dataset into training and testing subsets.

For comparative verification, six alternative machine learning techniques were selected. The results, evaluated using the first metric from Section 6.3.1, are presented as plots comparing the predicted and measured values for both training and testing samples in Fig. 6.18. The corresponding R^2 values are summarized in Table 6.9. Based on these results, it is evident that none of the other techniques provided predictions of comparable accuracy. In several cases, even the training data show considerable deviations from the ideal diagonal line, indicating model imprecision. Among the tested approaches, only random forest regression (RFR), multilayer perceptron (MLP), and extreme gradient boosting (XGBoost), shown in Figs. 6.18(b), (d), and (f), produce testing samples that partially align with the perfect prediction line, achieving R^2 values above 0.9 in training.

Regarding testing performance, the highest precision among the alternative models was achieved by the RFR, with $R^2 = 0.796964$, making it the second most suitable method after GPR. Nevertheless, the overall accuracy of RFR remains clearly inferior to that of GPR, which consistently outperformed all other tested techniques in both training and testing phases. The MLP model, despite a high R^2 in training, performed poorly on testing data, with a negative R^2 value, indicating

Table 6.9: Comparison of the Coefficient of Determination (R^2) Scores for Different Machine Learning Techniques Other than GPR

Technique	R^2**—Training data**	R^2**—Testing data**
Support vector regression (SVR)	0.698489	0.750561
Random forest regression (RFR)	0.946912	0.796964
Decision tree regression (DTR)	0.815512	0.567798
Multilayer perceptron (MLP)	0.954036	−1.384094
K-nearest neighbors (KNN)	0.615740	0.772583
Extreme gradient boosting (XGBoost)	0.937964	0.787466

severe overfitting. Similarly, other methods such as SVR, DTR, and KNN provided less reliable results, with considerably lower generalization capabilities.

These observations are consistent with findings reported in [102], [108], [109], which emphasize that methods like SVR, MLP, and XGBoost generally perform better with substantially larger datasets. In contrast, the characteristics of the present dataset, particularly the moderate size and the structure of the variables, align well with the strengths of GPR, as discussed in Section 4.2.4. Consequently, GPR is validated as the most effective and robust machine learning technique for the problem of no-load loss estimation in distribution transformers.

6.7 Summary of Chapter 6

This chapter presented the development of surrogate models for the estimation of transformer no-load losses, with a particular focus on Gaussian process regression (GPR) as the primary machine learning technique. The research was based on extensive experimental data collected from distribution transformers designed and manufactured over a wide timespan, providing a representative basis for statistical modeling. The dataset comprised measurements from 4 503 transformers corresponding to 322 unique designs, each record containing electrical and geometrical parameters together with repeated no-load loss measurements, enabling analysis of both average behavior and measurement variability.

The no-load test, which generated the measured data, was performed under rated voltage with the secondary winding open, capturing the core losses associated with magnetic flux in the transformer core. This standardized test ensured stable and repeatable values across a large number of designs, providing a reliable foundation for surrogate modeling.

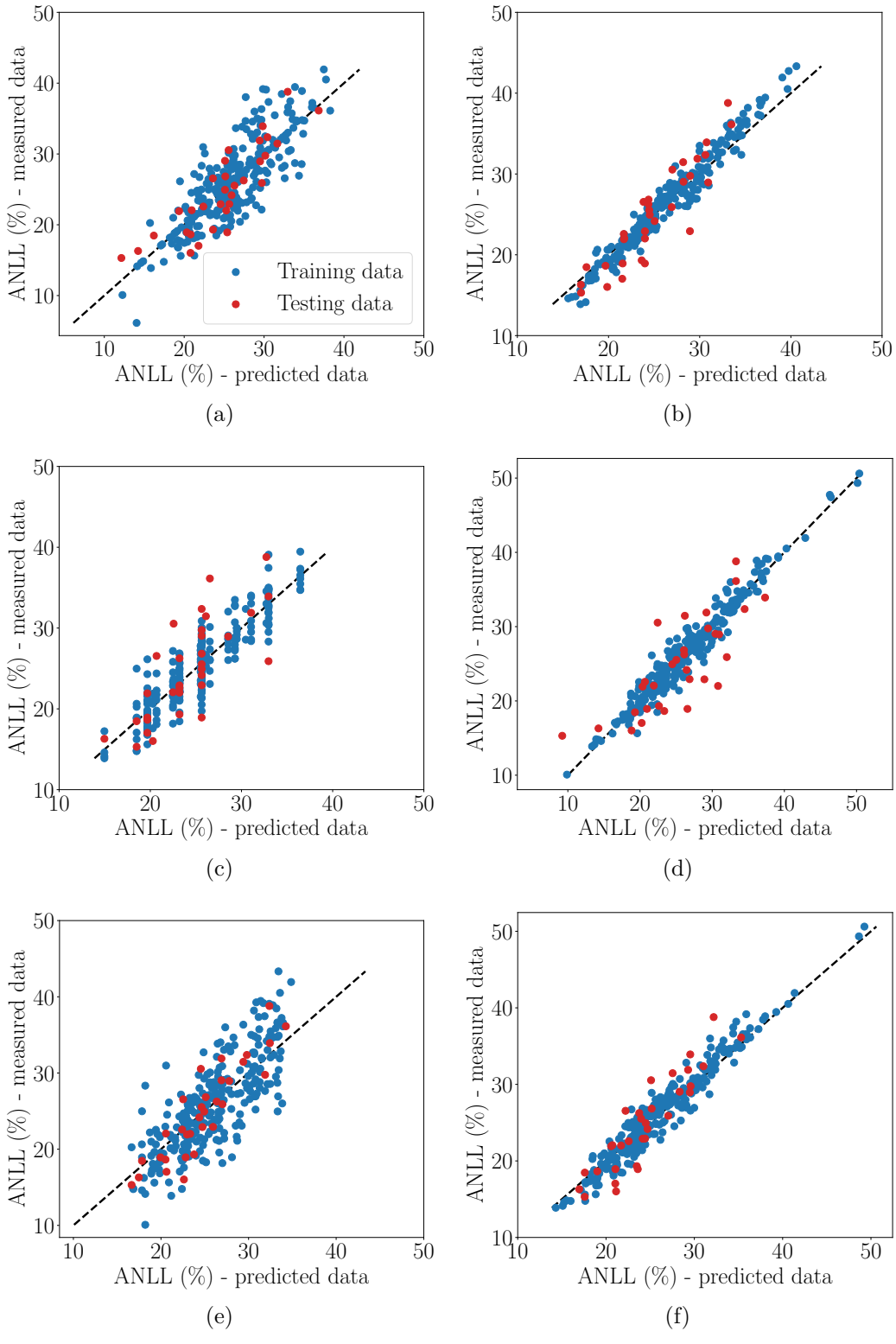


Fig. 6.18: Comparison of modeling ANLL using different machine learning techniques: (a) SVR, (b) RFR, (c) DTR, (d) MLP, (e) KNN, and (f) XGBoost.

Surrogate models were constructed using a systematic selection of input variables and careful kernel design. Continuous kernels were employed for numerical parameters, while a categorical kernel captured the effects of discrete features. Comparative testing of different kernel combinations demonstrated that mixed kernels, combining continuous and categorical components, consistently yielded higher predictive accuracy, as they effectively represented both numerical relationships and material-specific effects that purely numerical kernels could not capture. The kernel selection process was critical to ensuring model robustness and generalization, particularly given the moderate dataset size.

The initial surrogate modeling framework was developed by categorizing the input parameters into informative, basic, additional, and output variables. Three progressively refined models were formulated. The first was constructed with a minimal set of four fundamental inputs, the second incorporated three additional parameters to improve explanatory power, and the third employed a categorical kernel to capture the influence of the steel grade.

The improved framework, developed with the extended dataset, incorporated systematic filtration of outliers and inconsistent measurements to enhance data quality and model reliability. Additionally, an optimized selection of nine key input variables was performed to maximize predictive power while minimizing redundancy, ensuring that both continuous and categorical features were effectively represented. The surrogate models were designed not only to estimate the average additional no-load losses (ANLL) but also to predict the maximum and minimum deviations of repeated measurements within identical transformer designs, enabling quantification of uncertainty and supporting risk assessment during the design phase. Analysis of the learning curves demonstrated that model performance improved consistently with increasing dataset size, highlighting the strong dependence of GPR on adequate training data and confirming the robustness and generalization capability of the constructed surrogates.

Although alternative machine learning techniques such as random forest regression, multilayer perceptron, and extreme gradient boosting were also tested, none achieved the predictive accuracy or generalization capability of GPR for this dataset.

In conclusion, the chapter demonstrated that surrogate modeling based on measured transformer data provides an accurate and efficient alternative to analytical estimation of no-load losses. The GPR-based models, strengthened by systematic data preprocessing, optimized input selection, and careful kernel design, were validated as the most robust approach. They captured complex dependencies inherent in transformer design, provided enhanced reliability through uncertainty quantification, and offered practical applicability in design optimization, thereby establishing a strong methodological foundation for the subsequent parts of the thesis.

7 Conclusion

This doctoral thesis addressed the optimization of transformer design to meet increasingly stringent efficiency standards, such as the EcoDesign requirements of the European Union, while controlling the costs associated with increased material use to reduce losses. A fully automated design workflow was developed that integrates the simulated annealing (SA) algorithm with conventional analytical calculation models to systematically generate valid and cost-effective designs. Gaussian process regression (GPR) was employed to improve the prediction of no-load losses, allowing for more precise design margins and reducing material use across all transformer components, from the core to the tank. Together, these methods provide a practical, repeatable, and robust framework for modern transformer design.

The second chapter presented a comprehensive review of industrial transformer design practices, highlighting the combination of analytical methods, empirical corrections, structured computational models, and human experience. It detailed how input parameters derived from customer specifications guide material selection, component databases, and design variables, such as core geometry, winding configuration, and cooling arrangements, which are iteratively optimized to satisfy electrical, thermal, and mechanical requirements. Losses, short-circuit impedance, and thermal performance were estimated using analytical equations and empirical factors to ensure that the designs met technical constraints while remaining cost-effective. The chapter emphasized that, despite advances in simulation tools, industrial design continues to rely heavily on simplified models and engineering judgment, reinforcing the need for automated, optimization-driven approaches.

The third chapter focused on transformer no-load losses and conventional analytical calculation methods. Nominal core losses were estimated on the basis of core mass and specific lamination losses, then corrected using the building factor to account for additional losses. These additional losses arise from predictable factors, such as nonuniform flux distribution, lamination joints, stacking holes, and T-joints, as well as unpredictable effects including mechanical processing, steel tolerances, and assembly precision. Even transformers with identical designs can exhibit measurable differences in no-load losses, demonstrating the importance of considering both deterministic and stochastic influences. The chapter established that accurate estimation of no-load losses is critical for efficient transformer design and provides the foundation for data-driven and optimization-based methods.

The fourth chapter presented the theoretical underpinnings of the optimization and machine learning methods employed in the thesis. SA was applied as a robust and flexible algorithm capable of handling the complex, non-linear, and constrained design spaces typical of transformer engineering. To reduce computational effort, GPR was

introduced as a surrogate model for critical performance metrics, particularly no-load losses, offering accurate, interpretable predictions even with limited data. Integrating SA with GPR created a cohesive and automated framework that enables faster, more cost-effective, and technically compliant transformer designs while minimizing the reliance on manual adjustments.

The fifth chapter demonstrated the practical implementation of the automated SA-based workflow. The modular design of the optimizer was shown to efficiently explore the design space and produce feasible and cost-effective solutions, covering input parameters, design variables, iterative solution generation, constraint penalization, and final selection. Evaluations in multiple transformer ratings (50 kVA, 630 kVA, and 1,600 kVA) confirmed the algorithm's ability to quickly identify compliant designs while reducing material costs, with benefits especially evident for larger units. This work validated SA as a robust, repeatable and scalable methodology for modern transformer design, combining technical rigor with cost-effectiveness and reducing the dependence on the intuition of the designer.

The sixth and final chapter focused on the development and validation of surrogate models to predict transformer no-load losses, with GPR as the primary approach. Using an extensive data set of 4,503 transformers in 322 unique designs, the chapter outlined systematic model construction, including selection and further optimization of input parameters, kernel selection, and data quality improvement through the removal of outliers. Mixed kernels combining numerical and categorical features were essential to capture parametric and material-specific effects, enhancing predictive accuracy and generalization. The surrogate models were refined to predict not only average additional no-load losses but also measurement deviations, enabling uncertainty quantification and supporting risk assessment in the design process. Comparative evaluation with alternative machine learning methods, including random forest regression, multilayer perceptrons, and extreme gradient boosting, demonstrated GPR's superior accuracy and robustness. The results confirmed that data-driven surrogate modeling provides a reliable, practical and efficient alternative to conventional analytical methods, offering a strong foundation for integrating precise loss prediction into automated transformer design workflows.

In summary, this thesis presents a fully automated and data-driven transformer design methodology combining analytical techniques, robust optimization, and advanced surrogate modeling. By accurately predicting no-load and potentially load losses using GPR, exploring the design space, and incorporating manufacturing constraints, the framework enables cost-effective, high-performance transformer designs while minimizing reliance on empirical intuition. The work establishes a foundation for future research in high-efficiency transformer design and demonstrates the potential to integrate optimization and machine learning for industrial applications.

Bibliography

[1] P. Kellet, M. Oliva, M. Cuesto, and K. Pollari, "United to reduce global transformer losses – around the world, day and night", *Transformers Magazine*, vol. 7, no. 4, pp. 59–65, 2020.

[2] M. Cuesto, M. Oliva, and K. Pollari, "Driving down energy losses in transformers", *Transformers Magazine*, vol. 6, no. 4, pp. 67–71, 2019.

[3] *Commission regulation (EU) no 548/2014*, Brussels, 2014. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2014/548/oj.

[4] *Commission regulation (EU) no 2019/1783*, Brussels, 2019. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2019/1783/oj.

[5] A. D. Almeida, B. Santos, and F. Martins, "Energy-efficient distribution transformers in europe, Impact of EcoDesign regulation", *Energy Efficiency*, vol. 9, no. 2, pp. 401–424, 2016, ISSN: 1570-646X. DOI: 10.1007/s12053-015-9365-z. [Online]. Available: http://link.springer.com/10.1007/s12053-015-9365-z.

[6] M. Mrajca, V. Bilek, and J. Barta, "Transformer no-load losses estimation using machine learning from measured data", *IEEE Transactions on Industry Applications*, 2025.

[7] M. Mrajca and P. Silling, "Advanced transformer design optimization via simulated annealing", in *Proceedings I of the 31st Conference STUDENT EEICT 2025*, Brno, Czech Republic: Brno University of Technology, Faculty of Electrical Engineering and Communication, 2025, pp. 311–315, ISBN: 978-80-214-6321-9. [Online]. Available: https://www.eeict.cz/eeict_download/archiv/sborniky/EEICT_2025_sbornik_1.pdf.

[8] M. Mrajca, V. Bilek, J. Barta, and R. Cipin, "Transformer no-load losses calculation using machine learning", in *2024 International Conference on Electrical Machines (ICEM)*, Torino, Italy: IEEE, 2024, pp. 1–7, ISBN: 979-8-3503-7060-7. DOI: 10.1109/ICEM60801.2024.10700285. [Online]. Available: https://ieeexplore.ieee.org/document/10700285/.

[9] M. Mrajca, "Eddy currents additional losses comparison in oil distribution transformer windings", in *Proceedings I of the 29th Conference STUDENT EEICT 2023*, Brno, 2023, pp. 327–331, ISBN: 978-80-214-6153-6. [Online]. Available: https://www.eeict.cz/eeict_download/archiv/sborniky/EEICT_2023_sbornik_1.pdf.

[10] M. Mrajca, "Advanced turowski's method for eddy loss calculation in transformer foil winding", in *Proceedings I of the 29th Conference STUDENT EEICT 2023*, Brno, 2023, pp. 332–336, ISBN: 978-80-214-6153-6. [Online]. Available: https://www.eeict.cz/eeict_download/archiv/sborniky/EEICT_2023_sbornik_1.pdf.

[11] M. Mrajca and R. Cipín, "Axial channel effect on leakage reactance in split-winding transformer", in *Proceedings I of the 28th Conference STUDENT EEICT 2022*, Brno, 2022, pp. 331–335, ISBN: 978-80-214-6029-4. [Online]. Available: https://www.eeict.cz/eeict_download/archiv/sborniky/EEICT_2022_sbornik_1_v2.pdf.

[12] M. Mrajca, *Design of oil distribution transformer (Návrh olejového distribučního transformátoru)*, Master's thesis, Supervisor: Cestmir Ondrusek, Accessed: 2025-08-31, Brno, Czech Republic, 2021. [Online]. Available: https://www.vut.cz/studenti/zav-prace/detail/133160.

[13] M. Mrajca, *Production comparison of high voltage winding of dry transformers (Porovnání technologií výroby vinutí vysokého napětí suchých transformátorů)*, Bachelor's thesis, Supervisor: Cestmir Ondrusek, Accessed: 2025-08-31, Brno, Czech Republic, 2019. [Online]. Available: https://www.vut.cz/studenti/zav-prace/detail/119651.

[14] SGB-SMIT Group, *SGB-SMIT group – the pure-play transformer specialist from europe*, https://www.sgb-smit.com/, Accessed: 2025-08-31, 2025.

[15] ABB Group, *ABB – global technology leader in electrification and automation*, https://global.abb/, Accessed: 2025-08-31, 2025.

[16] Končar – Elektroindustrija d.d., *Končar – distribution and special transformers*, https://koncar.hr/en, Accessed: 2025-08-31, 2025.

[17] Lemi Trafo JSC, *Lemi Trafo – design and manufacture of three-phase distribution transformers*, Accessed: 2025-08-31, 2025. [Online]. Available: https://www.lemi-trafo.com/en/home/.

[18] Elpro-Energo Transformers s.r.o., *Elpro-Energo Transformers – manufacturer of top czech transformers*, https://www.czechtrafo.cz/en/, Accessed: 2025-08-31, 2025.

[19] E. I. Amoiralis, P. S. Georgilakis, M. A. Tsili, A. G. Kladas, and A. T. Souflaris, "Complete software package for transformer design optimization and economic evaluation analysis", *Materials Science Forum*, vol. 670, pp. 535–546, 2010, ISSN: 1662-9752. DOI: 10.4028/www.scientific.net/MSF.670.535. [Online]. Available: https://www.scientific.net/MSF.670.535.

[20] S. Kulkarni and S. Khaparde, *Transformer Engineering: Design, Technology, and Diagnostics*, 2nd Edition. Boca Raton: CRC Press, 2013, ISBN: 978-1-4398-5418-1.

[21] L. Li and X. Liu, "Calculation and analysis of short-circuit performance of a split-winding transformer with stabilizing windings", in *2018 IEEE International Conference on Applied Superconductivity and Electromagnetic Devices (ASEMD)*, 2018, pp. 1–2. DOI: 10.1109/ASEMD.2018.8558969. [Online]. Available: https://ieeexplore.ieee.org/document/8558969.

[22] C.-E. Salceanu, C. Dobrea, D. Ocoleanu, M. Nicola, D. Iovan, and M.-C. Nitu, "Experimental study of the dynamic short-circuit withstand capability of an 8400 kva power transformer specially designed for photovoltaic applications", *Machines*, vol. 11, no. 10, 2023, ISSN: 2075-1702. DOI: 10.3390/machines11100969. [Online]. Available: https://www.mdpi.com/2075-1702/11/10/969.

[23] Krempel GmbH, *Krempel – electrical insulation materials*, https://krempel.com/en, Accessed: 2025-08-31, 2025.

[24] Weidmann Group, *Weidmann Group – home*, https://www.weidmann-group.com/, Accessed: 2025-08-31, 2025.

[25] Tuboly-Astronic AG, *Tuboly-Astronic – machines for transformer production*, https://www.tuboly-astronic.ch/, Accessed: 2025-08-31, 2025.

[26] BR Technologies, *BR Technologies – coil winding machine experts*, https://www.br-tech.de/, Accessed: 2025-08-31, 2025.

[27] S. A. S. Alkadhim, "Three phase transformer: Connection and configuration", *SSRN Electronic Journal*, Jul. 2020. DOI: 10.2139/ssrn.3647143. [Online]. Available: https://ssrn.com/abstract=3647143.

[28] *IEC 60076-1:2011 - Power transformers – part 1: General*, https://webstore.iec.ch/en/publication/588, IEC Standard, International Electrotechnical Commission, 2011.

[29] *IEC 60076-11 - Power transformers — part 11: Dry-type transformers*, https://webstore.iec.ch/en/publication/29711, IEC Standard, International Electrotechnical Commission, 2018.

[30] S. Dohnal, *Comparison of additional losses calculation methods in transformer multilayer winding (Porovnání metod výpočtu přídavných ztrát ve vícevrstvém vinutí transformátoru)*, Bachelor's thesis, Supervisor: Miroslav Mrajca. Accessed: 2025-08-31, Brno, Czech Republic, 2024. [Online]. Available: https://www.vut.cz/studenti/zav-prace/detail/160627.

[31] J. Zdarsky, *Analysis of current higher harmonics impact on losses in transformer foil winding (Analýza vlivu vyšších harmonických proudu na ztráty ve foliovém vinutí transformátoru)*, Bachelor's thesis, Supervisor: Miroslav Mrajca. Accessed: 2025-08-31, Brno, Czech Republic, 2024. [Online]. Available: https://www.vut.cz/studenti/zav-prace/detail/160629.

[32] S. Trombik, *Loss and heat calculation of oil transformer top cover near bushings (Výpočet ztrát a oteplení víka olejového transformátoru v blízkosti průchodek)*, Supervisor: Miroslav Mrajca. Accessed: 2025-08-31, Brno, Czech Republic, 2024. [Online]. Available: https://www.vut.cz/studenti/zav-prace/detail/160628.

[33] R. Kuchler, *Transformers: Fundamentals of Their Calculation and Design (Die Transformatoren: Grundlagen für ihre Berechnung und Konstruktion)*, 2nd ed. Berlin: Springer, 1966, ISBN: 9783642524974.

[34] Ferokotao d.o.o., *Ferokotao – corrugated wall tanks for distribution transformers*, https://www.ferokotao.hr/, Accessed: 2025-08-31, 2025.

[35] E. Jezierski, *Transformers: Theoretical Fundamentals (Transformátory: teoretické základy)*. Prague: Academia, 1973, ISBN: 8320404878.

[36] *IEC 60076-2:2011 - Power transformers – part 2: Temperature rise for liquid-immersed transformers*, https://webstore.iec.ch/en/publication/599, IEC Standard, International Electrotechnical Commission, 2011.

[37] C. Wang, Q. Sun, A. Al-Abadi, and W. Wu, "A combined computational fluid dynamics and thermal–hydraulic modeling approach for improving the thermal performance of corrugated tank transformers: A comparative study", *Energies*, vol. 17, no. 8, 2024, ISSN: 1996-1073. DOI: 10.3390/en17081802. [Online]. Available: https://www.mdpi.com/1996-1073/17/8/1802.

[38] G. Liu, Z. Zheng, D. Yuan, L. Li, and W. Wu, "Simulation of fluid-thermal field in oil-immersed transformer winding based on dimensionless least-squares and upwind finite element method", *Energies*, vol. 11, no. 9, 2018, ISSN: 1996-1073. DOI: 10.3390/en11092357. [Online]. Available: https://www.mdpi.com/1996-1073/11/9/2357.

[39] Super-efficient Equipment and Appliance Deployment (SEAD) Initiative, "Distribution transformers report, part 2: Test method review", Clean Energy Ministerial, Tech. Rep., 2013.

[40] G. Bertotti, *Hysteresis in Magnetism: For Physicists, Materials Scientists, and Engineers*. Academic Press, 1998, ISBN: 9780120932702.

[41] F. Fiorillo, "Characterization and measurement of magnetic materials", *Academic Press (Monograph)*, 2004.

[42] AK Steel / Cleveland-Cliffs, "Selection of electrical steels for magnetic cores", AK Steel Corporation, Tech. Rep., 2016, Accessed 2025-08-31. [Online]. Available: https://www.brown.edu/Departments/Engineering/Courses/ENGN1931F/mag_cores_dataAKSteel-very%20good.pdf.

[43] E. Elgamli and F. Anayi, "Advancements in electrical steels: A comprehensive review of microstructure, loss analysis, magnetic properties, alloying elements, and the influence of coatings", *Applied Sciences*, vol. 13, no. 18, 2023, ISSN: 2076-3417. DOI: 10.3390/app131810283. [Online]. Available: https://www.mdpi.com/2076-3417/13/18/10283.

[44] C. P. Steinmetz, "On the law of hysteresis", *Transactions of the American Institute of Electrical Engineers*, 1892.

[45] International Electrotechnical Commission, *IEC 60404-8-7:2020 - magnetic materials – part 8-7: Specifications for individual materials – cold-rolled grain-oriented electrical steel strip and sheet delivered in the fully-processed state*, https://webstore.iec.ch/en/publication/63884, IEC Standard, 2020.

[46] Baosteel, "Baosteel technical manual for electrical steel products", Baosteel (Shougang Group), Technical Data Sheet, 2018, Accessed 2025-08-31. [Online]. Available: https://ecommerce.ibaosteel.com/portal/eleSteel/cn/pdf/technicalManual1118.pdf.

[47] N. Ashbahani, I. Daut, and M. I. Yusoff, "Building factor of grain-oriented silicon iron (3% sife) with different thickness on 100kva three phase distribution transformer core", *Advanced Materials Research*, vol. 488-489, pp. 537–541, 2012, ISSN: 1662-8985. DOI: 10.4028/www.scientific.net/AMR.488-489.537. [Online]. Available: https://www.scientific.net/AMR.488-489.537.

[48] I. A. Hernandez, J. M. Cañedo, J. C. Olivares-Galvan, and E. Betancourt, "Novel technique to compute the leakage reactance of three-phase power transformers", *IEEE Transactions on Power Delivery*, vol. 31, no. 2, pp. 437–444, 2016. DOI: 10.1109/TPWRD.2015.2412494. [Online]. Available: https://ieeexplore.ieee.org/document/7058430.

[49] T. Norio, X. Wang, Z. Cheng, L. Lin, and J. Wang, "Calculation and validation of iron loss in laminated core of power and distribution transformers", *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 33, no. 1/2, pp. 137–146, 2013, ISSN: 0332-1649. DOI: 10.1108/COMPEL-10-2012-0251. [Online]. Available:

https://www.emerald.com/insight/content/doi/10.1108/COMPEL-10-2012-0251/full/html.

[50] S. Mousavi, M. Shamei, A. Siadatan, F. Nabizadeh, and S. Mirimani, "Calculation of power transformer losses by finite element method", in *2018 IEEE Electrical Power and Energy Conference (EPEC)*, Toronto, Ontario, Canada: IEEE, 2018, pp. 1–5, ISBN: 978-1-5386-5419-4. DOI: 10.1109/EPEC.2018.8598292. [Online]. Available: https://ieeexplore.ieee.org/document/8598292/.

[51] P. Osinski and P. Witczak, "Analysis of core losses in transformer working at static var compensator", *Energies*, vol. 16, no. 12, 2023, ISSN: 1996-1073. DOI: 10.3390/en16124584. [Online]. Available: https://www.mdpi.com/1996-1073/16/12/4584.

[52] K. Dawood, I. Kaymaz, and S. Tursun, "Evaluation of no-load losses in the single-sheet, double-sheet, and triple-sheet step lap joints of the transformer core", in *Forthcoming Networks and Sustainability in the AIoT Era*, J. Rasheed, A. M. Abu-Mahfouz, and M. Fahim, Eds., Cham: Springer Nature Switzerland, 2024, pp. 14–24, ISBN: 978-3-031-62881-8. DOI: 10.1007/978-3-031-62881-8_2.

[53] E. Hajipour, P. Rezaei, M. Vakilian, and M. Ghafouri, "Power transformer no-load loss prediction with FEM modeling and building factor optimization", *Journal of Electromagnetic Analysis and Applications*, vol. 3, no. 10, pp. 430–438, 2011. DOI: 10.4236/jemaa.2011.310068.

[54] I. Hernandez, J. C. Olivares-Galvan, P. S. Georgilakis, and J. M. Cañedo, "Core loss and excitation current model for wound core distribution transformers", *International Transactions on Electrical Energy Systems*, vol. 24, no. 1, pp. 30–42, 2014. DOI: 10.1002/etep.1687. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/etep.1687.

[55] K. Dawood and S. Tursun, "Effects of the step-lap joints configurations on the no-load losses of the transformer", in *2023 12th International Conference on Power Science and Engineering (ICPSE)*, Eskisehir, Turkiye: IEEE, 2023, pp. 60–64, ISBN: 979-8-3503-4238-3. DOI: 10.1109/ICPSE59506.2023.10329298. [Online]. Available: https://ieeexplore.ieee.org/document/10329298/.

[56] K. Karsai, D. Kerenyi, and L. Kiss, *Large Power Transformers*, Rev. engl. version. Budapest: Akademiai Kiado, 1987, ISBN: 9630541122.

[57] C. Modaragamage, A. Shanmuganathan, H. Liyanage, H. Senevirathne, and R. Lucas, "Effect of size and position of pin holes on transformer core loss", in *KDU IRC 2020 (13th Internationa Research Conference)*, Oct. 2020, pp. 150–158. [Online]. Available: https://ir.kdu.ac.lk/handle/345/2971.

[58] J. Kozlicek, *Analysis of additional losses in transformer sheets caused by stacking holes (Analyza pridavnych ztrat v transformatorovych plechsich vlivem stohovacich der)*, Bachelor's thesis, Supervisor: Miroslav Mrajca. Accessed: 2025-08-31, Brno, Czech Republic, 2024. [Online]. Available: https://www.vut.cz/studenti/zav-prace/detail/160630.

[59] A. J. Moses, "Importance of rotational losses in rotating machines and transformers", *Journal of Materials Engineering and Performance*, vol. 1, no. 2, pp. 235–244, 1992, ISSN: 1059-9495. DOI: 10.1007/BF02648622. [Online]. Available: http://link.springer.com/10.1007/BF02648622.

[60] S. Zurek, "Two-dimensional magnetisation problems in electrical steels", Supervisor: T. Meydan. Accessed: 2025-08-31, Doctoral thesis, Wolfson Centre for Magnetics Technology, Cardiff School of Engineering, Cardiff University, Mar. 2005. DOI: 10.13140/RG.2.2.33615.76960.

[61] L. Zhu, X. Zhang, Q. Yang, and X. Zhang, "Design and analyze the optimum operating point between magnetic flux density and vibration noise of transformer cores", *The Open Electrical & Electronic Engineering Journal*, vol. 8, pp. 552–558, Dec. 2014. DOI: 10.2174/1874129001408010552.

[62] E. teNyenhuis, R. Girgis, and G. Mechler, "Other factors contributing to the core loss performance of power and distribution transformers", *IEEE Transactions on Power Delivery*, vol. 16, no. 4, pp. 648–653, 2001, ISSN: 08858977. DOI: 10.1109/61.956752. [Online]. Available: http://ieeexplore.ieee.org/document/956752/.

[63] H. G. G. Maschinenfabrik, *Georg — specialist in coil processing & roll grinding*, https://www.georg.com/en/, Accessed: 2025-08-31, 2025.

[64] R. Corin, J.-P. Lecointe, C. Demian, and J. Blaszkowski, "Analysis of the grain oriented electrical steel heterogeneities and their influence on the epstein frame loss measurements", *International Journal of Applied Electromagnetics and Mechanics*, vol. 69, no. 2, pp. 293–302, 2022. DOI: 10.3233/JAE-210184. [Online]. Available: https://journals.sagepub.com/doi/abs/10.3233/JAE-210184.

[65] S. Wada, T. Yagisawa, and N. Asai, "Building factors of transformer cores made from low loss Si-steel sheets", *Physica Scripta*, vol. 39, no. 3, pp. 403–405, 1989, ISSN: 0031-8949. DOI: 10.1088/0031-8949/39/3/029. [Online]. Available: https://iopscience.iop.org/article/10.1088/0031-8949/39/3/029.

[66] F. Glover, "Tabu search—part i", *ORSA Journal on Computing 1*, vol. 3, pp. 190–206, Aug. 1989, ISSN: 0899-1499. DOI: 10.1287/ijoc.1.3.190. [Online]. Available: https://pubsonline.informs.org/doi/10.1287/ijoc.1.3.190.

[67] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing", *Science*, vol. 220, no. 4598, pp. 671–680, 1983. DOI: 10.1126/science.220.4598.671. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.220.4598.671.

[68] L. Yang and Q. Li, "Research on logistics distribution vehicle path optimization based on simulated annealing algorithm", *Advances in Multimedia*, vol. 2022, pp. 1–8, 2022-5-6, ISSN: 1687-5699. DOI: 10.1155/2022/7363279. [Online]. Available: https://www.hindawi.com/journals/am/2022/7363279/.

[69] M. Fernández-Delgado, M. Sirsat, E. Cernadas, S. Alawadi, S. Barro, and M. Febrero-Bande, "An extensive experimental survey of regression methods", *Neural Networks*, vol. 111, pp. 11–34, 2019, ISSN: 0893-6080. DOI: 10.1016/j.neunet.2018.12.010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608018303411.

[70] C. Ling, W. Kuo, and M. Xie, "An overview of adaptive-surrogate-model-assisted methods for reliability-based design optimization", *IEEE Transactions on Reliability*, vol. 72, no. 3, pp. 1243–1264, 2023. DOI: 10.1109/TR.2022.3200137. [Online]. Available: https://ieeexplore.ieee.org/document/9875995.

[71] V. L. Deringer, A. P. Bartok, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csanyi, "Gaussian process regression for materials and molecules", *Chemical Reviews*, vol. 121, no. 16, pp. 10 073–10 141, 2021. DOI: 10.1021/acs.chemrev.1c00022. [Online]. Available: https://doi.org/10.1021/acs.chemrev.1c00022.

[72] K. Liu, X. Hu, Z. Wei, Y. Li, and Y. Jiang, "Modified gaussian process regression models for cyclic capacity prediction of lithium-ion batteries", *IEEE Transactions on Transportation Electrification*, vol. 5, no. 4, pp. 1225–1236, 2019. DOI: 10.1109/TTE.2019.2944802. [Online]. Available: https://ieeexplore.ieee.org/document/8853281.

[73] J. Wang, "An intuitive tutorial to gaussian processes regression", *Computing in Science & Engineering*, pp. 1–8, 2023. DOI: 10.1109/MCSE.2023.3342149.

[74] GPyTorch Developers, *GPyTorch documentation (v1.11)*, https://gpytorch.ai, Accessed: 2025-08-31, 2024.

[75] H. Chen, L. Ding, and R. Tuo, "Kernel packet: An exact and scalable algorithm for gaussian process regression with matern correlations", *J. Mach. Learn. Res.*, vol. 23, no. 1, Jan. 2022, ISSN: 1532-4435. DOI: 10.48550/arXiv.2203.03116. [Online]. Available: https://arxiv.org/abs/2203.03116.

[76] A. Ghasempour and M. Martínez-Ramón, "Electric load forecasting using multiple output gaussian processes and multiple kernel learning", in *2023 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, 2023, pp. 1–6. DOI: 10.1109/ISIEA58478.2023.10212291. [Online]. Available: https://ieeexplore.ieee.org/document/10212291.

[77] M. G. Genton, "Classes of kernels for machine learning: A statistics perspective", *Journal of Machine Learning Research*, vol. 2, pp. 299–312, 2001. [Online]. Available: https://jmlr.org/papers/v2/genton01a.html.

[78] R. Boloix-Tortosa, J. J. Murillo-Fuentes, F. J. Payán-Somet, and F. Pérez-Cruz, "Complex gaussian processes for regression", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5499–5511, 2018. DOI: 10.1109/TNNLS.2018.2805019. [Online]. Available: https://ieeexplore.ieee.org/document/8307269.

[79] A. McHutchon and C. E. Rasmussen, "Gaussian process training with input noise", in *NIPS'11:Proceedings of the 25th International Conference on Neural Information Processing Systems*, Granada, Spain: Curran Associates Inc., 2011, pp. 1341–1349, ISBN: 9781618395993.

[80] Z. Liu, C. Lyu, J. Huo, S. Wang, and J. Chen, "Gaussian process regression for transportation system estimation and prediction problems: The deformation and a hat kernel", *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 22 331–22 342, 2022. DOI: 10.1109/TITS.2022.3155527. [Online]. Available: https://ieeexplore.ieee.org/document/9788557.

[81] M. Liu, G. Chowdhary, B. Castra da Silva, S.-Y. Liu, and J. P. How, "Gaussian processes for learning and control: A tutorial with examples", *IEEE Control Systems Magazine*, vol. 38, no. 5, pp. 53–86, 2018. DOI: 10.1109/MCS.2018.2851010.

[82] M. M. Noack, H. Krishnan, M. D. Risser, and K. G. Reyes, "Exact gaussian processes for massive datasets via non-stationary sparsity-discovering kernels", *CoRR*, vol. abs/2205.09070, 2022. DOI: 10.21203/rs.3.rs-1746398/v1. [Online]. Available: https://arxiv.org/abs/2205.09070.

[83] K. McBride and K. Sundmacher, "Overview of surrogate modeling in chemical process engineering", *Chemie Ingenieur Technik*, vol. 91, no. 3, pp. 228–239, 2019. DOI: 10.1002/cite.201800091. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cite.201800091.

[84] H. Mu, K. Shen, and C. Ma, "Introduction of thin gauge grain-oriented electrical steel products of baosteel and their application in power transformers", *Electrical Steel*, vol. 3. No. 4. Pp. 7–11, 2021, ISSN: 2096-7101. [Online]. Available: http://www.bwjournal.com/dgg/EN/Y2021/V3/I4/7.

[85] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines", *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953-6-1, ISSN: 0021-9606. DOI: 10.1063/1.1699114. [Online]. Available: https://pubs.aip.org/aip/jcp/article-abstract/21/6/1087/202680/Equation-of-State-Calculations-by-Fast-Computing.

[86] C. Watkins and M. McAleer, "Pricing of non-ferrous metals futures on the london metal exchange", *Applied Financial Economics*, vol. 8, no. 2, pp. 127–141, 1998. DOI: 10.1080/096031098333275. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/09603100600756514.

[87] P.-S. Chow, T.-M. Choi, and T. C. E. Cheng, "Impacts of minimum order quantity on a quick response supply chain", *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, vol. 42, no. 4, pp. 868–879, 2012. DOI: 10.1109/TSMCA.2012.2183351. [Online]. Available: https://ieeexplore.ieee.org/document/6155096.

[88] D. Snell, "Noise generated by model step lap core configurations of grain oriented electrical steel", *Journal of Magnetism and Magnetic Materials*, vol. 320, no. 20, pp. 887–890, 2008, ISSN: 0304-8853. DOI: 10.1016/j.jmmm.2008.04.169. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0304885308005209.

[89] Z. Valkovic and A. Rezic, "Improvement of transformer core magnetic properties using the step-lap design", *Journal of Magnetism and Magnetic Materials*, vol. 112, no. 1-3, pp. 413–415, 1992, ISSN: 03048853. DOI: 10.1016/0304-8853(92)91216-G. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/030488539291216G.

[90] *IEC 60076-8:1997 - Power transformers – part 8: Application guide*, https://webstore.iec.ch/en/publication/606, IEC Standard, International Electrotechnical Commission, 1997.

[91] F. Brailsford, "Current and power relationships in the measurement of iron losses in a three-limb transformer core", *Proceedings of the IEE - Part II: Power Engineering*, vol. 101, pp. 409–416, 82 1954. DOI: 10.1049/pi-2.1954.0090. [Online]. Available: https://digital-library.theiet.org/doi/abs/10.1049/pi-2.1954.0090.

[92] A. Carlson, J. Fuhr, G. Schemel, and F. Wegscheider, *Testing of Power Transformers*. Duesseldorf, Germany: ABB Handbook, Proprint GmbH, 2003, ISBN: 3-00-010400-3.

[93] E. Woschnagg, "Unsymmetrische Leistungsverteilung an leerlaufenden Grosstransformatoren und Drosselspulen [asymmetrical power distribution in unloaded large transformers and reactors]", German, *Brown Boveri Mitteilungen*, vol. 56, no. 7, pp. 314–322, 1969.

[94] "IEEE standard test code for liquid-immersed distribution, power, and regulating transformers", *IEEE Std C57.12.90-2021 (Revision of IEEE Std C57.12.90-2015)*, pp. 1–119, 2022. DOI: 10.1109/IEEESTD.2022.9707753. [Online]. Available: https://ieeexplore.ieee.org/document/9707753.

[95] J. R. Gardner, G. Pleiss, D. Bindel, K. Q. Weinberger, and A. G. Wilson, "Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration", in *32nd Conference on Neural Information Processing Systems (NeurIPS 2018)*, 2018. DOI: 10.48550/arXiv.1809.11165. [Online]. Available: https://arxiv.org/abs/1809.11165.

[96] D. C. Liu and J. Nocedal, "On the limited memory bfgs method for large scale optimization", *Mathematical Programming*, vol. 45, no. 1, pp. 503–528, 1989. DOI: 10.1007/BF01589116. [Online]. Available: https://link.springer.com/article/10.1007/BF01589116.

[97] J. Wågberg, D. Zachariah, T. B. Schön, and P. Stoica, *Prediction performance after learning in gaussian process regression*, Accessed: 2025-08-31, 2017. DOI: https://doi.org/10.48550/arXiv.1606.00386. [Online]. Available: https://arxiv.org/abs/1606.03865.

[98] O. Renaud and M.-P. Victoria-Feser, "A robust coefficient of determination for regression", *Journal of Statistical Planning and Inference*, vol. 140, no. 7, pp. 1852–1862, 2010, ISSN: 0378-3758. DOI: 10.1016/j.jspi.2010.01.008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378375810000194.

[99] A. Wang, R. Murdock, S. Kauwe, *et al.*, "Machine learning for materials scientists: An introductory guide towards best practices", *Chemistry of Materials*, vol. 32, 12 2020. DOI: 10.26434/chemrxiv.12249752.v1. [Online]. Available: https://chemrxiv.org/engage/chemrxiv/article-details/60c74ab3337d6ccdf3e27961.

[100] A. Botchkarev, "A new typology design of performance metrics to measure errors in machine learning regression algorithms", *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 14, pp. 045–076, 2019, ISSN: 45-76. DOI: 10.28945/4184. [Online]. Available: https://www.informingscience.org/Publications/4184.

[101] R. Trinchero, M. Larbi, H. M. Torun, F. G. Canavero, and M. Swaminathan, "Machine learning and uncertainty quantification for surrogate models of integrated devices with a large number of parameters", *IEEE Access*, vol. 7, pp. 4056–4066, 2019. DOI: 10.1109/ACCESS.2018.2888903. [Online]. Available: https://ieeexplore.ieee.org/document/8584446.

[102] H. Taghavifar and A. Mardani, "Gaussian process with automatic relevance determination predictive model for energy management of electric direct-drive wheels: Experimental validation", *IEEE Transactions on Vehicular Technology*, vol. 73, no. 2, pp. 1910–1917, 2024. DOI: 10.1109/TVT.2023.3316917. [Online]. Available: https://ieeexplore.ieee.org/document/10258308.

[103] I. Roman, R. Santana, A. Mendiburu, and J. A. Lozano, "Evolving gaussian process kernels for translation editing effort estimation", in *Learning and Intelligent Optimization*, N. F. Matsatsinis, Y. Marinakis, and P. Pardalos, Eds., Cham: Springer International Publishing, 2020, pp. 304–318, ISBN: 978-3-030-38629-0. DOI: 10.1007/978-3-030-38629-0_25. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-38629-0_25.

[104] C. Tantithamthavorn, S. McIntosh, A. E. Hassan, and K. Matsumoto, "An empirical comparison of model validation techniques for defect prediction models", *IEEE Transactions on Software Engineering*, vol. 43, no. 1, pp. 1–18, 2017. DOI: 10.1109/TSE.2016.2584050. [Online]. Available: https://ieeexplore.ieee.org/document/7497471.

[105] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of bayesian optimization", *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016. DOI: 10.1109/JPROC.2015.2494218. [Online]. Available: https://ieeexplore.ieee.org/document/7352306.

[106] T. Viering and M. Loog, "The shape of learning curves: A review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, pp. 7799–7819, 2023. DOI: 10.1109/TPAMI.2022.3220744. [Online]. Available: https://ieeexplore.ieee.org/document/9944190.

[107] H. Shan, *Understanding bias-variance trade-off from learning curve*, Accessed: 2025-08-31, Feb. 2021. [Online]. Available: https://hshan0103.medium.com/understanding-bias-variance-trade-off-from-learning-curve-a64b4223bb02.

[108] Y. Zhou, Y. Liu, D. Wang, and X. Liu, "Comparison of machine-learning models for predicting short-term building heating load using operational parameters", *Energy and Buildings*, vol. 253, p. 111 505, 2021, ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2021.111505. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778821007891.

[109] J. L. A. Gardner, Z. Faure Beaulieu, and V. L. Deringer, "Synthetic data enable experiments in atomistic machine learning", *Digital Discovery*, vol. 2, pp. 651–662, 3 2023. DOI: 10.1039/D2DD00137C. [Online]. Available: https://pubs.rsc.org/en/content/articlelanding/2023/dd/d2dd00137c.

Symbols and Abbreviations

Symbols

AC	power frequency applied voltage
A	transformer core cross-section
B	magnetic flux density (in transformer core)
BIL	basic impulse level voltage
c	categorical input
$c_k(\mathbf{x})$	nonnegative penalty function
d	relative difference between the measured voltages
$d(x_i, x_j)$	Euclidean distance between samples
$\delta(x_i^{(k)}, x_j^{(k)})$	Kronecker delta function
E	induce voltage
$\hat{E}$	peak applied voltage
f	frequency
$f(x)$	function to be modeled
$f'(\mathbf{x})$	penalized function
$\hat{H}$	peak magnetic field intensity
H_{rms}	root-mean-square magnetic field intensity
I	identity matrix
I_0	no-load current during no-load test
I_2	transformer secondary current
I_{10}	transformer no-load primary current
I_μ	transformer magnetizing current
K	kernel function

K_{Cat}	kernel function for categorical input components
K_{Cont}	kernel function for continuous input components
k	constraint
k_{a}	constant for anomalous (excess) losses calculation
k_{e}	constant for eddy current losses calculation
k_{h}	constant for hysteresis losses calculation
k_{BF}	building factor (for core losses calculation)
l	lengthscale hyperparameter
N	number of turns of the transformer winding
$\mathcal{N}$	Gaussian (normal) distribution
n	number of training points
$O(n^2)$	memory size
$O(n^3)$	number of operations
P_{10}	transformer no-load losses
P_{a}	anomalous (excess) losses
P_{e}	eddy current losses
P_{Fe}	transformer core (steel) losses
$P_{\text{Fe,n}}$	nominal transformer core losses
P_{h}	hysteresis losses
P_{j10}	transformer no-load Joule loss
P_{m}	measured no-load losses
P_0	sinusoidal waveform equivalent no-load losses
P_t	probability of accepting new candidate
R_1	transformer primary resistance
R^2	coefficient of determination

U_{m}	highest voltage for equipment
U_{rms}	root-mean-square phase voltage
$\bar{U}$	mean voltage recalculated to sinusoidal root-mean-square
x	continous input
x_i	input sample i
x_j	input sample j
x_*	new input
$\mathbf{x}_t$	current solution
$\mathbf{x}_0$	initial solution
$\mathbf{x}'$	new candidate solution
$\mathbf{x}^*$	optimal solution
$\mathbb{R}^n$	search space of n dimensions
ϕ	total magnetic flux
μ	mean function
μ_*	posterior mean function
σ	covariance function
$\sigma_{\mathbf{f}}$	covariance scaling factor
σ_*	posterior covariance function
σ_n^2	noise variance
t	iteration in discrete time step
T_t	temperature parameter at iteration
w_k	penalty weight
y	observing
$p(y\|f(x))$	likelihood of observing
$\hat{B}$	peak magnetic flux density (in transformer core)

Abbreviations

2D	two-dimensional
AC	alternate current
ANLL	additional no-load losses
ANN	artificial neural networks
BUT	Brno University of Technology
CFD	computational fluid dynamics
CO	constrained optimization
CRGO	cold-rolled grain-oriented (steel)
CT	current transformer
DC	direct current
DPP	diamond pattern paper
DTR	decision tree regression
EU	European Union
FEM	finite element method
GP	Gaussian process
GPR	Gaussian process regression
HiB	high permeability (steel)
IEC	International Electrotechnical Commission
INLP	integer nonlinear programming
KNN	k-nearest neighbors algorithm
L-BFGS	Limited-memory Broyden–Fletcher–Goldfarb–Shanno (algorithm)
LME	London Metal Exchange

LV	low voltage
MAE	mean absolute error
MIP	mixed-integer programming
MOQ	minimum order quantity
MLP	multilayer perceptron
MV	medium voltage
PEI	peak efficiency index
RBF	radial basis function
RMSE	root mean squared error
RFR	random forest regression
RQ	rational quadratic
RMS	root-mean-square
SA	simulated annealing
SR	standardized residual
SST	single sheet test
STD	standard deviation
SVR	support vector regression
VT	voltage transformer
XGBoost	extreme gradient boosting