Politécnico
de Viseu

Escola Superior
de Tecnologia
e Gestão de Viseu

Smart Kiln Drying: Leveraging AI for Precision in Moisture Management

José Afonso de Almeida Ramos

Trabalho de Projeto
Mestrado em Engenharia Informática - Sistemas de Informação

Trabalho efetuado sob a orientação de
Professor Doutor Rui Pedro Monteiro Amaro Duarte

Fevereiro de 2026

Politécnico
de Viseu

Escola Superior
de Tecnologia
e Gestão de Viseu

Smart Kiln Drying: Leveraging AI for Precision in Moisture Management

José Afonso de Almeida Ramos

Trabalho de Projeto
Mestrado em Engenharia Informática - Sistemas de Informação

Trabalho efetuado sob a orientação de
Professor Doutor Rui Pedro Monteiro Amaro Duarte

Fevereiro de 2026

To my family,
who believed in me even when I doubted myself...

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Rui Pedro Duarte, for his guidance, critical insights, and continuous support throughout the development of this thesis. His availability, technical rigor, and constructive feedback were essential to the successful completion of this work.

I am also deeply thankful to SONAE Arauco for providing the industrial data and the opportunity to work on a real-world case study. In particular, I would like to acknowledge the factory director, José António Silva, and my contact point from SONAE, Leonardo Porto, for their unwavering help and unconditional availability to discuss and solve the problems that arose when writing this thesis. I would also like to mention the engineers and technical staff who shared their operational knowledge and practical expertise, which significantly contributed to the understanding of the drying process and to the interpretation of the results.

My appreciation also extends to the faculty members and colleagues who contributed through discussions, suggestions, and academic support during the course of this project.

Finally, I would like to thank my family and friends, especially my mother, father and girlfriend for their patience, encouragement, and unwavering support even on the hardest parts of this journey. Their motivation and understanding were fundamental during the most demanding phases of this work.

Abstract

This research studies the application of data-driven *Machine Learning* (ML) techniques to the modeling and analysis of an industrial wood drying process. Effective management of *Moisture Content* (MC) plays an essential role in the wood processing industry to ensure that product quality and operational efficiency is achieved. However, direct MC measurements in industrial drying systems are often sparse and sometimes inconsistently recorded or subject to uncertainty which limits their direct use. To address these limitations, temperature was adopted as a proxy variable for the drying state, leveraging its strong physical relationship with MC removal and its reliable, high-frequency availability in industrial environments. Using historical operational data from a traditional industrial drying system, a structured modeling framework is developed and evaluated offline. Two machine learning approaches are implemented and compared: a feedforward Artificial Neural Networks (ANN) and a Long Short-Term Memory (LSTM) network. The ANN model represents a static nonlinear regression approach based on instantaneous process snapshots, while the LSTM model explicitly captures temporal dependencies through sequential input representations. Both models are evaluated using chronologically split datasets to assess their generalization capability under realistic industrial conditions. In addition to predictive modeling, an offline control recommendation framework is introduced to explore how the predictive models respond to variations in controllable process variables. This framework enables the comparative analysis of model sensitivity, stability, control effort, and recommendation consistency without deploying a real-time or closed-loop control system. The results demonstrate that while ANN-based models can generate larger instantaneous corrective responses, their behavior is more variable and less stable in control-oriented scenarios. In contrast, LSTM-based models exhibit improved generalization and more consistent, conservative behavior due to their ability to incorporate process history. These findings highlight the importance of temporal modeling for advisory and decision support applications in industrial wood drying processes.

Keywords: Wood Drying, Machine Learning, Artificial Neural Networks, Long Short-Term Memory, Industrial Processes, Temperature Modeling, Moisture Content

Resumo

Esta investigação estuda a aplicação de técnicas de ML orientadas por dados à modelação de um processo industrial de secagem de madeira. A gestão eficaz do MC é essencial para garantir qualidade do produto e eficiência operacional. No entanto, as medições diretas de MC em sistemas industriais são frequentemente escassas ou inconsistentes, limitando a sua utilização direta em modelos de elevada resolução temporal.

Para ultrapassar estas limitações, a temperatura foi adotada como variável *proxy* do estado de secagem, explorando a sua relação física com a remoção de MC e a sua elevada disponibilidade industrial. Com base em dados históricos de operação, foi desenvolvido e avaliado offline um enquadramento estruturado de modelação.

Foram implementadas duas abordagens: uma ANN feedforward e uma rede LSTM. A ANN representa uma regressão não linear estática baseada em instantâneos do processo, enquanto a LSTM capta dependências temporais através de entradas sequenciais. Ambos os modelos foram avaliados com divisão cronológica dos dados para analisar a sua capacidade de generalização em condições industriais realistas.

Foi ainda introduzido um enquadramento offline de recomendações de controlo para analisar a resposta dos modelos a variações nas variáveis controláveis. Este enquadramento permite comparar sensibilidade, estabilidade e esforço de controlo, sem implementação em tempo real.

Os resultados mostram que, embora a ANN produza correções de maior magnitude, apresenta comportamento mais variável. A LSTM demonstra melhor generalização e maior consistência, evidenciando a relevância da modelação temporal em aplicações de apoio à decisão na secagem industrial de madeira.

Palavras-Chave: Secagem de Madeira, Machine Learning, Redes Neuronais, Memórias de Curto-Longo Prazo, Processos industriáis, Modelagem de Temperatura, Humidade

Contents

List of Tables

List of Figures

List of Acronyms

Adam	Adaptive Moment Estimation
AI	*Artificial Intelligence*
ANN	Artificial Neural Networks
CV	*Computer Vision*
DLASS	Deep LSTM Autoencoder Minimizing Perturbed Error
EMC	*Equilibrium Moisture Content*
KPI	Key Performance Indicators
LSTM	Long Short-Term Memory
LSTM AE	Deep Long Short-Term Memory Autoencoders
MAE	Mean Absolute Error
MC	*Moisture Content*
MDF	*Medium-density Fibreboard*
ML	*Machine Learning*
MSE	Mean Squared Error
NIR	*Near-Infrared Spectroscopy*
NIR-HSI	Near-Infrared Hyperspectral Imaging
$\mathbf{R}^2$	Coefficient of Determination
RGB	*Red, Green, Blue*
RMSE	Root Mean Sqared Error
RNN	Recurrent Neural Network

Chapter 1

Introduction

The wood processing industry plays an essential role in modern manufacturing, supplying raw materials for construction, furniture production, and engineered wood products among others. Due to its scale and economic relevance, this sector faces increasing pressure to improve process efficiency, ensure consistent product quality, and reduce its environmental impact. Among the many challenges that are encountered in wood processing, the control and reduction of *Moisture Content* (MC) during drying operations remains as a particularly important issue that needs to be addressed. Proper management of MC is essential to ensure that dimensional stability, mechanical performance, and long term durability and integrity of wood products. Issues in the drying process, including uneven drying, can result in serious defects on the pieces composed of the processed wood, such as warping, cracking, or even internal stress that may lead to future damage. These defects would cause material waste, increased operational costs and environmental worries. For these reasons, industrial drying processes must operate under carefully controlled conditions to achieve target MC levels in a reliable, repeatable, and sustainable manner. Kiln drying is one of the most widely adopted industrial techniques when it comes for drying and MC removal, whether applied to wood or other products, such as those in the food processing industry, and it employs controlled combinations of temperature, airflow and humidity to accelerate the drying process. Despite its undisputed effectiveness, kiln drying also remains a complex and energy-intensive operation that, if performed incorrectly can lead to significant losses. The interaction between multiple process variables, material heterogeneity and thermal inertia makes precise regulation difficult to achieve, particularly under ever-changing operating conditions. As a result, maintaining consistent drying behavior across production cycles continues to be a major industrial challenge that this work aims to study and address. In recent years, advances in sensor technology and data availability have enabled the collection of large volumes of data from industrial drying systems. At the same time, developments in *Artificial Intelligence* (AI) and *Machine Learning* (ML) have provided the industry with new tools that are more capable of extracting patterns and relationships from complex, nonlinear datasets. These developments create opportunities to better understand drying behavior, improve predictive modeling and support data-driven analysis of industrial processes. However, the direct application of ML techniques to industrial wood drying is constrained by practical data limitations. In particular, direct measurements of MC are often available at low temporal resolution, inconsistently logged, or subject to uncertainty. In contrast, temperature measurements are typically recorded continuously and at high-frequency, making them more suitable for high-resolution modeling. For this reason, temperature is commonly used in industrial practice as an indirect indicator of drying progress and moisture removal dynamics. This thesis investigates the application of data-driven ML techniques to the modeling and analysis of an industrial wood drying process using historical operational data. Temperature is adopted as a proxy variable for the drying state, enabling the development of predictive models under realistic industrial constraints. Two modeling approaches are explored and compared: a feedforward ANN and a LSTM network, representing static and temporal modeling paradigms, respectively. In addition to predictive modeling, this work introduces an offline control recommendation framework to analyze how different models respond to variations in controllable process variables. Rather than implementing a real-time or closed-loop control system, the framework is used to evaluate model sensitivity, stability, and control-oriented behavior in a reproducible and non-intrusive manner. Through a case study based on data from an industrial drying system operated by SONAE Arauco, this research aims to contribute both academically and practically by clarifying the strengths and limitations of static and temporal machine learning models in industrial drying contexts. The findings provide insight into the role of temporal modeling for advisory and decision-support applications and establish a foundation for future developments toward more advanced control-oriented systems.

1.1 Motivation

Environmental awareness and cost reduction are constant concerns in modern industrial systems, and the wood processing industry is no exception. Wood is a valuable and versatile natural resource, widely used in applications ranging from energy production to construction and furniture manufacturing. However, its industrial use requires extensive preparation due to its material properties, and MC plays a central role in this process. MC represents the amount of water contained within the wood material and has a direct impact on dimensional stability, mechanical strength, and long-term durability. Freshly cut wood typically presents high MC levels, which naturally decrease over time as the material exchanges moisture with the surrounding environment given the right circumstances, like a dryer environment surrounding the wood. This process continues until the wood reaches its *Equilibrium Moisture Content* (EMC), a condition determined by ambient temperature and relative humidity [Laboratory, 1999]. While EMC reflects a natural balance, it is usually insufficient for most industrial applications, making further moisture reduction necessary. To achieve target MC levels, the industry has developed several drying techniques, ranging from natural air drying to more controlled industrial processes. Natural drying is economically attractive but slow and highly dependent on environmental conditions. At the other extreme, chemical or assisted drying methods can achieve precise results but are associated with higher costs and environmental concerns. As a compromise, traditional kiln drying remains the most widely adopted industrial solution. Kiln drying heavily relies on controlled combinations of temperature, airflow and humidity to accelerate MC removal through evaporation. Although effective, this process is complex. Many critical variables interact in a nonlinear manner, and their effects on MC removal are not instantaneous. In industrial environments, some of these variables are not always measured with sufficient spatial or temporal resolution, while others are difficult to regulate precisely in real time. As a result, predicting final MC levels and determining optimal adjustments to controllable parameters often depend heavily on operator experience rather than systematic modeling. This variability in process control can lead to uneven drying, reduced product quality, increased energy consumption and, in some cases, the need for additional corrective treatments. As a consequence, improving MC management has become increasingly important, from both economic and environmental points of view. Recent advances in AI and ML offer promising tools to address these challenges. By leveraging historical operational data and sensor measurements, data-driven models can capture complex relationships between process variables that are difficult to describe analytically. Several studies have already demonstrated the potential of AI and ML techniques to improve MC-related predictions and support decision-making in drying processes [Rahimi and Avramidis, 2022]. The motivation for this work lies in exploring how such data-driven approaches can contribute to an improved understanding and management of industrial wood drying processes. By investigating predictive modeling and control-oriented analysis under realistic industrial constraints, this research aims to support the development of more consistent, efficient, and sustainable drying practices.

1.2 Contextualization

Wood Processing industries play a vital role in the global economy, providing essential materials for a wide range of applications, from energy production to furniture manufacturing. This sector is as economically significant as it is resource-intensive, urging the industry to balance productivity with environmental awareness. As wood processing relies heavily on the responsible use of natural resources, companies within this sector face rising pressure to adopt increasingly sustainable practices in order to reduce operational costs without compromising product quality [UNECE, 2023].

One of the most critical aspects of preparing and processing wood for industrial applications is managing its MC. Freshly cut wood naturally has a high MC, which significantly compromises its suitability for further processing. Consequently, MC reduction is an essential step in achieving the properties required to produce high-quality products across a range of applications such as *Medium-density Fibreboard* (MDF) production, which is essentially a type of board made from fine wood fibers that are pressed and stuck together mainly used for making furniture [Magalhães et al., 2021]. However, reducing wood MC in a cost-effective and environmentally friendly manner remains a significant challenge to the wood processing industry.

To address this challenge, a wide range of wood drying techniques has been developed, ranging from traditional natural drying - which is slow, typically does not achieve low or desirable MC levels, but is inexpensive - to more energy-demanding or less environmentally friendly methods, such as kiln drying and chemical drying. In particular, kiln drying, which involves the controlled use of hot air, temperature, and humidity levels to expel moisture from the wood, offers an efficient and scalable approach that is widely adopted in industrial settings [Keey et al., 2012]. While this technique enables companies to achieve lower MC levels in a shorter time, it also requires precise control over variables such as temperature, humidity, and airflow. This makes its management challenging, especially since these variables are often not directly measured or precisely regulated. Figure 3.1 presents a typical kiln layout.

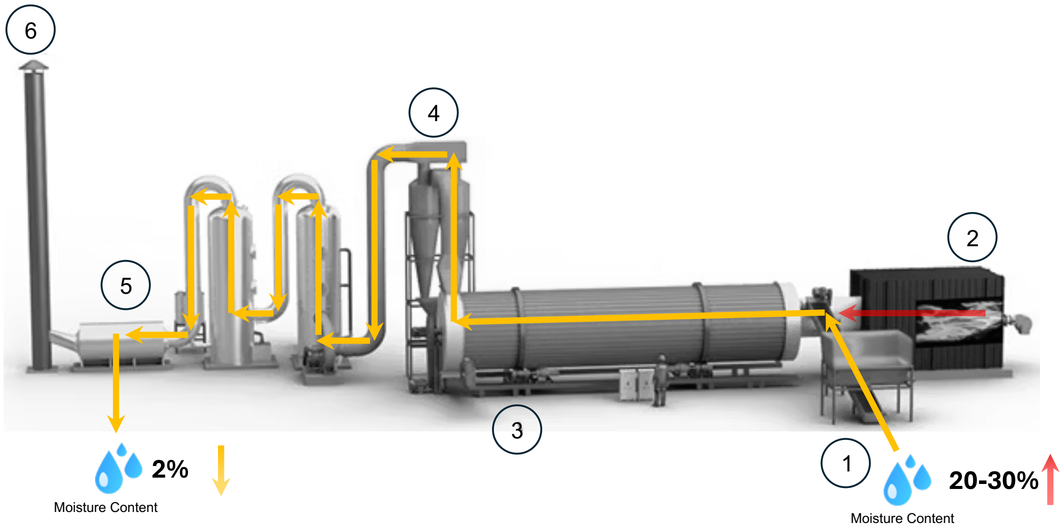

Figure 1.1: Representation of a general view of a traditional kiln system.

The basic and most common drying process using a traditional kiln can be understood as follows. The raw wood material is introduced into the system in 1. Then, a heat source 2 is applied to the raw material, supplying hot air for the whole system. 3 represents a rotary dryer, where the raw wood material is dried with the aid of the hot airflow coming from the heat source. 4 corresponds to a sorting system to sort unwanted materials from the final dried product. 5 represents the final phase of a traditional kiln system, and it varies depending on the specific industry in which the kiln is employed. Finally, 6 is a chimney that allows the hot airflow, now containing water vapor released from the drying wood, to be discharged into the atmosphere.

For Companies operating kilns, achieving precise MC levels at a final, as illustrated in 5 is an operational and economic challenge. Although kiln drying significantly accelerates the drying process, maintaining ideal condition within the system is quite expensive, resource-intensive, and prone to variability. This variability can lead to inconsistent wood moisture levels, affecting quality and durability of the end product. Consequently, over-dried or under-dried can result in material waste, costly reprocessing, or defective end products. Traditional methods struggle to account for dynamic variables such as temperature, humidity, and airflow. To address this issue, companies may resort to chemical treatments as a final resort to further reduce MC. However, these treatments raise environmental concerns and introduce higher operational costs, stepping away from the primary goal of efficiency and sustainability.

Given these challenges, there is a strong motivation to study improvements in kiln drying technology that could provide more precise control over the drying conditions. Such improvements would help reduce costs and align with both industry and global sustainability objectives and standards by lowering natural resource usage and waste, as well as reducing reliance on chemical treatments. In recent years, advances in sensor technologies, data analytics, and AI have shown significant potential to improve the efficiency and accuracy of these drying operations. These innovations could allow companies to better monitor and adjust key variables in real time, achieving more consistent MC levels and minimizing losses [da Silva Ferreira et al., 2024][Wang et al., 2023].

The kilns operating at SONAE Arauco provide a valuable case study for these potential improvements. By examining the company's specific challenges and the limitations of current kiln system technology, this research aims to identify practical solutions to optimize drying schedules and minimize operational costs. Furthermore, this study will explore how machine learning technologies could be implemented within the SONAE Arauco system to enhance performance and sustainability outcomes. A simplified representation of this system is shown in Figure 1.2.

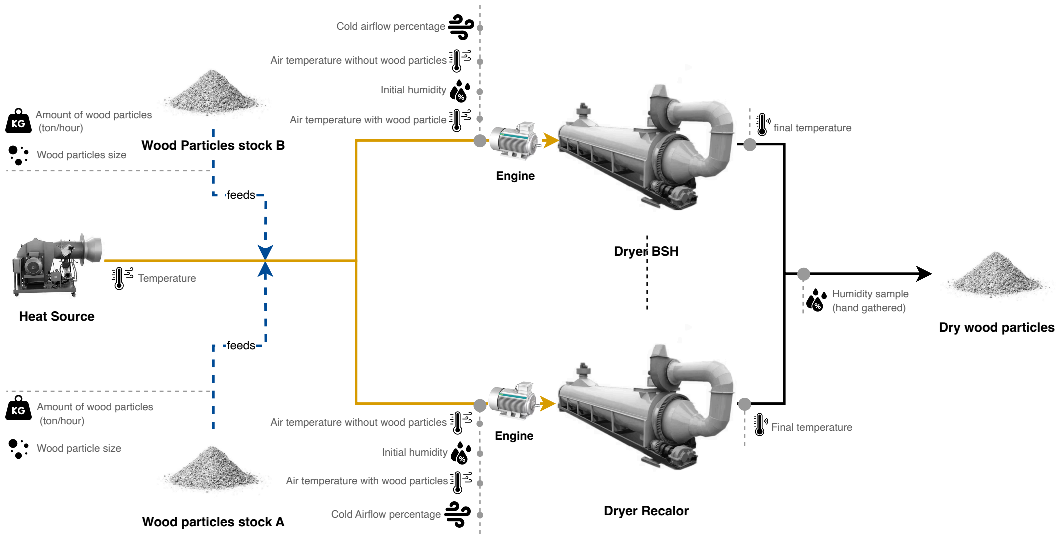

Figure 1.2: Representation of the SONAE Arauco simplified kiln drying system.

Considering a single dryer and assuming at this stage that all dryers are perfectly mirrored, the SONAE Arauco system is centered around an engine that controls the airflow of the entire kiln system, from the heat source, from which the temperature is known until its exit. It is clear that two different stockpiles feed the dryers, and here, two different measurements are provided: the wood particle size and the wood flow, or the amount of wood particles. Before the wood particles are fed into the system, a last temperature measurement is taken, as illustrated in Figure 1.2. Subsequently, the initial humidity is measured, along with a new temperature, which is now affected by the wood particles that enter the system. A controllable variable can also be controlled at this system stage: the cold airflow percentage, which determines how much ambient air enters the system. After the drying process, another temperature measurement and a periodic humidity samples are manually collected to verify that the target MC has been achieved.

This thesis therefore complements a critical research gap in the precise management of MC of wood drying techniques using traditional kiln processes to produce MDF-based products. This study seeks to contribute meaningfully to both industrial knowledge and SONAE Arauco's operational efficiency.

1.3 Problem Statement

Industrial wood drying involves a complex balance between operational efficiency, cost control, product quality, and environmental sustainability. Although kiln drying technology is well established and studied, widely adopted across the wood processing industry, significant limitations remain in achieving precise and consistent reduction of wood MC. These limitations arise from the strong coupling between multiple process variables, nonlinear system behavior and practical constraints in industrial monitoring and control.

In traditional kiln systems, key variables such as temperature, airflow, and humidity interact dynamically and influence MC removal in a delayed and nonlinear manner. While these variables are known to be critical, they are not always measured or regulated with sufficient resolution at the system level. As a result, operators often rely on empirical knowledge and manual adjustments to manage the drying process, which can lead to variability in final MC levels, increased energy consumption and reduced product consistency.

Recent advances in sensor technology, data analytics and AI offer promising opportunities to address these challenges. ML techniques enable data-driven modeling of complex industrial processes by learning relationships directly from historical operational data. However, despite growing research interest, practical and scalable solutions that effectively integrate predictive modeling with control-oriented analysis in real industrial drying environments remain limited. Challenges related to data quality, temporal dependencies, interpretability and industrial feasibility continue to restrict widespread adoption.

Within this context, the present research aims to investigate the potential of AI and ML techniques to support improved modeling and understanding of industrial wood drying behavior under realistic operational constraints. Rather than replacing existing control systems, this work focuses on predictive analysis and advisory control exploration based on historical data.

Specifically, this research seeks to address the following questions:

1. To what extent can AI and ML techniques improve the prediction of the drying state in industrial kilns, contributing to reduced variability and improved consistency when compared to traditional static approaches?

2. To what extent can data-driven models support the understanding of how controllable process variables influence drying outcomes, thereby enabling the generation of informative and interpretable control recommendations?

By addressing these questions, this thesis aims to evaluate whether sensor-integrated, data-driven modeling frameworks can enhance decision support in industrial kiln drying. The ultimate goal is to contribute to more consistent drying outcomes, improved operational efficiency, and reduced reliance on costly or environmentally aggressive treatments, while remaining aligned with practical industrial constraints.

1.4 Objectives

The main objective of this thesis is to investigate how AI and ML techniques can support improved management of the wood drying process in industrial kiln systems. Enhancing process understanding, predictive capability and decision support under realistic industrial constraints constitutes the main focus ot this work, with the ultimate aim of contributing to more consistent drying outcomes, improved operational efficiency and reduced environmental impact. To achieve this overall objective, the following specific objectives were defined:

- *Objective 1*: Develop and evaluate data-driven predictive models capable of estimating the drying state of an industrial wood drying system using historical operational data, adopting temperature as a proxy variable for MC due to its higher availability, reliability, and physical relationship with MC removal.

- *Objective 2*: Compare static and dynamic machine learning modeling approaches, specifically feedforward ANN and LSTM networks, in order to assess their ability to generalize under changing industrial operating conditions and to capture the temporal dynamics of the drying process.

- *Objective 3*: Design and assess an offline control recommendation framework that explores how predictive models respond to variations in controllable process variables, enabling the analysis of model sensitivity, stability, and interpretability for advisory and decision-support purposes, without deploying a real-time or closed-loop control system.

1.5 Expected Results

The expected results of this thesis focus on improving understanding, predictive capability, and decision support for industrial wood drying processes through the application of data-driven machine learning techniques. Rather than proposing a fully deployed industrial control solution, this work aims to provide validated insights and frameworks that can support future system integration.

- *R1*: The development and evaluation of ML models capable of accurately estimating the drying state of an industrial wood drying system using historical operational data. By adopting temperature as a proxy variable for MC, the models are expected to demonstrate the feasibility of data-driven prediction under realistic industrial constraints, while highlighting the strengths and limitations of different modeling approaches.

- *R2*: A detailed analysis of the influence of controllable and non-controllable process variables on the predicted drying state. Through feature relevance and sensitivity analysis, this result is expected to provide insights into which variables most strongly affect model predictions, supporting improved process understanding and informed operational decision-making.

- *R3*: The implementation and offline evaluation of a control recommendation framework based on machine learning predictions. This framework is expected to enable the assessment of how different predictive models behave and, consequently, respond to variations in controllable process variables, allowing comparison of control effort, stability. and consistency without deploying a real-time or closed-loop control system.

1.6 Work Plan

The work developed in this thesis followed a structured and sequential plan designed to progressively address the research objectives while accounting for practical constraints associated with industrial data and offline analysis. The main stages of the work are summarized below.

The first stage consisted of an extensive bibliographic review focused on wood drying processes, moisture content behavior, and the application of artificial intelligence and machine learning techniques in drying and industrial process modeling. This review established the theoretical foundation of the work and supported the selection of suitable modeling approaches. In the second stage, historical operational data from an industrial wood drying system operated by SONAE Arauco was collected and analyzed. This phase involved understanding the structure of the dataset, identifying available process variables, and assessing data quality and limitations, particularly regarding moisture content measurements. The third stage focused on data preprocessing and preparation. This included data cleaning, handling of missing and anomalous values, removal of non-informative variables, and data standardization with special attention given to preserving temporal consistency and ensuring suitability for machine learning modeling. In the fourth stage, feature relevance and selection analyses were performed using permutation-based methods. These analyses aimed to identify the most influential process variables for predictive modeling, considering both static and temporal modeling formulations. The fifth stage involved the implementation of predictive machine learning models. Two approaches were developed and evaluated: a ANN and a LSTM network. Both models were trained and tested using chronologically split datasets to assess predictive performance and generalization under realistic industrial conditions. The sixth stage consisted of the development of an offline control recommendation framework. Using the trained predictive models, this framework explored how variations in controllable process variables affect the predicted drying state, enabling comparative analysis of control sensitivity, effort, and stability between modeling approaches. Finally, the last stage involved the evaluation and interpretation of results. Predictive performance metrics, feature relevance outcomes, and control recommendation behavior were analyzed and discussed. The findings were consolidated to assess the suitability of static versus temporal modeling approaches and to identify directions for future work and potential industrial integration.

1.7 Thesis Structure

This thesis is structured into five main chapters:

1. Introduction - This chapter introduces the context and motivation for the research. It presents the problem statement, defines the research objectives, outlines the expected results, and describes the work plan adopted throughout the thesis.

2. Literature Review - This chapter reviews existing scientific literature related to wood drying processes and MC management. It also introduces key concepts and examines previous applications of AI and ML techniques in drying processes and related industrial domains.

3. Methodology - This chapter details the methodological approach adopted in the thesis. It describes the bibliographic research strategy, the overall implementation framework, and the assumptions under which the study was conducted. The chapter also explains the modeling approach, evaluation strategy, and design considerations guiding the work.

4. Data Preparation and Model Implementation - This chapter presents the acquisition and preprocessing of industrial data, feature selection procedures, and the implementation of ML models. It details the development of both static and temporal predictive models, including ANN and LSTM networks.

5. Results and Discussion - This chapter presents and analyzes the experimental results obtained from predictive modeling and control recommendation analyses. The performance of different modeling approaches is evaluated and discussed in the context of industrial applicability, generalization capability, and control-oriented behavior.

6. Conclusion and Future Work - The final chapter summarizes the main findings and contributions of the thesis. It discusses the limitations of the study and proposes directions for future research, including potential extensions toward real-time implementation and enhanced industrial integration.

Chapter 2

Literature Review

The wood processing system faces complex challenges in its operations every day. One of the main challenges is achieving precise MC control, especially when using traditional kiln-drying techniques. This chapter explores the theoretical framework and research developments related to kiln drying processes, MC reduction techniques, and the importance of emerging technologies such as AI, ML, and advanced sensor integration.

Earlier studies in this field, such as those by [Laboratory, 1999] and later by [Keey et al., 2012], provided a foundational understanding of kiln drying processes and the effects of various drying parameters on MC levels. These studies demonstrated the importance of managing airflow, temperature, and humidity levels while highlighting the limitations of precision control, which can lead to variations in MC that increase operational costs and affect product quality.

The influence of MC on wood properties extends beyond drying. More recently, [Magalhães et al., 2021] explored how MC affects the internal bond strength and wood swelling behavior of MDF. This study showed that higher MC levels in MDF panels lead to reduced internal bonding strength and increased swelling, compromising the structural integrity of wood-based products. These findings further emphasize the importance of precise MC management for maintaining product quality, particularly in industries where wood is exposed to varying environmental conditions.

This review examines both traditional and emerging techniques for controlling MC in wood processing, discussing the challenges and recent advancements in sensor integration, AI and ML technologies that can improve the efficiency of traditional kiln systems. Additionally, it contextualizes these advancements with the SONAE Arauco case study, where new approaches will be investigated and tested to improve MC control in an industrial setting.

2.1 Definitions

In the context of this research, there are concepts and technologies that play an important role in understanding the methodologies employed to improve the precision of wood drying processes. One of the fundamental aspects is MC, which represents the percentage of water contained inside the wood. The ability to accurately determine and control MC levels is critical ensuring product quality, reducing waste, and optimizing energy efficiency. MC is calculated using a standard formula, that determines the difference between the wet and dry weight of the wood, providing information regarding the effectiveness of the chosen drying techniques and schedules.

When it comes to wood drying techniques, kiln drying is widely used in industrial applications, offering a controlled environment in which temperature, airflow, wood flow, and humidity can be adjusted to accelerate the drying process. Unlike other drying methods, kiln drying allows for greater efficiency and uniformity; however, achieving consistent results requires precise management of the controllable variables involved such as those mentioned above. One of the critical benchmarks in this process is reaching the EMC, a point at which wood neither gains nor loses moisture based solely on ambient conditions and variables. Stabilizing MC at this stage ensures that further drying can be conducted effectively to reach optimal and target levels for industrial use.

An important application of wood drying technology, and the focus of this thesis, is in the production of MDF, an engineered product composed of wood fibers bonded with resin under high temperature and pressure. MDF is highly valued and known for its uniformity and stability, being the main material used in furniture manufacturing, yet it remains highly sensitive to moisture. This makes precise control of MC an essential factor in ensuring the durability and integrity of MDF-based products.

To address the known challenges in optimizing kiln wood drying, this study integrates AI techniques, which allow the computer to perform tasks that typically require human intelligence, such as decision-making, pattern recognition, and real-time variable adjustments. In this study, AI is used to enhance kiln operations by processing large amounts of data and making predictive adjustments to optimize drying conditions. A smaller subfield of AI, known as ML, plays a critical role in this optimization, allowing systems to learn from data and improve their predictive capabilities over time. In kiln drying, ML models analyze variables in available kiln data - such as temperature, MC, wood flow, and airflow - to predict final MC levels and make dynamic adjustments to maintain consistency.

One of the advanced technologies supporting to ML in this field is *Computer Vision* (CV), a field that enables machines to interpret and analyze visual data. By using *Red, Green, Blue* (RGB) imaging and *Near-Infrared Spectroscopy* (NIR), CV techniques provide real-time monitoring capabilities of the drying processes, allowing AI driven systems to detect anomalies and optimize conditions accordingly. A more advanced approach, Near-Infrared Hyperspectral Imaging (NIR-HSI), further improves this capability by capturing detailed spatial and spectral data, offering deeper information about moisture distribution inside the wood.

Evaluating the performance of AI-driven drying systems requires a robust set of Key Performance Indicators (KPI)s, which evaluate efficiency and effectiveness. The most commonly used statistical metrics are Mean Squared Error (MSE) and Root Mean Sqared Error (RMSE), both of which measure the accuracy of MC predictions by calculating the deviation between predicted and actual values. Another important metric is the Mean Absolute Error (MAE), which calculates the average absolute difference between predicted and actual MC levels, giving a way to measure the prediction accuracy.

By integrating these concepts, this research aims to revolutionize kiln drying operations through AI driven optimization, ensuring greater precision, energy efficiency, and sustainability in wood drying processes. The synergy between AI, ML, and advanced monitoring technologies provides a promising path in achieving high quality drying outcomes, consistently reaching target MC levels in a constant manner while minimizing environmental impact and operational costs.

2.2 Related Work

Traditional kiln drying methods have been widely researched, with foundational studies such as those by [Keey et al., 2012] and [Laboratory, 1999] that explored how the control of drying variables - temperature, humidity, and airflow - affects variability in MC levels of wood. Despite its effectiveness, kiln drying faces some challenges in maintaining precise control over these variables in an industrial settings, leading to significant variability in MC that affects final product quality and increases operational costs due to reprocessing or additional treatments, often involving chemicals. Recent advancements have explored solutions using machine learning models to predict drying rates and improve process control [Rahimi and Avramidis, 2022, Wu and Avramidis, 2006].

The impact of MC on wood-based products is further supported by [Magalhães et al., 2021], who investigated and studied how MC influences the internal bond strength and swelling behavior of MDF. Their study showed that increased levels of MC compromise bond strength and increase swelling. They demonstrated that bond strength can be regained after drying. However, the process shows a slight oscillatory behavior, reinforcing the need for careful MC management at all stages of production.

In recent years, the integration of sensor technology and AI into drying processes has offered promising solutions to the limitations of traditional kiln drying. Real-time data collection from sensors can provide powerful insights and enables precise monitoring of kiln variables [Rahimi and Avramidis, 2022]. At the same time, AI-driven control systems can perform predictive adjustments to optimize the whole process and improve drying conditions. Similar approaches have shown success in other industries. For example, AI and sensor integrations have been deeply reviewed and analyzed to show how much can machine learning models improve energy efficiency, thus turning drying processes more sustainable and cost-effective [Đaković et al., 2024]. Also, for instance, in the food processing industry, studies have demonstrated that moisture consistency and stability can be significantly improved using AI driven systems while, at the same time, reducing energy use and enhancing product quality [Zhang et al., 2024]. Emerging AI systems have also shown potential for improving real-time monitoring and process optimization in sectors such as agriculture and energy. In the context of drying processes, as previously mentioned, AI enabled optical sensors have been applied to smart and precision food drying, for example, indicating that these technologies can be extended to other industrial applications [da Silva Ferreira et al., 2024, Yadav et al., 2021]. Within the wood processing industry, recent studies have applied ML models to predict MC in kiln-dried timbers, demonstrating the possibility and potential benefits of AI integration in wood drying processes [Rahimi and Avramidis, 2022] [Rahimi et al., 2024].

Building on these studies, the present research will apply AI and sensor technology in a case study at SONAE Arauco, aiming to assess the benefits of these innovations to kiln conditions optimization. The focus is to achieve lower costs, consistent MC and improve environmental impacts since there is a lack on literature regarding practical AI integration in industrial kiln systems and case studies regarding operational cost reduction and environmental impact mitigation, as well as, AI-enabled sensing technologies in wood drying processes to enhance real-time monitoring and control [da Silva Ferreira et al., 2024, Wang et al., 2023].

The study [da Silva Ferreira et al., 2024] explores the use of AI-enabled optical sensing technologies for precise monitoring and variable control in food drying processes. The authors discussed three optical sensing systems - RGB imaging with CV, NIR, and NIR-HSI. Each one of these techniques was evaluated in terms of mechanisms, applications, advantages, and limitations. The study demonstrates that integrating these techniques with AI allows more accurate real-time monitoring, thereby improving drying efficiency and final product quality. The major challenges with these solutions are the high costs and significant computational requirements. Although the article's primary focus is on the food industry, more specifically on food drying processes, the discussed techniques can be adapted for wood drying processes. The implementation of AI powered optical sensors could allow real-time monitoring of MC during the wood drying process, providing precise insights to make adjustments to drying conditions in order to achieve the desired MC levels. This capability can result in more uniform drying and improved final product quality. However, it is crucial to consider the differences in the physical and chemical properties between of food and wood when trying to replicate and adapt these techniques.

The study by [Wang et al., 2023], it presents an MC prediction approach for Sugi wood drying using Deep Long Short-Term Memory Autoencoders (LSTM AE) with perturbed error minimization. Autoencoders are neural networks designed to learn efficient data representations in an unsupervised manner and consist of two main components: an encoder and a decoder. The encoder is responsible for taking the input data and encoding it into a format more suitable for transmission or storage, while the decoder is responsible for the opposite: taking the encoded data and decoding it into the original format. The primary objective of an autoencoder is to minimize the reconstruction error, ensuring that the output closely matches the original input [Bank et al., 2023]. LSTM AE are a specific class of autoencoders in which both the encoder and decoder are implemented using LSTM networks [Malhotra et al., 2015]. LSTMs are a type of Recurrent Neural Network (RNN) particularly suited to learning temporal dependencies in sequential data [Hochreiter and Schmidhuber, 1997]. LSTM AEs are well suited for time-series data or sequences where patterns evolve over time and are widely used in for tasks such as anomaly detections or trend prediction. The proposed model referred to as a Deep LSTM Autoencoder Minimizing Perturbed Error (LSTM), is designed to extract hidden representations from historical MC data, aiming to improve the prediction accuracy. The results indicate that this model outperforms traditional models in terms of accuracy in predicting MC during Sugi wood drying processes. This article is particularly relevant as it provides valuable insights into the prediction of MC during wood drying using advanced machine learning techniques. However, it is important to highlight the differences in MDF and Sugi wood. MDF is an engineered material in contrast to Sugi wood which is a natural material which also means that MDF lacks natural cellular structure causing its molecular behavior to be fundamentally different making the MDF drying process a unique challenge.

Another relevant study, [Chai and Li, 2023], focuses on the use of ANNs for wood drying process predictions. ANNs are versatile and are capable of modeling complex non-linear relationships between input variables, such as temperature and humidity, and output variables, such as MC [Haykin, 1998]. ANNs are also easier to implement when compared to LSTMs; however they lack robustness in capturing situations when long-terms dependencies, such as how past and evolving process conditions influence future MC levels. The primary objective of this study is to model the relationships between key drying process parameters - such as temperature, airflow, and humidity - and how it affects the resulting levels of MC. The study aims to improve the precision of MC predictions, optimize wood drying schedules, and enhance the overall efficiency of the wood drying process. It provides a solid foundation for the application of ANNs to the wood processing industry, demonstrating their ability to accurately capture non-linear relationships and making then a suitable modeling approach for the challenges associated with MDF drying.

Environmental awareness is also a major concern of the SONAE Arauco operational system. The study by [Đaković et al., 2023] addresses energy efficiency in various drying applications using ML highlighting how these techniques can enhance energy efficiency while improving product quality. The primary objective of the study is to explore the intersection between energy-saving methods and advanced ML techniques across different drying systems. The study demonstrates that ML models can achieve significant energy savings by optimizing key drying parameters such as temperature, airflow, and humidity, thereby minimizing energy wastage. For example, the application of LSTM networks to predict and adjust drying schedules proved to reduce energy consumption in industrial dryers. However, the study also highlights that such approaches are not trivial to implement, as they require substantial computational resources and large, high-quality datasets for effective and efficient model training. Another challenge identified, which is also relevant to this study is the integration of ML-based solutions with existing industrial systems and sensor technologies.

Chapter 3

Methodology

This chapter presents the methodology adopted to address the objectives of this thesis. It describes the overall research approach, the modeling framework employed, and the evaluation procedures defined to study the application of artificial intelligence and machine learning techniques to industrial wood drying processes. The methodological choices described in this chapter were guided by both theoretical considerations and practical constraints arising from the use of real industrial data.

The chapter begins with an overview of the bibliographic research conducted to establish a solid theoretical foundation and to identify relevant studies related to wood drying processes and machine learning applications in this domain. This review supports the selection of appropriate modeling approaches and situates the present work within the existing scientific literature.

The implementation process is then described at a conceptual level, outlining the sequence of methodological steps adopted, from data acquisition and preprocessing to the selection and configuration of predictive modeling techniques. Particular emphasis is placed on the adoption of ANN and LSTM networks, motivated by their suitability for modeling nonlinear and time-dependent industrial processes.

The methodology focuses on offline analysis under controlled assumptions, defining how models are trained, tested, and evaluated without real-time system interaction. This approach enables a systematic and reproducible assessment of predictive behavior while respecting industrial constraints and limitation related to data availability.

Finally, this chapter presents the evaluation methods and performance metrics used to compare different modeling approaches. These criteria were selected to ensure a transparent and consistent assessment framework, forming the basis for the analysis and discussion of results presented in the following chapter.

3.1 Bibliographic Research

A bibliographic review was conducted as a foundational component of this research, with the objective of establishing a solid theoretical and scientific basis for the application of AI and ML techniques to industrial wood drying processes. This review aimed to identify relevant methodologies, state-of-the-art approaches, and documented challenges related to the modeling, prediction, and optimization of drying systems.

Through a systematic analysis of the existing literature, this research sought to gain insights into current trends, limitations, and research gaps associated with kiln wood drying, moisture-related process behavior, and data-driven control strategies. Particular attention was given to studies addressing nonlinear modeling, time-dependent processes, and the application of AI-based techniques in industrial drying contexts.

The findings of this bibliographic review informed the methodological decisions adopted in this thesis, particularly with regard to the selection of predictive modeling approaches and the definition of evaluation criteria. By grounding the proposed methodology in established scientific knowledge, the review ensured alignment with current research directions in both wood science and machine learning, while also identifying opportunities for further investigation.

3.1.1 Bibliographic Research Methodology

The bibliographic research conducted for this thesis followed a structured and systematic approach to identify, analyze, and synthesize relevant scientific literature related to the application of AI and ML techniques in drying processes, with particular emphasis on the wood processing industry.

The review process began with an extensive search across major academic databases, like Google Scholar, IEEE Xplore, ScienceDirect, Scopus, and SpringerLink. These platforms were selected to ensure comprehensive coverage of peer-reviewed literature spanning both engineering and applied machine learning domains.

A set of targeted keywords was defined to capture the relevant research scope. Examples of the search terms used include ”neural networks wood drying”, ”machine learning moisture content prediction”, ”artificial intelligence kiln drying optimization”, ”energy efficiency in drying processes” and ”AI-driven monitoring and control systems". Boolean operators such as AND and OR were employed to refine search queries and improve result relevance.

To ensure scientific rigor and relevance, priority was given to peer-reviewed journal articles, conference proceedings and authoritative books, predominantly published within the last decade. In addition, classical and foundational works were included where necessary to support theoretical concepts, particularly in the areas of neural networks, time-series modeling, and drying theory.

Special attention was given to studies investigating ANN, LSTM networks, and autoencoder-based approaches in the context of drying processes and dynamic system modeling. Once relevant literature was identified, each study was carefully reviewed to extract methodological approaches, key findings and, reported limitations. This analysis enabled the identification of common challenges, emerging trends, and gaps within the existing body of research.

Finally, the reviewed literature was organized into thematic categories to facilitate synthesis and interpretation. This thematic organization supported the contextualization of the present work within the broader research landscape and provided guidance for the methodological decisions described in the subsequent sections.

3.1.2 Topics and Areas of Research

The topics explored were selected to align with the scope of this thesis, aiming to provide the most relevant knowledge to support and enrich this study. The areas of focus included:

- AI and ML techniques: Fundamental concepts and architectures related to ANN, LSTM and autoencoders, with an emphasis on their suitability for regression tasks and dynamic process analysis.

- Wood Drying Processes: An overview of traditional and modern wood drying methods, including kiln drying and natural drying, with a discussion of their advantages, limitations, and industrial relevance. Particular attention was given to the challenges associated with achieving consistent moisture content and the implications for product quality.

- Energy Efficiency: Studies addressing energy consumption in drying processes and the potential of ML-based solutions to improve efficiency and reduce operational costs.

- Sensor Integration and Data Acquisition: Research on industrial sensor systems, real-time data collection, and the role of sensor data in supporting data-driven modeling and predictive control frameworks.

3.1.3 Search Strategy and Quality Assessment

The literature search strategy was defined at two complementary levels: the search methodology and the quality assessment of the retrieved studies.

At the search methodology level, database queries were refined using Boolean operators and iterative keyword adjustments, as mentioned earlier, in order to improve precision and relevance. Citation chaining was also employed to identify influential and frequently cited studies that contributed significantly to the development of the field of study.

Regarding quality assessment, each selected document was evaluated for scientific credibility, methodological rigor, and relevance to the objectives of this thesis. Indexing platforms, such as Scopus, were used as supporting tools during this evaluation.

Finally, the studies were critically examined to determine their applicability to industrial wood drying processes and data-driven modeling approaches.

3.2 Implementation Process

The implementation process adopted in this thesis follows a structured, data-driven methodology aimed at modeling and analyzing industrial wood drying processes using machine learning techniques. The general workflow of the proposed framework is illustrated in Figure 3.1, which provides a conceptual overview of the main methodological stages, ranging from data acquisition and preprocessing to predictive modeling and offline control analysis.

It is important to emphasize that the proposed framework was developed and evaluated exclusively in an offline context using historical industrial data. The objective of this implementation was not the deployment of a real-time control system, but rather the systematic assessment of predictive modeling approaches and the exploration of their potential applicability to control-oriented decision support.

This workflow ensures methodological clarity and reproducibility, while allowing every stage of the process to be analyzed independently.

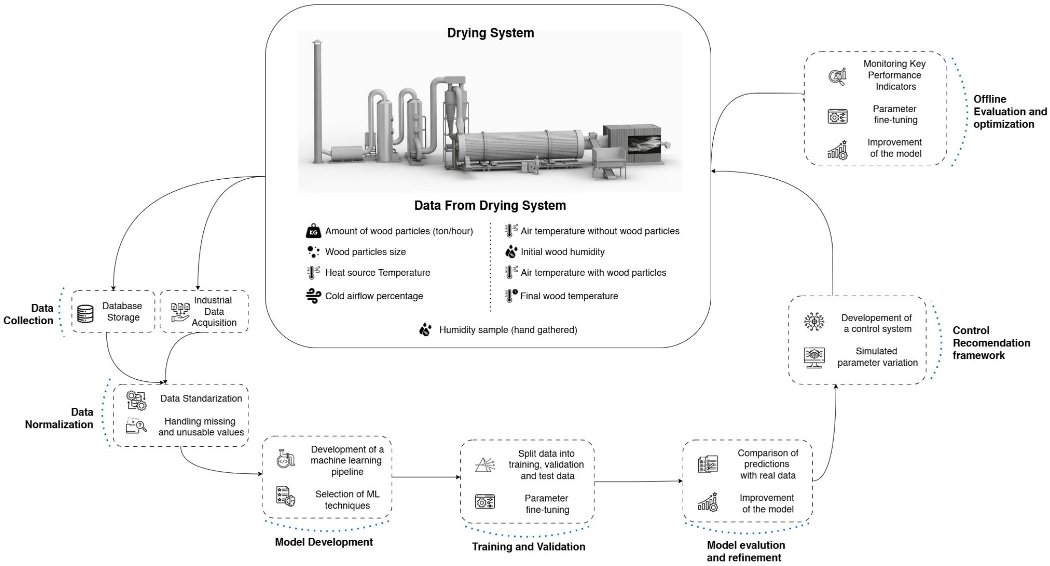

Figure 3.1: Conceptual workflow of the proposed data-driven framework for predictive modeling and control recommendation of an industrial wood drying kiln.

3.2.1 Data Acquisition and Industrial Context

The data used in this work originates from an industrial wood drying system operated by SONAE Arauco. The system is equipped with a set of sensors designed to monitor key operational conditions during normal industrial operation. These measurements provide the foundation for the development and evaluation of data-driven modeling approaches within an industrial context.

The present study is based exclusively on historical process data and does not include real-time data acquisition or online system interaction. Although the drying system is capable of supporting real-time monitoring, the analysis was intentionally restricted to offline data to ensure reproducibility and to avoid interference with production processes.

During data collection, different types of process variables were recorded at varying temporal resolutions. In particular, direct measurements related to wood MC were available only at relatively large time intervals and were not consistently logged, introducing uncertainty and potential measurement bias. In contrast, temperature-related variables were continuously measured and recorded with higher temporal resolution and greater reliability.

Given these constraints, temperature was selected as the primary modeled variable in this study, serving as an indirect indicator of the drying state. This methodological choice enables the use of dense and reliable time-series data while remaining aligned with industrial practice, where temperature is a central control and monitoring variable. The rationale for adopting temperature as a proxy for moisture-related behavior is further discussed in Section 3.2.2.

3.2.2 Temperature as a Proxy for Moisture Content

The primary objective of this thesis is the prediction and control of wood MC during the drying process. However, in the industrial context considered in this work, direct moisture measurements present significant practical limitations. MC data were collected at relatively large time intervals (approximately once per hour) and, in some cases, were not consistently recorded or logged. These limitations introduce errors related to data sparsity, temporal misalignment and potential measurement errors.

In contrast, temperature measurements within the drying system were continuously monitored through integrated sensors and recorded at a higher temporal resolution. According to operational expertise provided by plant operators, the temporal evolution of temperature inside the dryer is strongly and consistently related to moisture removal, as heat transfer mechanisms directly impact evaporation rates and drying efficiency. This relationship is well established in the literature on kiln drying processes, where temperature profiles are commonly used to characterize drying stages and to indirectly infer moisture evolution [Keey et al., 2012, Wu and Avramidis, 2006].

Given these considerations, temperature was adopted in this study as a proxy variable for the drying state. This methodological choice enables the use of dense and reliable time-series data while avoiding the introduction of uncertainty errors associated with sparse and inconsistent MC measurements. Furthermore, this approach remains aligned with industrial practice, where drying processes are primarily monitored and controlled through temperature-related variables rather than direct MC measurements.

3.2.3 Data Preprocessing

Before proceeding to predictive model development, the collected data underwent a preprocessing stage to ensure consistency, numerical stability, and suitability for machine learning algorithms. This stage was designed to address common issues associated with industrial datasets, such as missing values, sensor noise, and redundant information.

The preprocessing procedure included the identification and removal of invalid or unusable records, as well as the handling of missing values to prevent issues during model training. Variables that were constant over time or deemed irrelevant for the modeling objectives were eliminated in order to reduce dimensionality and avoid unnecessary computational complexity.

Additionally, data cleaning procedures were applied to reduce the impact of anomalous or incorrect sensor readings. These steps aimed to minimize noise and improve data reliability while preserving the underlying process dynamics. The resulting dataset provided a consistent and structured basis for subsequent modeling and evaluation stages.

3.2.4 Feature Relevance Analysis and Variable Selection

Given the large volume and heterogeneity of the available industrial drying dataset, a feature relevance analysis was incorporated into the methodology to support informed feature selection for predictive modeling. This step aimed to reduce the influence of unimportant or redundant variables, minimize noise and ensure that the selected input features were both statistically relevant and physically meaningful within the context of the drying process.

Traditional feature selection techniques based on linear correlation are often insufficient for complex industrial systems, where strong nonlinearities and interactions between variables are present [Zhang et al., 2025]. For this reason, a permutation-based feature importance approach was adopted [scikit-learn developers, 2024], as it is well suited for nonlinear machine learning models and does not rely on model-specific internal parameters [Breiman, 2001].

In this approach, a predictive model is first trained using the original dataset. Subsequently, the values of each input variable are randomly permuted while all remaining variables are left unchanged. The resulting change in model performance is then used as an estimate of the relative importance of the permuted variable. Variables whose permutation leads to a greater degradation in predictive performance are interpreted as having a stronger influence on the model output.

To ensure methodological consistency across modeling approaches, feature relevance analysis was conducted separately for each predictive model considered in this work, namely a feedforward ANN and a LSTM network. This allowed feature selection to account for both static and temporal modeling perspectives without imposing assumptions regarding linearity or time independence.

3.2.5 Predictive Model Development

Following the data preprocessing and feature relevance analysis, predictive models were defined to estimate the evolution of the drying process using historical industrial data. The objective of this stage was to establish a modeling framework capable of representing kiln behavior under different assumptions regarding process dynamics, rather than identifying optimal predictive configurations.

Two predictive modeling approaches were considered in this work: a feedforward ANN and a LSTM network. These models were selected based on insights obtained from the bibliographic review, which highlighted the widespread use of ANN architectures for nonlinear regression in drying processes and the increasing relevance of LSTM networks for modeling temporal dependencies in industrial time-series data.

The predictive task was formulated as a regression problem, in which the models estimate the output temperature of the drying system at a given time step. As discussed previously, temperature was adopted as a proxy variable for the drying state due to its strong physical relationship with MC evolution, as well as the higher reliability and temporal resolution of temperature measurements in the available industrial data.

Artificial Neural Network (ANN) Model

The ANN model was defined as a multilayer perceptron designed to capture static nonlinear relationships between process variables and the target temperature. At each time step, the model receives a vector of instantaneous process measurements as input and produces a single temperature estimation as output.

The network architecture consists of multiple fully connected layers, organized to provide sufficient representational capacity for nonlinear mapping. Input features are standardized before training to ensure numerical stability and consistent scaling. The ANN formulation does not explicitly incorporate temporal information, operating under the assumption that the instantaneous process state contains sufficient information to estimate the target variable.

Within the scope of this methodology, the ANN serves as a baseline static modeling approach, enabling subsequent comparison with temporally aware models without imposing assumptions regarding process memory or delayed effects.

Long Short-Term Memory (LSTM) Model

To explicitly account for temporal dynamics related to the drying process, an LSTM-based model was defined. LSTM networks are a class of recurrent neural networks designed to capture long-term dependencies in sequential data through gated memory mechanisms, making them particularly suitable for modeling industrial time series behavior.

In the adopted formulation, the input data are structured into sequences using a sliding window approach. Each input sample consists of a sequence of consecutive time steps of process variables, and the model is trained to estimate the target temperature at the following time step. This sequence-to-one configuration allows the model to learn how historical process states influence future temperature evolution.

The LSTM architecture comprises stacked recurrent layers, followed by a dense output layer. Input features are standardized prior to sequence construction to ensure consistent scaling across time steps. By incorporating historical information, the LSTM model is designed to represent delayed effects, process inertia, and dynamic interactions that cannot be captured by static modeling approaches.

Modeling Scope and Design Considerations

The ANN and LSTM models were developed independently using representative architectures, without performing exhaustive hyperparameter optimization or architecture search. This design choice reflects the exploratory nature of the study and the practical constraints associated with computational cost and industrial data availability.

The objective of this modeling stage was not to identify optimal predictive configurations, but rather to investigate how different modeling approaches respond to the same industrial dataset when evaluated under consistent methodological conditions. Both models were trained using the same preprocessed feature sets and target definitions to ensure comparability and methodological consistency.

Relation to Control Recommendation Framework

The predictive models defined in this section form the basis for the offline control recommendation framework described in subsequent sections. For the ANN model, control analysis is performed by evaluating the sensitivity of the predicted temperature to variations in controllable process variables at a single time step. In contrast, the LSTM model evaluates the impact of parameter variations within the context of recent process history, capturing the influence of temporal dependencies.

By employing both static and temporal predictive models, this methodological framework enables a comparative analysis of model behavior and responsiveness in control-oriented scenarios, without introducing assumptions regarding real-time deployment or closed-loop system operation.

3.2.6 Training, Validation, and Testing Strategy

To ensure a realistic and methodologically sound assessment of predictive performance, a structured training, validation, and testing strategy was adopted. Given the time-dependent nature of the drying process, particular emphasis was placed on preserving the temporal ordering of the data throughout the modeling pipeline.

During initial exploratory phases, some predictive models were trained and evaluated using the full available dataset. These runs were intentionally performed to support preliminary validation tasks, feature relevance analysis, and sensitivity exploration. The resulting performance metrics correspond to in-sample error estimates and were not interpreted as indicators of model generalization capability.

For all comparative evaluations showed in this thesis, the dataset was first sorted in chronological order based on timestamp information and subsequently divided into training and testing subsets using a fixed percentage portion. The earliest split of the time series (approximately 80%) was used for testing while the remaining portion (approximately 20%) was used for testing. This chronological split reflects the practical constraint that future process states cannot be used to predict past behavior and provides a realistic evaluation scenario for potential industrial deployment.

Input features were standardized using parameters estimated exclusively from the training data, and the same transformations were subsequently applied to the testing data in order to prevent information leakage. For the LSTM model, the use of sliding window sequences resulted in a reduced effective number of samples compared to the ANN model. This difference arises naturally from the temporal modeling formulation and was accounted for during evaluation.

3.2.7 Control Recommendation Framework

Building on the predictive models defined in the previous sections, an offline control recommendation framework was developed to analyze how model predictions respond to variations in controllable process variables. The objective of this framework was not to replace existing industrial control systems, but rather to explore model sensitivity and generate interpretable recommendations that could support the operator decision-making process.

A subset of variables was identified as controllable based on operational knowledge, including actuator-related parameters associated with wood flow and air flow. During the control analysis, these variables were systematically perturbed within predefined bounds, while all non-controllable variables were held constant. The predictive models were then used to estimate the resulting variation in the predicted output temperature.

For the ANN based framework, control recommendations were derived using instantaneous snapshots of the process state. Each controllable variable was independently adjusted over a limited range, and the predicted temperature response was evaluated relative to a predefined target value. This formulation reflects the static nature of the ANN model and provides insight into pointwise sensitivity between input variables and the predicted output.

In contrast, the LSTM-based framework incorporated recent process history through temporal input sequences. Control perturbations were applied to the most recent time step within each sequence, and predictions were evaluated in the context of preceding system dynamics. This approach enables the LSTM model to account for delayed effects, process inertia, and temporal dependencies when generating control-oriented recommendations.

The adopted control strategy is intentionally simple and transparent, relying on bounded perturbation analysis rather than optimization-based or model predictive control techniques. This design choice prioritizes interpretability and robustness, facilitates comparison between modeling approaches, and avoids assumptions regarding actuator dynamics, closed-loop stability, or real time system integration.

3.2.8 Offline Evaluation and Refinement

The control recommendation framework was evaluated through offline analysis using historical industrial data. For each target variable, a fixed number of representative time points were selected from the test period, and control recommendations were generated independently for each instance. This approach enabled a systematic and repeatable assessment without interfering with industrial operation.

Evaluation metrics were defined to characterize both predictive adjustment and control effort. Predictive adjustment was quantified as the change in absolute prediction error relative to a predefined target temperature. A recommendation was considered successful if the predicted temperature fell within a specified tolerance range around the target value. Control effort was assessed based on the magnitude of parameter changes and the number of controllable variables modified in each recommendation.

This evaluation framework was applied consistently to both ANN-based and LSTM-based control formulations, enabling a comparative analysis under identical conditions. The purpose of this analysis was to examine model behavior in control-oriented scenarios rather than to emulate a closed-loop control system.

It is important to emphasize that no real-time actuation or industrial deployment was performed. Actuator constraints, system delays, and safety mechanisms were not explicitly modeled, and the generated recommendations should be interpreted strictly as advisory outputs. The results obtained from this offline evaluation are intended to inform future research and provide guidance regarding the feasibility of model-based decision support in industrial drying applications.

3.3 Functional and Non-Functional Requirements

The design and implementation of an AI-driven kiln drying system require clearly defined functional and non-functional requirements. These requirements will ensure that the system meets performance expectations, complies with industrial standards, effectively integrates industrial operations, and provides accurate and practical user benefits.

3.3.1 Functional Requirements

The functional requirements define the system's core capabilities and ensure that it can effectively and reliably predict, monitor, and adjust the drying process in real time:

- **Accurate MC prediction**: The system must reliably predict MC at different drying stages using AI/ML models.

- **Real-time monitoring and data acquisition**: Sensors must continuously collect real-time data of the kiln system.
- **Automated Dynamic Adjustments**: To maintain optimal drying conditions, the system must dynamically adjust kiln settings (e.g., airflow and wood flow) based on AI predictions.
- **User Alerts and Recommendations**: The system must provide timely alerts and actionable recommendations in the event of abnormal conditions, ensuring process stability and operational efficiency.
- **Data logging and analysis**: The system must store historical process data to support performance analysis and continuous model improvements.

3.3.2 Non-Functional Requirements

The non-functional requirements define the system's reliability, efficiency, and usability.

- **Scalability**: The system must support integration with different kiln types and configurations without requiring major modifications. The software architecture should also accommodate future extensions, including additional ML models and sensor integrations.
- **Energy Efficiency**: The AI-driven control mechanisms must aim to optimize energy consumption while ensuring that kiln drying performance and product quality are maintained.
- Robustness and Fault Tolerance: The system must correctly handle missing or noisy data ensuring that prediction accuracy is not significantly affected.
- **Security and Data Privacy**: All sensitive data must be securely stored and transmitted, in compliance with applicable cybersecurity standards.

3.4 Evaluation Methods

To ensure the reliability and efficiency of the AI-driven kiln drying system, a structured evaluation framework was implemented. This evaluation process involved the analysis KPIs, benchmarking the system against existing methods, and the collection of operator feedback to assess its real-world applicability.

3.4.1 Performance Metrics

The performance of the predictive model was evaluated using multiple quantitative metrics to assess both accuracy and efficiency. MSE and RMSE were employed to measure the precision of the AI model's MC predictions, highlighting any deviations from the observed values. Additionally, MAE was also used to quantify the average discrepancy between predicted and actual MC levels. In addition to prediction accuracy, the evaluation considered energy consumption reduction, analyzing the changes achieved through AI optimized process relative to traditional kiln operations. Finally, process consistency was assessed by examining the model's ability to maintain stable and uniform MC levels across varying drying conditions.

3.4.2 Benchmarking

To evaluate the effectiveness of the AI-driven system, it was benchmarked against more traditional kiln drying techniques. This assessment involved a comparison with traditional methods, evaluating the AI system's ability to optimize drying schedules and improve MC prediction accuracy over conventional empirical approaches. Additionally, the system was tested against other ML techniques such as Random Forest, in order to asses the model robustness and superiority. Computational efficiency examined, considering processing time and resource utilization of the AI system were analyzed to ensurethe feasibility of real-time implementation in industrial kiln operations.

3.4.3 User Feedback and Practical Validation

Beyond quantitative analysis, qualitative feedback was collected to evaluate the system's usability in real-world scenarios. Operator usability tests were conducted at SONAE Arauco, where kiln operators assessed the ease of integration with existing workflows, the clarity of system recommendations and the overall user experience. Industrial deployment trials were performed to test system performance under live operating conditions measuring its performance in live environments. Finally, system adaptability was examined by evaluating the AI's responsiveness to variations in wood composition, initial MC levels and external environmental factors ensuring that the model is robust and flexible across diverse conditions in mind.

Chapter 4

Implementation of Machine Learning Models for Industrial Wood Drying

This chapter presents the practical implementation of the data-driven framework developed in this thesis for modeling and analyzing an industrial wood drying process. Building upon the methodological foundations established in Chapter 3, the focus here is on how the proposed approaches were concretely applied using real industrial data, rather than on theoretical formulation or performance outcomes. The implementation was carried out using historical process data collected from an operating industrial drying line. Consequently, the work reflects realistic industrial conditions, including measurement noise, incomplete data, operational variability, and confidentiality constraints. These factors strongly influenced the implementation decisions described in this chapter and guided the selection of modeling strategies, preprocessing techniques, and evaluation procedures. This chapter details the complete implementation pipeline, beginning with data acquisition and preprocessing, followed by feature selection and the construction of predictive machine learning models. ANN and LSTM networks were implemented to represent static and dynamic modeling approaches, respectively. The implementation emphasizes robustness, reproducibility, and interpretability under industrial constraints, rather than exhaustive optimization or real-time deployment. All analyses and model executions described in this chapter were performed offline using historical data. No closed-loop control or real-time actuation was implemented; instead, the objective was to assess the feasibility of predictive modeling and control-oriented analysis in an industrial context, providing a solid foundation for the evaluation and discussion of results presented in Chapter 5.

4.1 Data Acquisition

The data used in this study was obtained from an industrial wood drying system operated by SONAE Arauco. The dataset consists of historical process records collected during normal industrial operation, reflecting realistic production conditions, including inherent process variability, sensor noise and, operational constraints typical of large-scale industrial drying systems. The drying line is equipped with multiple sensors distributed throughout the process, continuously monitoring key operational variables. These sensors provide measurements of parameters such as air temperature, wood particle temperature, airflow conditions, and wood flow rates. In addition to automatically logged sensor data, certain process-related measurements were obtained through manual acquisition procedures, depending on operational availability and measurement feasibility. As a result, the dataset combines both high-frequency automated measurements and low-frequency manually collected data. Although the industrial system supports real-time data acquisition, this study focuses exclusively on the analysis of historical data. This design choice was made to ensure reproducibility, avoid interference with ongoing production processes and enables a controlled and systematic evaluation of machine learning models under offline conditions. Consequently, no real-time data streaming, online learning or direct interaction with industrial control systems was considered within the scope of this study. A key limitation of the available dataset concerns direct measurements of wood MC. Moisture data was recorded at relatively large time intervals and were not consistently logged across all operating periods. This limitation introduces data sparsity, temporal misalignment, and potential measurement uncertainty, restricting the direct use of MC as a primary modeling target. In contrast, temperature-related variables were continuously monitored and recorded at a significantly higher temporal resolution, providing a dense and reliable representation of process dynamics. Based on data availability, measurement reliability, and established industrial practice, temperature was selected as the primary modeled variable and used as a proxy for the drying state of the system. Temperature evolution within the dryer is strongly linked to moisture removal mechanisms, as heat transfer directly governs evaporation rates and drying efficiency. This relationship is well established in the literature on industrial drying and kiln operation, where temperature profiles are commonly used to characterize drying stages and to indirectly infer moisture evolution. The resulting dataset provides a realistic foundation for the implementation of data-driven predictive models, while accurately reflecting the practical constraints and measurement limitations typical of industrial wood drying environments. The preprocessing procedures applied to address data quality issues and to prepare the dataset for modeling are described in the following section.

4.2 Data Preprocessing

Before implementing the machine learning models, the available industrial data underwent a preprocessing phase aimed at addressing data quality issues and ensuring suitability for predictive modeling. Given the real-world, industrial origin of the dataset, this stage was essential to handle noise, missing values, inconsistencies, and redundant information commonly encountered in industrial process data. Unlike artificially generated datasets, industrial data are collected under varying operating conditions and are subject to sensor faults, communication interruptions, and manual intervention. As a result, preprocessing decisions directly influence model stability, convergence behavior, and generalization capability. For this reason, preprocessing was treated as a critical component of the implementation rather than as a purely technical preliminary step. The preprocessing stage aimed to transform raw process measurements into a structured, numerically stable, and consistent dataset while preserving the underlying dynamics of the drying operation. Particular attention was given to avoid aggressive data manipulation that could distort physical relationships or introduce artificial patterns. The main preprocessing steps included data cleaning, handling of missing and invalid values, removal of non-informative variables and data standardization. Each preprocessing operation was applied uniformly across all experiments to ensure methodological consistency and comparability between models. Additionally, preprocessing procedures were designed to be compatible with both static models, such as ANN, and dynamic models, such as LSTM networks, which impose stricter requirements on temporal continuity and numerical stability. The individual preprocessing steps and their implementation details are described in the following subsections.

4.2.1 Data Cleaning and Consistency Checks

An initial exploratory analysis of the dataset revealed the presence of invalid records, inconsistent entries, and periods with incomplete sensor information. Such issues are characteristic of data collected under normal industrial operation and arise from factors such as sensor maintenance activities, communication interruptions, calibration procedures, and logging inconsistencies. To address these issues, a systematic data cleaning procedure was applied. Records containing invalid timestamps were removed to ensure correct temporal alignment across all process variables. Additionally, observations with clearly erroneous values, such as physically implausible readings or values outside known operational limits, were also excluded from further analysis. These filtering steps were applied conservatively to minimize the removal of valid operating data while reducing the influence of measurement errors. Variables that showed constant or near-constant values, over extended periods of time, were also identified and removed. Such variables do not provide informative variation for predictive modeling and may negatively affect learning performance by introducing redundant inputs or numerical instability, particularly in neural network models. Removing these values improves model efficiency and reduces the risk of overfitting to non-informative features. The data cleaning process ensured temporal consistency across variables and improved the overall reliability of the dataset. By reducing the impact of incorrect measurements and non-informative inputs, this step provided a cleaner and more stable foundation for subsequent preprocessing stages and model implementation.

4.2.2 Handling of Missing and Anomalous Values

Missing values were present in several variables due to sensor unavailability, communication interruptions, and missing value logging during certain operating periods. Such gaps are common in industrial environments and must be handled carefully to avoid introducing artificial patterns or misleading information into the dataset. Records containing excessive amounts of missing information and data were discarded, as aggressive imputation in such cases would introduce significant uncertainty and could distort the underlying process dynamics. For the remaining data, missing values were handled conservatively by removing records containing isolated missing entries rather than applying interpolation or forward-filling techniques. This decision was motivated by the non-stationary and highly dynamic nature of the drying process, where artificial reconstruction of missing values could introduce physically inconsistent patterns and bias model learning. Although this approach reduced the effective dataset size, it ensured that only fully observed and physically consistent process states were used for model training and evaluation. This choice reflects the objective of modeling realistic operating conditions rather than maximizing sample count at the expense of data reliability. In addition to missing data, anomalous sensor readings were identified through a combination of domain knowledge and statistical inspection. These anomalies included sudden spikes, abrupt drops, and/or isolated values inconsistent with known physical behavior and operational constraints of the drying process. Such observations are often associated with temporary sensor faults or logging errors rather than genuine process events. Identified anomalous values were either removed or filtered during preprocessing to reduce noise and prevent distortion of model training. This step was particularly important for neural network models, which are highly sensitive to outliers and may exhibit unstable training behavior or overfitting when exposed to extreme or non-representative values. Overall, this preprocessing stage improved data reliability while preserving a representative depiction of normal operating conditions. By avoiding overly aggressive imputation and carefully handling anomalies, the dataset retained its physical relevance and suitability for both static and dynamic predictive modeling.

4.2.3 Variable Filtering and Dimensionality Reduction

Given the large number of available process variables, an initial filtering stage was applied to reduce input dimensionality before performing feature relevance analysis. This step aimed to remove variables that were clearly irrelevant to the drying process or unrelated to the modeling objectives, based on operational knowledge and an understanding of the industrial system. Variables associated with administrative signals, redundant status indicators, or measurements unrelated to thermal or material transport phenomena were excluded at this stage. The purpose of this preliminary filtering was not to perform detailed feature selection, but rather to eliminate variables that would not reasonably contribute to predictive modeling of the drying process. This filtering step was intentionally conservative. Only variables with negligible informational value or no physical relevance to the drying dynamics were removed, while potentially informative variables were retained even if their contribution was uncertain at this stage. This approach minimized the risk of prematurely discarding relevant process information and ensured that subsequent feature relevance analysis was conducted on a sufficiently rich input space. The refined dataset resulting from this stage served as the input for the feature relevance analysis described in Section 4.3, where data-driven methods were applied to further assess the contribution of individual variables to predictive performance.

4.2.4 Data Standardization

To ensure numerical stability and consistent learning behavior during model training, all input variables were standardized prior to predictive modeling. Standardization was performed by transforming each variable to have zero mean and unit variance, using statistics computed exclusively from the training dataset. This normalization step is particularly important for neural network models, as the heterogeneous scales of industrial process variables can otherwise lead to unstable gradients, slower convergence, or biased learning toward variables with larger numerical magnitudes. By placing all inputs on a common scale, the learning algorithms are able to treat each variable with comparable importance during optimization. To prevent information leakage, the standardization parameters derived from the training data were applied unchanged to the validation and test datasets. This ensured that model evaluation accurately reflected performance under realistic deployment conditions, where future data are normalized using statistics computed from historical observations only. Data standardization was applied consistently across all experiments and modeling approaches. For the LSTM model, standardization was performed prior to sequence construction to guarantee that temporal inputs remained numerically stable across all time steps. Following this stage, the dataset was considered suitable for feature relevance analysis and predictive model implementation.

4.2.5 Dataset Characterization and Dimensionality Evolution

To ensure methodological transparency and to explicitly address dataset representativeness, a quantitative characterization of the industrial dataset was performed after preprocessing.

The raw dataset contained a total of 21067 time-stamped samples and 312 recorded process variables, reflecting real operating conditions of an industrial wood drying line. These variables included temperature measurements at different stages of the dryer, airflow-related signals, pressure indicators, motor currents, fan speeds, setpoints, and material transport variables.

Dataset Reduction After Cleaning

During preprocessing, invalid timestamps were removed, columns where the values were constant were eliminated and non-informative and irrelevant variables were filtered. As a consequence, a total of 69 constant variables were removed, reducing the dimensionality from 312 to 243 candidate variables, while preserving the entirety of the time samples.

The percentage reduction in dimensionality was:

$$\frac{312-243}{312} \times 100 \approx 22.1\%$$

This reduction improved numerical stability and eliminated redundant signals without compromising temporal coverage.

Training and Testing Split

The dataset was sorted chronologically and divided using an 80/20 temporal split, ensuring that future studies were never used to predict past states. This divided the full dataset into two sets:

- **Training set:** 16853 samples
- **Testing set:** 4214 samples

This approach simulates realistic industrial deployment conditions.

LSTM Effective Sample Size

For the LSTM model, input data was structured into fixed-length sequences using a sliding window of W = 30 time steps.

Because sequence construction requires "W" past observations to predict the next one, the effective number of usable samples becomes:

- **Training sequences:** 16823
- **Testing sequences:** 4184

This reduction is expected when it comes to temporal modeling and it reflects the cost of incorporating process memory.

Final Feature Set Used for Modeling

After permutation-based feature relevance analysis, a reduced feature set was defined for predictive modeling.

Both ANN and LSTM models were trained using 32 selected variables derived from the filtered dataset. There variables included temperature measurements at critical dryer zones, airflow-related signals, combustion indicators, motor-related measurements, and selected actuator positions relevant to process control.

This dimensionality reduction from 243 candidate variables to just 32 modeling features represents an additional reduction of:

$$\frac{243-32}{243} \times 100 \approx 86.8\%$$

This substantial reduction improves interpretability and reduces overfitting risk while preserving the most informative process variables.

4.3 Feature Selection

Following the dataset characterization presented in Section 4.2.5, a structured feature selection procedure was implemented to identify the most relevant process variables for predictive modeling.

Although the dataset had already been reduced during preprocessing, the remaining candidate variables still represented a high-dimensional industrial system. In complex thermal processes such as wood drying, multiple signals may be correlated, indirectly related, or partially redundant. Including all available variables could increase model complexity without necessarily improving predictive performance, while also reducing interpretability and increasing overfitting risk.

For these reasons, feature selection was treated as a deliberate implementation stage rather than a secondary optimization step.

The objectives of this phase were:

- Reduce dimensionality while preserving predictive capability
- Improve training stability and generalization
- Identify the process variables that most strongly influence temperature evolution

To achieve this, a permutation-based feature importance approach was adopted. This method was selected because:

- It is suitable for nonlinear models
- It does not assume linear relationships
- It evaluates variables using the same trained predictive model
- It provides an intuitive measure of practical impact

Rather than relying on correlation metrics or domain-only filtering, feature importance was assessed using the actual predictive behavior of the implemented models.

Because the ANN and LSTM models represent fundamentally different modeling paradigms (static vs. temporal), feature selection was performed independently for each architecture.

4.3.1 Feature Selection Procedure for the ANN Model

For the ANN model, feature relevance was evaluated using instantaneous input vectors corresponding to individual time steps.

After training the ANN on the training dataset, feature importance was computed using the test dataset in order to ensure that relevance estimates reflected behavior under unseen operating conditions.

The procedure followed these steps:

1. Train the ANN model using the selected training data
2. Compute baseline predictive performance on the test set
3. For each input variable:
   - Randomly permute its values across the test samples
   - Recompute model predictions
   - Measure the degradation in predictive performance
4. Rank variables according to the magnitude of performance degradation

If permuting a variable significantly worsened prediction accuracy, that variable was considered highly influential. Conversely, variables whose permutation produced negligible performance change were interpreted as weak contributors.

This approach allowed the identification of process variables that most strongly influence static nonlinear temperature prediction.

Importantly, this procedure evaluates variables in the context of the trained ANN model. Therefore, relevance is defined in terms of actual predictive contribution rather than theoretical importance.

Based on this analysis, a reduced ANN-specific feature set was defined and subsequently used for all ANN-based experiments.

4.3.2 Feature Selection Procedure for the LSTM Model

Because the LSTM model operates on temporal sequences rather than isolated time steps, the feature selection procedure was adapted accordingly.

In the LSTM formulation, each input sample consists of a sequence of consecutive time steps. Therefore, permuting a variable at a single time index would disrupt sequence coherence and produce unrealistic perturbations.

To preserve temporal structure, permutation was applied consistently across entire sequences. The procedure followed these steps:

1. Train the LSTM model using sequential training data
2. Compute baseline prediction loss
3. For each input variable:
   - Permute its values across all time steps within the sequences
   - Preserve the remaining variables and sequence structure
   - Recompute model predictions
   - Measure the increase in prediction error
4. Rank variables according to performance degradation

This sequence-consistent permutation approach ensures that feature relevance reflects the influence of both instantaneous and delayed temporal effects.

Variables that caused significant increases in prediction error were interpreted as strongly influencing the dynamic behavior learned by the LSTM model.

Because the LSTM incorporates process memory, the resulting relevance pattern differs from that of the ANN. Some variables exhibit stronger importance under temporal modeling due to delayed or accumulative effects.

Based on this analysis, a reduced LSTM-specific feature set was defined and fixed for all subsequent LSTM experiments.

4.3.3 Implementation Considerations and Stability of Feature Selection

To ensure robustness, feature importance was evaluated using multiple permutation repetitions. This aims to reduces the impact of random fluctuations and ensures that relevance rankings are more stable rather than noise-driven.

Feature retention was conservative. Variables were selected based on consistent influence rather than isolated peaks in importance. This approach reduces the risk of excluding variables that may contribute under specific operating conditions.

It is also important to mention that feature selection was not designed to minimize the number of inputs aggressively, but rather designed to identify a stable and interpretable subset of variables that preserves predictive power while reducing unnecessary complexity.

The final feature sets used for modeling are described in Section 4.2.5.

4.3.4 Role of Feature Selection in Control Analysis

Beyond predictive performance, feature selection also supports interpretability within the control recommendation framework.

By identifying variables with strong predictive influence, the analysis provides insight into:

- Which process signals most strongly affect temperature evolution
- Which actuators may have greater leverage for control
- Which variables exhibit negligible practical impact

This interpretability is essential when transitioning from pure prediction toward advisory or control-oriented applications.

The quantitative feature importance rankings and their interpretation are presented and discussed in Chapter 5.

4.4 Machine Learning Prediction Models

This section describes the concrete implementation of the predictive models developed in this work. While Chapter 3 introduced the conceptual rationale behind the selected modeling approaches, the present section focuses on the practical implementation choices, architectural configurations, and training procedures adopted for the industrial dataset under study.

Two distinct machine learning architectures were implemented: a feedforward ANN and a LSTM network. These models represent two fundamentally different modeling paradigms. The ANN captures static nonlinear relationships between process variables and the target temperature at a given time step, whereas the LSTM explicitly incorporates temporal dependencies by processing sequences of historical observations. Both models were implemented under consistent experimental conditions in order to allow a controlled and fair comparison of their predictive and control-oriented behavior.

4.4.1 Problem Formulation and Target Definition

The predictive task addressed in this work was formulated as a supervised regression problem. Given a set of process variables describing the operational state of the drying system, the objective of the model is to estimate the outlet temperature at a given time step. As discussed in previous sections, temperature was adopted as a proxy for the drying state due to its strong physical relationship with moisture removal mechanisms and its reliable, high resolution availability in the industrial dataset. This formulation reflects a practical industrial scenario in which the current or future thermal state of the drying process must be inferred from measured operational variables. The regression-based approach allows the models to learn continuous nonlinear mappings between process inputs and the target temperature without imposing explicit physical equations. For the ANN model, inputs consist of instantaneous measurements of the selected process variables at a single time step. Each input vector represents a snapshot of the current process state, and the model produces a scalar output corresponding to the predicted temperature at that same time step. This formulation assumes that the instantaneous state contains sufficient information to estimate the target variable and serves as a baseline representation of static nonlinear relationships within the system. For the LSTM model, inputs are defined as sequences of consecutive time steps constructed using a sliding window approach. Each input sample consists of a fixed-length sequence of past process observations, representing recent process history. The model outputs a single scalar temperature prediction corresponding to the subsequent time step. This sequence-to-one formulation enables the LSTM model to capture temporal dependencies, delayed effects, and process inertia that are characteristic of industrial drying systems and cannot be represented by static models. By defining distinct input-output formulations for the ANN and LSTM models, this work enables a structured comparison between static and dynamic predictive modeling approaches under consistent industrial conditions.

4.4.2 Artificial Neural Network (ANN) Implementation

The ANN model was implemented as a feedforward multilayer perceptron configured to approximate nonlinear relationships between the selected input variables and the target temperature. The network receives a vector composed of the 32 features selected during the feature relevance analysis stage and produces a single continuous output.

The adopted architecture consists of two fully connected hidden layers with 64 and 32 neurons, respectively, followed by a linear output layer. This configuration was selected after preliminary experimentation and specially based on the state of the art study, balancing representational capacity and computational efficiency. The objective was not to identify an optimal architecture through extensive hyperparameter tuning, but rather to establish a representative static modeling baseline capable of capturing nonlinear dependencies without excessive model complexity.

Input variables were standardized prior to training using parameters computed exclusively from the training dataset. This step was essential to ensure numerical stability and to prevent scale dominance among heterogeneous industrial variables. The model was trained using MSE as the loss function and the Adaptive Moment Estimation (Adam) optimization algorithm.

Adam is a gradient-based optimization method that automatically adjusts the learning rate of each model parameter during training. Instead of using a fixed learning rate, Adam adapts the update step size based on past gradients, allowing faster convergence and improved stability. This makes it particularly suitable for neural network training, especially when dealing with heterogeneous and scaled industrial data.

Early stopping mechanisms were also employed based on validation performance within the training set, reducing the risk of overfitting while preserving generalization capability.

The ANN formulation treats each time step independently. Each input sample corresponds to an instantaneous snapshot of process variables, and the model produces a temperature estimate for that same time step. No explicit temporal information is provided to the network. Consequently, the ANN assumes that the instantaneous process state encapsulates sufficient information to approximate the drying condition. This characteristic makes the ANN suitable as a baseline model for static nonlinear regression but inherently limits its ability to represent delayed process effects and dynamic system inertia.

During implementation, the ANN exhibited fast convergence and stable optimization behavior. However, preliminary experiments performed prior to feature selection revealed sensitivity to redundant and weakly informative variables, reinforcing the importance of dimensionality reduction before model training.

4.4.3 Long Short-Term Memory (LSTM) Implementation

To address the temporal nature of the drying process, an LSTM-based neural network was implemented. The LSTM architecture was selected due to its capability to retain and update internal memory states, enabling the modeling of sequential dependencies that arise from process inertia, thermal accumulation, and delayed system responses.

In this implementation, the input data was transformed into sequences using a fixed-length sliding window approach. A window size of W=30 time steps was adopted, meaning that each input sample will consist of 30 consecutive observations of the selected process variables mentioned earlier, and the model is trained in order to predict the temperature at the subsequent time step. This configuration allows the network to learn how recent historical process behavior influences near-future temperature evolution.

The choice of a 30-step window was guided by preliminary experimentation and engineering considerations. Shorter windows were found to insufficiently capture process inertia, whereas substantially longer windows increased computational cost without consistent performance gains. The selected window length represents a compromise between temporal depth and practical feasibility.

The adopted architecture of the LSTM implementation consists of two stacked LSTM layers with 64 and 32 units, respectively, followed by a dense output layer producing a single scalar prediction. The stacked configuration allows the network to extract hierarchical temporal patterns, where lower layers capture short-term fluctuations and higher layers model longer-term dependencies.

As with the ANN, input variables were standardized using training data statistics. Sequence construction was performed separately for training and testing subsets to avoid any temporal leakage across dataset boundaries. The LSTM model was trained using MSE as the loss function and the Adam optimizer, with early stopping applied to prevent overfitting.

When compared to the ANN, the LSTM required significantly higher computational resources and training time due to the sequential processing of input data. Additionally, the transformation into sequences resulted in a reduced effective number of training samples, as described in Section 4.2.5. Despite these additional costs, the LSTM formulation provides the structural capability to represent dynamic behavior that cannot be captured by static inputâĂŞoutput mappings.

4.4.4 Experimental Consistency and Model Comparison Design

To ensure methodological rigor, both predictive models were implemented under comparable experimental conditions, like using the same chronologically ordered dataset, with an 80/20 temporal split between training and testing sets. Standardization procedures were applied consistently and performance evaluation metrics were identical across models. The same target definition and feature sets were also adopted for both architectures.

This controlled setup ensures that differences observed in Chapter 5 can be attributed to the modeling paradigm itself instead of the discrepancies in preprocessing, dataset composition, or evaluation procedures. Therefore, the ANN serves as a static baseline, while the LSTM represents a temporally informed modeling approach more familiar with dynamic approaches.

4.4.5 Scope and Implementation Boundaries

It is important to clarify the scope of the implemented models. The objective of this work was to evaluate predictive behavior and control-oriented sensitivity in an offline setting using historical industrial data. The models were not deployed in real-time industrial environments, nor were they integrated with existing control hardware.

Actuator dynamics, real-time latency, safety constraints, and closed-loop stability were not enforced within the modeling framework. Instead, the models were developed to support offline predictive analysis and to serve as the foundation for the control recommendation framework described in subsequent sections.

By limiting the scope to offline evaluation, this study isolates the predictive and sensitivity characteristics of static and dynamic modeling approaches under realistic industrial data conditions, without introducing additional layers of implementation complexity.

Chapter 5

Results and Discussion

This chapter presents and discusses the results obtained from the implementation of the machine learning models described in Chapter 4. The primary objective of this chapter is to evaluate predictive performance, analyze feature relevance outcomes, and assess the behavior of the proposed control recommendation framework under offline industrial conditions.

The results are presented in a structured manner to support a transparent and systematic comparison between static and temporal modeling approaches. Particular attention is given to the comparison between ANN and LSTM networks, highlighting differences in predictive accuracy, generalization capability, stability, and sensitivity to process variables. All evaluations are performed using historical industrial data that were not used during model training, following the chronological splitting strategy previously described.

In addition to predictive accuracy, this chapter examines the implications of model behavior for control-oriented applications. Offline control recommendation experiments are analyzed to evaluate how each modeling approach responds to variations in controllable process variables and to assess the feasibility of using data-driven models to support operator decision-making.

5.1 Experimental Setup and Evaluation Framework

This section describes the experimental setup adopted to evaluate the predictive performance of the implemented machine learning models and to ensure that the comparison between static and temporal modeling approaches is both fair and methodologically rigorous. The evaluation framework was designed to reflect realistic industrial deployment constraints while maintaining reproducibility and analytical consistency.

All experiments were conducted using the historical industrial dataset characterized in Chapter 4. The dataset consisted of 21067 time-stamped samples after preprocessing, with 243 candidate variables prior to feature selection. As described previously, a reduced set of 32 features was defined and fixed for predictive modeling.

Prior to model training, the dataset was sorted chronologically to preserve the temporal evolution of the drying process. A time-based data splitting strategy was then applied, allocating approximately 80% of the earliest observations (16853 samples) for training and internal validation, and reserving the remaining 20% (4214 samples) for testing. This chronological split prevents information leakage across time and mirrors realistic industrial conditions, where models must predict future process behavior using only past information.

Both the ANN and the LSTM models were trained and evaluated using the same chronological division to ensure consistency. For the LSTM model, the construction of temporal sequences through a sliding window (W = 30) reduced the effective number of usable samples to 16823 training sequences and 4184 testing sequences. This reduction is inherent to sequence-based modeling and was explicitly considered during comparative result interpretation.

Input features were standardized using statistics computed exclusively from the training data. The same transformation parameters were subsequently applied to validation and test sets to avoid information leakage. Model training incorporated internal validation procedures, including early stopping mechanisms, to reduce overfitting and improve generalization stability. The test dataset was excluded from all training and tuning procedures and was used solely for final performance assessment.

5.1.1 Evaluation Metrics

To evaluate predictive performance, a few regression metrics were employed, providing additional perspectives on model accuracy, robustness and relevance. The selected metrics were the MAE, MSE, and Coefficient of Determination (R^2). The MAE measures the average magnitude of prediction errors and provides an intuitive interpretation of the model accuracy by showing deviations in the same physical units as the target variable. This makes MAE particularly suitable for industrial drying applications, where absolute temperature deviations are what operators directly interpret on their day to day tasks. The MSE penalizes larger prediction errors more heavily, highlighting instances of significant deviation between predicted and true values. This metric is useful in identifying model instability and large prediction failures that may be crucial in control oriented contexts. The R^2 quantifies the proportion of variance in the observed data explained by the model. While it does not directly show error magnitude, R^2 provides a normalized measure of model fit and supports comparative analysis between different modeling approaches when interpreted alongside MAE and MSE. All of this metrics were computed only on the test dataset. No model selection or tuning decisions were made with test performance in mind, in order to make sure that an unbiased and transparent assessment of generalization capability.

5.2 Predictive Performance Results

This section presents the predictive performance obtained for the ANN and LSTM models described in Chapter 4. The objective of this analysis is to evaluate the generalization capability of both modeling approaches when applied to unseen industrial data and to compare their suitability for representing the thermal dynamics of the drying process.

All results reported in this section correspond exclusively to the test dataset defined in Section 5.1. This dataset represents the most recent chronological segment of the available historical data and was strictly excluded from all training, validation, and model selection procedures. Consequently, the performance metrics presented here reflect true out-of-sample predictive capability under future operating conditions.

Results are reported separately for Dryer 1 and Dryer 2, allowing independent assessment of model behavior across distinct thermal subsystems of the drying line. This separation is important because each dryer operates under different physical constraints, airflow patterns, and combustion interactions, which may influence model performance.

The comparative analysis focuses on three aspects:

1. absolute prediction accuracy,
2. error stability and sensitivity to large deviations, and
3. variance explanation capability.

These dimensions are examined jointly through the evaluation metrics defined previously, enabling a structured comparison between static nonlinear regression (ANN) and temporal sequence modeling (LSTM with $W = 30$).

5.2.1 Overall Predictive Performance on Test Data

Table 5.1 presents the predictive performance of the ANN and LSTM models on the test dataset for the two analyzed dryers. For the LSTM model, the notation $W = 30$ refers to the temporal window length used during sequence construction.

Each LSTM input sequence contains the previous 30 consecutive time steps of process variables, and the model predicts the dryer temperature at the subsequent time step. This window length was selected as a compromise between capturing sufficient process memory and maintaining computational stability, given the available sampling frequency and dataset size.

Model	**Dryer**	**MAE**	**MSE**	**R^2**	**Number of Samples**
ANN	Dryer 1	82.04	11156.05	-7.59	4214
LSTM (W=30)	Dryer 1	12.28	557.02	0.57	4214
ANN	Dryer 2	39.13	2376.76	-1.36	4214
LSTM (W=30)	Dryer 2	11.15	213.01	0.79	4214

Table 5.1: Predictive performance of ANN and LSTM models on the test dataset

The results reveal a substantial performance gap between static and temporal modeling approaches. The ANN model exhibits very large prediction errors on unseen data. For Dryer 1, the mean absolute error exceeds 82 °C, while for Dryer 2 it reaches approximately 39 °C. Considering that industrial drying control decisions often depend on deviations of only a few degrees Celsius, such errors are operationally unacceptable.

The negative R^2 values obtained by the ANN model further reinforce this conclusion. An R^2 below zero indicates that the model performs worse than a naive baseline predictor that simply outputs the mean temperature of the dataset. In practical terms, this means that although the ANN was capable of fitting the training data, it failed to capture patterns that remain valid under new operating conditions.

In contrast, the LSTM model demonstrates significantly improved generalization performance. For Dryer 1, the MAE decreases from 82.04 °C to 12.28 °C, representing an error reduction of approximately 85%. For Dryer 2, the reduction is from 39.13 °C to 11.15 °C, corresponding to a decrease of approximately 71%. Additionally, the R^2 values of 0.57 and 0.79 indicate that the LSTM model explains a substantial proportion of the variance in the observed temperature data.

These results indicate that incorporating temporal context through a sliding window of past observations allows the LSTM model to capture process inertia, delayed thermal responses, and dynamic interactions between variables. Such effects are inherently inaccessible to static models based solely on instantaneous process snapshots.

5.2.2 Training vs Testing Performance Analysis

Model	Dataset	Dryer	MAE	MSE	R^2	Number of samples
ANN	Training	Dryer 1	0.93	1.81	0.95	16853
ANN	Testing	Dryer 1	82.04	11156.05	-7.59	4214
LSTM (W = 30)	Training	Dryer 1	3.78	22.65	0.38	16823
LSTM (W = 30)	Testing	Dryer 1	12.28	557.02	0.57	4214
ANN	Training	Dryer 2	1.18	4.62	0.96	16853
ANN	Testing	Dryer 2	39.13	2376.76	-1.36	4214
LSTM (W = 30)	Training	Dryer 2	2.30	12.31	0.90	16823
LSTM (W = 30)	Testing	Dryer 2	11.15	213.01	0.79	4214

Table 5.2: Training and testing performance comparison

A comparison between training and testing performance provides crucial insight about the generalization capability and robustness of the predictive models. While high accuracy on training data may suggest good learning capabilities, industrial applicability also depends on stable performance under unseen and sometimes unclear operating conditions.

The ANN model shows extremely strong performance on the training dataset. For Dryer 1, it achieves an MAE of 0.93°C and an R^2 of 0.95. For Dryer 2, the results are also high, with an MAE of 1.18°C and an R^2 of 0.9 which indicate that the network is capable of fitting the training data almost perfectly.

However, when the model is evaluated on the test dataset, the performance of the ANN deteriorates drastically; for Dryer 1, the MAE increases from 0.93°C to 82.04°C and for Dryer 2, the MAE increases from 1.18°C to 39.13°C. The corresponding R^2 values become negative in both cases, confirming that the model performs worse than a baseline predictor that just outputs the mean temperature.

This extreme difference between training and testing performance is a clear indication of severe overfitting. The ANN model successfully captures patterns specific to the training period but it fails to extract stable relationships that remain valid under different operating conditions and because it relies exclusively on instantaneous input vectors, it lacks contextual awareness of process evolution. In a system characterized by thermal inertia, delayed effects, and regime shifts, this limitation becomes critical.

The LSTM model exhibits a fundamentally different behavior. While its training performance is less extreme âĂŤ for example, an MAE of 3.78°C for Dryer 1 and 2.30°C for Dryer 2 âĂŤ the degradation between training and testing is significantly smaller. On the test dataset, MAE values increase to 12.28°C and 11.15°C, respectively. Although the increase is noticeable, it remains within a stable and interpretable range. Importantly, the R^2 values remain positive and relatively high

(0.57 and 0.79), confirming that the model preserves predictive structure under new data.

From a generalization standpoint, the LSTM model demonstrates significantly higher robustness. The inclusion of temporal context through sequential inputs effectively acts as an implicit regularization mechanism. By conditioning predictions on recent process history, the model avoids extreme reliance on isolated correlations that may not persist across different operational regimes.

In industrial environments where process conditions vary over time, this stability is more valuable than near-perfect historical fit. The results therefore reinforce the importance of temporal modeling for dynamic process systems such as wood drying, where delayed interactions and accumulated thermal effects play a dominant role.

5.2.3 Interpretation of Static versus Temporal Modeling Approaches

The contrasting behaviors observed between the ANN and LSTM models reflect fundamental differences between static and temporal modeling paradigms when applied to industrial drying systems.

Feedforward ANN models treat each time step as an independent observation. The prediction at a given instant is based solely on the instantaneous values of the selected process variables. While this formulation can be effective for systems where outputs depend primarily on current inputs, it is inherently limited when modeling processes governed by cumulative physical phenomena.

Industrial wood drying is characterized by thermal inertia, delayed moisture migration, and progressive heat transfer. The temperature observed at a given moment is not only a function of current airflow or actuator positions, but also of the recent evolution of these variables. These delayed and accumulated effects cannot be fully inferred from a single process snapshot.

This limitation is clearly reflected in the experimental results. Although the ANN achieves near-perfect training accuracy, its performance collapses on the test dataset. The model effectively captures correlations specific to the training period but fails to learn stable dynamic relationships that remain valid under different and dynamic operating conditions. Without temporal context, the model has no information about whether the system is heating up, stabilizing, or cooling down, leading to unstable and inconsistent predictions when exposed to unseen conditions.

In contrast, the LSTM model incorporates explicit temporal structure by processing sequences of past observations. In this work, a sliding window of W = 30 time steps was used, meaning that each prediction is conditioned on the 30 most recent measurements of the selected variables. This design allows the model to capture short-term process memory while remaining computationally manageable. The window length was selected as a compromise between representing sufficient historical context and avoiding excessive model complexity.

By analyzing recent process evolution rather than isolated measurements, the LSTM model learns how temperature trajectories develop over time. This enables it to represent delayed thermal responses, progressive energy accumulation, and gradual transitions between operating states. The improved generalization performance observed in Section 5.2.1 confirms that incorporating temporal memory is not simply an architectural refinement but a structural necessity for this type of industrial system.

Importantly, the LSTM does not attempt to fit each observation independently. Instead, it learns patterns of evolution, resulting in smoother predictions that remain physically plausible across changing conditions. The reduced performance gap between training and testing further supports the interpretation that temporal modeling provides a stabilizing effect by constraining predictions to remain coherent with recent process history.

Overall, the results demonstrate that temporal modeling is essential for reliable prediction in industrial wood drying applications. While static ANN models may provide useful insights under stable conditions or for exploratory analysis, sequence-based approaches such as LSTM networks are substantially more suitable for real-world deployment scenarios, where operating conditions evolve continuously and system dynamics are strongly path-dependent.

5.3 Feature Relevance Results

This section presents the results of the feature relevance analysis performed for both the ANN and LSTM models. The objective of this analysis is to quantify the relative contribution of each input variable to temperature prediction and to examine how feature sensitivity differs between static and temporal modeling approaches.

As described in Section 4.3, feature relevance was evaluated using a permutation-based importance methodology. For the ANN model, feature importance corresponds to the normalized degradation in predictive performance when the values of a single input variable are randomly permuted while all others are kept unchanged. For the LSTM model, permutation was applied consistently across all time steps of the input sequences, and importance was measured as the increase in prediction loss resulting from this disruption.

The relevance analysis was conducted on the reduced candidate set of 243 variables obtained after preprocessing. Based on the permutation results, a final subset of 32 variables was selected for predictive modeling. These variables were retained because they consistently showed significant influence on model performance across multiple permutation runs and across both dryer outputs.

Due to industrial confidentiality requirements, individual sensor identifiers are not disclosed in this document. However, it is important to clarify that the original variable names were fully available and explicitly used during preprocessing, feature selection, modeling, and control analysis. Only the reporting format has been anonymized for publication purposes. All numerical results presented in this section are derived directly from the real industrial dataset described in Chapter 4.

To facilitate interpretation while respecting confidentiality constraints, variables are discussed according to their physical function within the process, including:

- Temperature measurements at different dryer stages
- Airflow and combustion-related signals
- Motor currents and mechanical load indicators
- Actuator positions and setpoint references
- Material transport and feed rate variables

The following subsections present the ranked feature relevance results for each modeling approach and discuss the observed differences in sensitivity patterns between static and temporal formulations.

5.3.1 Feature Relevance Patterns in the ANN Model

Rank	Feature Name	Normalized Importance	Variable Category
1	Bsh -4 corrente motor rotação (°C)	0.061	Mechanical / Drive
2	Recalor -5 ventilador - velocidade (%)	0.051	Airflow Control
3	T2-72a5:- 631tt5160 temp 2 right combustor ()	0.049	Thermal
4	Coolair.av_tempbshclient ()	0.044	Thermal
5	Caldeira - start emergência chaminé (Horas)	0.041	Operational State
6	Recalor -5 ventilador - temperatura saída ar (%)	0.030	Air Temperature
7	Caldeira - start up chaminé (dias)	0.028	Operational State
8	S1-143b3 635tt2100 temp primary t.o. pump 3 ()	0.028	Thermal
9	T5-79a1 820tt0700 temperature fluegases eco 2.1 ()	0.025	Thermal
10	Caldeira - tempo abertura start emergência chaminé (sec)	0.024	Operational Timing

Table 5.3: ANN feature relevance results (Dryer 1)

The ANN feature relevance results for the dryer 1 reveal a highly concentrated importance structure, where a small number of instantaneous process variables account for a significant portion of the models predictive sensitivity. The most influential variables are primarily related to immediate thermal conditions, airflow control and mechanical operation. This concentration suggests that the ANN model relies heavily on short term statistical correlations between current measurements and the target temperature. While this behavior enables the ANN to achieve very low prediction error during training, it also makes the model highly sensitive to variations in a limited subset of signals. When these dominant variables change their behavior under new operating conditions, the ANN lacks sufficient contextual information to compensate which result in unstable predictions.

Rank	Feature Name	Normalized Importance	Variable Category
1	Recalor -5 ventilador - temperatura saída ar (%)	0.391	Air Temperature
2	Recalor -5 ventilador - velocidade (%)	0.071	Airflow Control
3	Caldeira - start up chaminé (dias)	0.032	Operational State
4	Caldeira - tempo abertura start emergÃłncia chaminé (sec)	0.027	Operational Timing
5	Eixo pu -5 silo humido - caudal de saída madeira (%)	0.024	Material Flow
6	Caldeira - start emergência chaminé (Horas)	0.022	Operational State
7	S1-143b3 635tt2100 temp primary t.o. pump 3 ()	0.019	Thermal
8	T2-72a5:- 631tt5160 temp 2 right combustor ()	0.017	Thermal
9	Recalor -5 ventilador - temperatura rolamento3 (°C)	0.016	Mechanical
10	Coolair.av_temprecalorclient ()	0.016	Thermal

Table 5.4: ANN feature relevance results (Dryer 2)

For Dryer 2, the ANN relevance distribution is even more concentrated, with a single feature accounting for nearly 40% of the total importance. This extreme dominance indicates that the ANN model bases its predictions primarily on one highly correlated signal, which significantly increases susceptibility to overfitting and poor generalization.

5.3.2 Feature Relevance Patterns in the LSTM Model

Rank	Feature Name	Normalized Importance	Variable Category
1	Bsh -2 temperatura entrada depois da mistura (°C)	0.0135	Thermal
2	Bsh -3 depressão secador (°C)	0.0118	Pressure / Depression
3	Bsh -4 corrente motor rotação (°C)	0.0058	Mechanical
4	Coolair.av_temprecalorclient ()	0.0045	Thermal
5	Recalor -5 ventilador - temperatura saída ar (%)	0.0043	Air Temperature
6	Recalor -4 temperatura saída1 (°C)	0.0043	Thermal
7	Coolair.av_tempbshclient ()	0.0035	Thermal
8	S1-145b3 xxxtt0xxx temp secondary t.o. pump 1 (kt press) ()	0.0034	Thermal
9	246m011 - velocidade (RPM)	0.0034	Mechanical
10	Bsh -2 posição clapete2 (°C)	0.0033	Airflow Actuation

Table 5.5: LSTM feature relevance results (Dryer 1, W = 30)

Unlike the ANN, the LSTM model shows a distributed relevance structure, where multiple variables contribute in a meaningful way to predict the model performance. There is not a dominant variable on the relevance ranking, and importance is spread across thermal, mechanical, airflow and pressure related features. This behavior shows that the temporal formulation of the LSTM, where variables influence predictions not only through their instantaneous values but also through their evolution over time. As a result, variables that might look weak in a static context become relevant when their temporal patterns are considered hence, gaining strength in a dynamic model.

Rank	Feature Name	Normalized Importance	Variable Category
1	Recalor -5 ventilador - temperatura saída ar (%)	0.0277	Air Temperature
2	Recalor -2 temperatura entrada depois da mistura (°c)	0.0104	Thermal
3	Recalor -3 temperatura entrada (°C)	0.0070	Thermal
4	Recalor -2 temperatura ar de entrada (mbar)	0.0068	Air Pressure
5	Bsh -5 temperatura saída1 (°C)	0.0060	Thermal
6	Eixo pu -1 alimentação serrim - tempo entre movimentos (BAR)	0.0027	Material Feed
7	T1-68b3 410tt0130 fire prot fls 1 ()	0.0026	Safety / Protection
8	Coolair.av_temprecalorclient ()	0.0025	Thermal
9	Recalor - setpoint temp. saída recalor (°C)	0.0023	Control Setpoint
10	T1-18p5:- 510pt1280 flow transmitter secondary air ()	0.0021	Airflow Measurement

Table 5.6: LSTM feature relevance results (Dryer 2, W = 30)

The relevance distribution for dryer 2 confirms the robustness of the LSTM model, with contributions emerging from multiple subsystems rather than from isolated measurements

5.3.3 Comparative Interpretation and Implications

The quantitative feature relevance results provide strong evidence for the fundamentally different ways in which ANN and LSTM models interpret industrial process data. The ANN model exhibits a high concentration of importance, sowing that it heavily relies on a small number of dominant instantaneous variables. This explains both its strong training performance and its poor generalization to unseen operating conditions. On the other hand, the LSTM model distributes relevance across a larger set of variables and subsystems. By incorporating temporal context through a sliding window of 30 time steps, the LSTM is able to capture cumulative effects, delayed responses and process inertia that are essential to industrial drying operations. From a practical perspective, this distributed sensitivity is highly desirable. Models that rely on a narrower set of variables are more susceptible to instability when those variables change, in contrast with models that integrate information from multiple sources over time, which tend to produce smoother and more reliable predictions. These findings reinforce the predictive performance results presented in Section 5.2 and support the conclusion that temporal modeling is essential for a robust and reliable prediction in industrial wood drying processes.

5.4 Control Recommendation Results

This section presents the results of the offline control recommendation analysis conducted using the predictive models described in previous chapters. The objective of this analysis was to evaluate how the ANN and LSTM models respond to systematic changes in controllable process variables and to assess their suitability for control oriented decision support in an industrial drying context. It is important to mention that the control framework implemented in this work operates entirely offline and does not interact with the real industrial control system. The generated recommendations should therefore be interpreted as advisory insights rather than actionable control commands. Nevertheless, this evaluation provides valuable information regarding model sensitivity, stability and practicality for future integration into decision support or advanced control strategies.

5.4.1 Offline Control Evaluation Setup

Control recommendations were generated for a representative subset of samples extracted directly from the available test dataset. For each selected time point, the predictive model was provided with the current process state (ANN) or the most recent temporal sequence (LSTM), and systematic perturbations were applied to a predefined subset of controllable variables. Controllable variables were identified based on operational knowledge and correspond to actuator related parameters such as airflow regulation, material feed rates and valve positions. Each controllable variable was independently changed within bounded ranges consistent with normal industrial operation, while all non controllable variables were held constant. For each change, the model was used to predict the resulting dryer temperature. The objective of the control recommendation was to reduce the deviation between the predicted temperature and a predefined target value.

5.4.2 Evaluation Metrics for Control Performance

The effectiveness of the control recommendations was assessed using multiple quantitative metrics designed to capture both outcome quality and control effort:

- **Mean improvement:** average reduction in absolute prediction error relative to the target temperature.
- **Median improvement:** robust measure of typical improvement.
- **Success rate:** percentage of cases in which the predicted temperature fell within a predefined tolerance band around the target value.
- **Mean control effort:** average magnitude of parameter changes applied.
- **Mean number of variable changes:** number of controllable variables modified per recommendation.

These metrics allow evaluation not only of whether a model can improve predictions, but also of how aggressively it does so and how consistent the recommendations are giving us a good idea on the strength of the used models.

5.4.3 Control Recommendation Results

This subsection presents the quantitative results obtained from the offline evaluation of the control recommendation framework described in Chapter 4. For each dryer, a fixed number of representative operating points from the test period were selected and control recommendations were generated independently using both the ANN based and LSTM based predictive models. Table 5.7 summarizes the aggregated performance metrics obtained for each modeling approach. These metrics include the mean and median predicted improvement relative to the target temperature, the percentage of successful recommendations within the defined tolerance band, the average control effort and the average number of controllable variables modified per recommendation.

Model	Dryer	Mean Improvement	Median Improvement	Success Rate (%)	Mean Effort	Mean Variables Changed
ANN	Dryer 1	4.10	3.42	38.5	11.76	12.0
LSTM (W = 30)	Dryer 1	0.22	0.15	0.0	8.94	9.6
ANN	Dryer 2	5.69	7.20	35.5	11.57	12.0
LSTM (W = 30)	Dryer 2	0.93	0.16	0.0	10.96	11.1

Table 5.7: Summary of offline control recommendation performance

The values reported in Table 5.7 are computed from real industrial data. Variable identities have been anonymized to comply with confidentiality constraints imposed by the industrial partner. It should be noted that the zero success rate observed for the LSTM based recommendations results from the strict tolerance criterion adopted in this study. Although the LSTM model produces consistent directional improvements, these are generally smaller and do not reach the tolerance threshold defined for success At a purely numerical level, the ANN based framework achieves higher average predicted improvements compared to the LSTM based approach for both dryers. However, this improvement is accompanied by lower success rates and higher control effort. In contrast, the LSTM based framework consistently produces smaller predicted improvements but with greater stability across operating points. A deeper interpretation of these results is provided in the following subsections.

5.4.4 Interpretation of Improvement and Success Rate Metrics

While the quantitative metrics presented in Table 5.7 provide a numerical summary of the control recommendation performance, a deeper interpretation is required to understand the practical implications of these results. In particular, the relationship between predicted improvement, success rate, and control effort reveals fundamental differences between the ANN based and LSTM based modeling approaches. A control recommendation was considered successful if the predicted temperature after applying the recommended parameter changes fell within a tolerance band of $\pm$ 3 °C around the target temperature. In other words, success was achieved when the absolute deviation between the predicted temperature and the target value did not exceed 3 °C. This tolerance band was selected to reflect realistic industrial acceptability, accounting for sensor uncertainty, natural process variability and the offline nature of the evaluation framework. The ANN based control framework consistently achieved higher mean and median predicted improvements for both dryers. These results indicate that the ANN model is highly sensitive to changes in controllable variables and is capable of generating large actions to correct the expected values in a single step. However, this apparent advantage is accompanied by relatively low success rates. In many cases, the ANN based recommendations overshoot the target temperature, producing predicted values that fall outside the defined tolerance band despite exhibiting large absolute improvements. This behavior can be attributed to the static formulation of the ANN model. Because the ANN operates exclusively on instantaneous process snapshots, it lacks awareness of the recent evolution of the system and does not account for process inertia or delayed effects, that would be present in a dynamic model. As a result of this, it frequently overcompensates aggressively by modifying several variables at the same time, which increase the likelihood of overcorrection. In contrast, the LSTM based control framework showed substantially smaller predicted improvements but it was capable of a more stable and consistent behavior across evaluated operating points. Although the success rate of the LSTM based recommendations was zero under the strict ± 3 °C tolerance criteria, this result does not indicate a failure of the model. Instead, it should be analyzed and interpreted as the conservative nature of the LSTM predictions, which prioritize gradual adjustments and continuity with recent process history over large instantaneous corrections. The temporal structure of the LSTM model acts as an implicit smoothing mechanism. By incorporating information from previous time steps, the model naturally dampens abrupt responses and limits the magnitude of recommended changes. Consequently, LSTM based recommendations often move the predicted temperature in the correct direction but fail to reach the tolerance threshold within a single adjustment step. From an industrial point of view, this trade off is extremely relevant. Large, aggressive control actions may achieve faster convergence but carry a higher risk of instability, excessive actuator usage or product quality degradation and waste. On the other hand, conservative adjustments may require multiple iterations but offer improved robustness and predictability. In practical deployments, such behavior could be combined with iterative control strategies to progressively reach the target while respecting operational constraints. Overall, the results highlight that higher numerical improvement does not necessarily imply better control performance. When stability, interpretability and operational feasibility are considered, the LSTM based framework shows characteristics that are more aligned with advisory decision support systems, while the ANN based framework is better suited for sensitivity exploration and short term corrective analysis.

5.4.5 Analysis of Control Effort and Recommendation Complexity

Beyond predictive improvement and success rate, the practical feasibility of control recommendations depends strongly on the required control effort and the complexity of the suggested actions. In this work, control effort is quantified through two complementary indicators: the average magnitude of parameter changes and the average number of controllable variables modified per recommendation. The results reported in Table 5.7 show that the ANN based control framework consistently requires higher control effort and changes a larger number of variables per recommendation for both dryers. On average, ANN based recommendations consist of changes to nearly all available controllable variables, with relatively large adjustment magnitudes. This behavior demonstrates that the model is extremely sensitive to instantaneous input variations and the tendency it has to overcompensate aggressively for deviations from the defined target temperature. This characteristic is directly linked to the static nature of the ANN model. Since the ANN does not incorporate temporal context, it handles each operating point independently and assumes that all corrective action must be applied immediately in order to achieve the target temperature,leading to the model often distributing corrective effort across multiple actuators simultaneously, increasing both the magnitude and the dimensionality of the recommended intervention. On the other hand, the LSTM based control framework exhibits lower average control effort and consistently modifies fewer variables per recommendation. By integrating recent process history in its predictions, the LSTM model silently takes ongoing trends and previously applied effects into account. This allows for a more specialized and moderate set of adjustments from the model, with it often focusing on a smaller subset of variables that have more influence in the current temporal context rather than trying to adjust everything at once at the same time. From an operational point of view, recommendations that involve less variables and smaller changes are generally easier to interpret, validate and implement. Industrial drying systems are subject to mechanical, safety and/or operational policy constraints that restrict frequent or large actuator movements. In this sense, the reduced control effort observed for the LSTM based framework aligns more closely with practical industrial requirements, even if the immediate predicted improvements are smaller.

5.4.6 Stability, Consistency and Industrial Relevance of Control Recommendations

In addition to numerical performance metrics, the stability and consistency of control recommendations have been proven to be critical factors when assessing the suitability for them to be industrially applied. Highly variable or wrong recommendations can sabotage operator trust and increase the risk of unintended process behavior. The ANN based control framework show substantial variability across the evaluated operating points. While some recommendations yield large predicted improvements, others result in excessive corrections or even fail to meet the wanted tolerance criteria that was implemented entirely. This variability is completely normal and a direct consequence of the models static nature, which does not take into account the dynamic nature of this industrial. In contrast, the LSTM based framework presents a significantly more consistent behavior across all the evaluated test instances. Although its recommendations are more conservative and often insufficient to reach the target within a single step, they are directionally coherent and less sensitive to small perturbations in input variables. This consistency comes from the model ability to integrate information over time and to capture delayed system responses. In the context of industrial wood drying, where process inertia, thermal lag and material variability play a dominant and important role, such consistency is not only highly desirable but essential. Drying operations usually prioritize product quality, equipment longevity, process safety over and less material wastage as well as lower environmental impact over rapidly reaching a target value. From this perspective, stable and predictable recommendations may be preferable to aggressive corrections that risk overshooting or oscillatory behavior. It is also important to emphasize that the control recommendation framework evaluated in this study is intentionally simplified and operates entirely offline. No actuator constraints, rate limits or closed loop feedback mechanisms were enforced. As such, the generated recommendations should be interpreted as advisory insights rather than actionable control commands. Nevertheless, the observed differences between ANN based and LSTM based behavior provide valuable guidance for future control oriented model design. Overall, these findings suggest that temporal, dynamic models like LSTM networks offer a more suitable foundation for advisory or decision support systems in industrial drying environments. However, static models like ANN still remain valuable for exploratory analysis and sensitivity assessment but may require additional constraints or supervisory logic in order to guarantee a more stable control behavior in practical deployments.

5.4.7 Summary of Control-Oriented Modeling Insights

The evaluation of the offline control recommendation framework highlights fundamental differences between static and dynamic modeling approaches when applied to industrial wood drying systems. While both ANN based and LSTM based predictive models are capable of capturing meaningful relationships between process variables and temperature evolution, their behavior in control oriented scenarios differs substantially. The ANN based framework shows a strong ability to generate large corrective actions, resulting in higher average predicted improvements, however, this behavior is also linked to increased variability, higher control effort and overshooting the target temperature becomes a really probable reality. These characteristics reflect the limitations of static snapshot based modeling when applied to dynamic industrial processes with significant inertia and delayed responses. On the other hand, the LSTM based framework shows more conservative but consistent behavior. By taking recent process history into account, the LSTM model produces smoother and more stable recommendations that align better with the gradual nature of industrial drying dynamics and all of the industry constraints. Although these recommendations almost never achieve the target temperature in a single step under strict tolerance targets, they provide directionally coherent guidance that could be effectively exploited in iterative or supervisory control strategies. It is also important to mention that the results highlight that predictive accuracy alone is not enough for assessing control suitability. Metrics such as stability, control effort and recommendation complexity play a crucial role in determining if a model can support practical decision making in industrial environments. With this perspective in mind, temporal models typically offer advantages that extend beyond raw prediction performance. All in all, the findings of this section support the use of temporal machine learning models as a promising foundation for data driven advisory systems in industrial drying applications while at the same time, highlighting the continued relevance of static models for sensitivity analysis and exploratory assessment. These insights support the conclusions presented in the next chapter and provide clear directions for future research toward closed loop and real time control integration.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis investigated the application of data driven machine learning techniques to the modeling and analysis of an industrial wood drying process using historical operational data provided by an industrial partner. The primary objective was to assess the suitability of static and temporal predictive models for representing drying behavior and supporting and information systems focused on control oriented analysis under realistic industrial constraints. Due to practical limitations in the availability, resolution and consistency of direct MC measurements, temperature was adopted as a proxy variable for the drying state of the system. This decision was guided by operational expertise and supported by established drying theory, which recognizes temperature evolution as a key indicator of MC removal dynamics in kiln based drying processes. The use of temperature enabled high resolution modeling while avoiding the uncertainty introduced by sparse and inconsistently logged MC data. Two predictive modeling approaches were implemented and evaluated: a feedforward ANN and a LSTM network. The ANN model was used as a baseline static nonlinear regression approach, while the LSTM model explicitly incorporated temporal dependencies through sequential input representations. Both models were trained and evaluated offline using chronologically split datasets to ensure a realistic assessment of predictive generalization. The results showed that while the ANN model achieved strong performance during training, its predictive accuracy degraded when evaluated on unseen operating periods. This behavior reflects the limitations of static models in capturing the dynamic and time dependent nature of industrial drying processes. In contrast, the LSTM model demonstrated improved generalization performance, highlighting the importance of incorporating temporal context when modeling drying dynamics. Beyond prediction, an offline control recommendation framework was developed to explore how predictive models respond to variations in controllable process variables. This framework was not intended for real time deployment or closed loop control, but rather to evaluate model sensitivity, stability, and interpretability in a control oriented context. The ANN based framework produced larger predicted improvements but showed higher variability and required greater control effort. The LSTM based framework generated more conservative and consistent recommendations, reflecting its temporal smoothing behavior and improved stability. Overall, the findings of this thesis suggest that temporal machine learning models provide a more robust foundation for advisory and decision support analysis in industrial wood drying applications. While static models remain useful for exploratory analysis and sensitivity assessment, their limitations in dynamic and control oriented scenarios must be carefully considered. The work presented establishes a structured basis for future research toward more advanced, real time, and closed loop control strategies.

6.2 Main Contributions

The main contributions of this thesis can be summarized as follows:

1. **Application of data driven modeling to an industrial wood drying process**
   This work presents a structured application of machine learning techniques to a real industrial wood drying system using historical operational data. Unlike laboratory scale data or simulated studies, the dataset reflects realistic production conditions, including noise, variability and incomplete measurements, thereby increasing the practical relevance of the analysis.

2. **Justified use of temperature as a proxy for MC**
   Due to limited availability and low temporal resolution of direct MC measurements, temperature was adopted as a proxy variable for the drying state. This decision was made based in both industrial expertise and established drying theory and it enabled high resolution modeling while avoiding the introduction of uncertainty associated with incorrect and less moisture measurements.

3. **Comparative analysis of static and temporal predictive models**
   Two different machine learning models were implemented and compared: a feedforward ANN and a LSTM network. By evaluating both models under consistent preprocessing, feature selection and data splitting strategies, the

study provides a clear insight into the strengths and weaknesses of static versus temporal modeling approaches for industrial drying processes.

4. **Implementation of permutation based feature relevance analysis for industrial data**
   A permutation based feature relevance methodology was applied to both ANN and LSTM models to identify influential process variables in nonlinear and time dependent contexts. This analysis supported informed feature selection and demonstrated differences in how static and temporal models distribute relevance across process variables.

5. **Development of an offline control recommendation framework**
   An offline control recommendation framework was designed to evaluate how predictive models respond to changes in controllable process variables. Although not intended for real time or closed loop control, this framework enabled systematic comparison of model sensitivity, stability, control effort and recommendation complexity.

6. **Quantitative and qualitative evaluation of control-oriented behavior**
   The study provided both numerical metrics and qualitative interpretation of control recommendations, highlighting trade offs between predicted improvement, stability and control effort. This analysis proved that higher predictive sensitivity does not necessarily translate into better control suitability, particularly in dynamic industrial environments.

7. **Methodological transparency and reproducibility**
   All stages of the modeling pipeline, including data preprocessing, feature selection, model training, evaluation and control analysis, were implemented using transparent and reproducible procedures. Despite confidentiality constraints on variable naming, the methodology itself remains fully reproducible and transferable to similar industrial drying systems.

6.3 Limitations

Despite the positive findings and insights obtained in this study, there are several limitations that must be acknowledged. These limitations are essentially associated with data availability, modeling scope and the offline nature of the implemented framework. First, the analysis was based exclusively on historical industrial data, and no real time data acquisition or online system integration was performed. As a result, the predictive models and control recommendation framework were evaluated entirely offline. While this approach made sure that reproducibility was possible and avoided interference with industrial operations, it does not account for real time execution constraints such as communication delays, actuator response times, or sensor update frequencies. Second, temperature was used as a proxy variable for MC rather than modeling MC directly. Although this choice was justified by industrial practice and supported by drying theory, it introduces an indirect relationship between the modeled variable and the primary drying objective. Direct moisture measurements, if available at higher temporal resolution and with consistent logging, could provide more precise modeling of the drying state. Third, the predictive models implemented in this work were designed as representative rather than optimized architectures. Extensive hyperparameter tuning, architecture search, or ensemble modeling was intentionally avoided in order to focus on comparative behavior between static and temporal modeling paradigms. Consequently, the reported performance does not represent the maximum achievable accuracy for either model type. Fourth, he control recommendation framework adopted in this study was intentionally simplified. Recommendations were generated using bounded perturbations of controllable variables without enforcing physical actuator constraints, rate limits or closed loop feedback. As a result, the recommendations should be interpreted as advisory insights rather than directly executable control actions. Additionally, the evaluation of control recommendations was made using a single step adjustment strategy. This approach does not capture the cumulative effects of repeated control actions or long term convergence behavior since more advanced strategies, such as iterative or receding horizon control, were outside the scope of this work. Finally, confidentiality constraints imposed by the industrial partner limited the presentation of detailed variable descriptions and sensor identifiers. While this does not affect the validity or reproducibility of this work, it restricts the level of process specific interpretation that can be presented in this document. Overall, these limitations define the scope of the present study and provide a clear context for the interpretation of the results. Addressing these constraints offers multiple opportunities for future research and practical system enhancement, which are discussed in the following section.

6.4 Future Work

The results and insights obtained in this thesis open several directions for future research and system development aimed at enhancing the applicability of data driven approaches to industrial wood drying processes.

A natural extension of this work is the integration of real time data acquisition and online model execution. Deploying predictive models in a live industrial environment would enable continuous monitoring and adaptive decision support. This integration would require addressing practical challenges related to sensor latency, data synchronization and computational constraints as well as making sure that robustness is still present, even under changing operating conditions.

Future work should also focus on direct modeling of MC, with the constraint of being necessary that higher resolution and more consistently logged MC measurements become available. The inclusion of reliable MC sensors could allow the development of models that directly predict MC evolution, reducing reliance on proxy variables and improving process interpretability.

Another promising direction is the development of closed loop control strategies based on the predictive models explored in this study. More advanced control frameworks, such as Model Predictive Control or reinforcement learning based approaches, could exploit temporal models like LSTM networks to generate multi step control actions while explicitly enforcing actuator constraints, safety limits and process stability requirements.

The control recommendation framework could also be expanded in order to support iterative or receding horizon strategies, where recommendations are updated in a continuously way as new measurements become available. This would make gradual convergence toward target conditions while accounting for process inertia and delayed system responses possible and more feasible.

Finally, future work could be made, in order to, explore the incorporation of energy related metrics and sustainability indicators once reliable energy consumption data is available and provided. This would allow the evaluation of trade offs between product quality, process efficiency and energy usage, supporting more holistic optimization strategies that would aligned with industrial and environmental objectives and modern ideologies.

Overall, the control framework and findings studied and presented in this thesis provide a solid foundation that can allow continued research towards intelligent, data driven control and support systems for industrial wood drying, bridging the gap between offline analysis and practical, real time deployment.

References

[Đaković et al., 2023] Đaković, D., Kljajić, M., Milivojević, N., Doder, Đ., and Anđelković, A. S. (2023). Review of energy-related machine learning applications in drying processes. *Energies*, 17(1):224. [Cited on page 18]

[Bank et al., 2023] Bank, D., Koenigstein, N., and Giryes, R. (2023). Autoencoders. *Machine learning for data science handbook: data mining and knowledge discovery handbook*, pages 353–374. [Cited on page 17]

[Breiman, 2001] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32. [Cited on page 25]

[Chai and Li, 2023] Chai, H. and Li, L. (2023). Prediction of wood drying process based on artificial neural network. *BioResources*, 18(4):8212. [Cited on page 17]

[da Silva Ferreira et al., 2024] da Silva Ferreira, M. V., Ahmed, M. W., Oliveira, M., Sarang, S., Ramsay, S., Liu, X., Malvandi, A., Lee, Y., and Kamruzzaman, M. (2024). Ai-enabled optical sensing for smart and precision food drying: Techniques, applications and future directions. *Food Engineering Reviews*, pages 1–29. [Cited on pages 6 and 16]

[Haykin, 1998] Haykin, S. (1998). *Neural networks: a comprehensive foundation*. Prentice Hall PTR. [Cited on page 17]

[Hochreiter and Schmidhuber, 1997] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780. [Cited on page 17]

[Keey et al., 2012] Keey, R. B., Langrish, T. A., and Walker, J. C. (2012). *Kiln-drying of lumber*. Springer Science & Business Media. [Cited on pages 4, 13, 15, and 24]

[Laboratory, 1999] Laboratory, F. P. (1999). *Wood handbook: wood as an engineering material*. [Cited on pages 3, 13, and 15]

[MagalhÃčes et al., 2021] MagalhÃčes, R., Nogueira, B., Costa, S., Paiva, N., Ferra, J. M., MagalhÃčes, F. D., Martins, J., and Carvalho, L. H. (2021). Effect of panel moisture content on internal bond strength and thickness swelling of medium density fiberboard. *Polymers*, 13(1). [Cited on pages 4, 13, and 15]

[Malhotra et al., 2015] Malhotra, P., Vig, L., Shroff, G., Agarwal, P., et al. (2015). Long short term memory networks for anomaly detection in time series. In *Esann*, volume 2015, page 89. [Cited on page 17]

[Rahimi and Avramidis, 2022] Rahimi, S. and Avramidis, S. (2022). Predicting moisture content in kiln dried timbers using machine learning. *European Journal of Wood and Wood Products*, 80(3):681–692. [Cited on pages 3, 15, and 16]

[Rahimi et al., 2024] Rahimi, S., Avramidis, S., Sassani, F., and Nasir, V. (2024). Machine learning-based prediction of internal moisture variation in kiln-dried timber. *Wood Material Science & Engineering*, 19(2):499–510. [Cited on page 16]

[scikit-learn developers, 2024] scikit-learn developers (2024). Permutation feature importance. `https://scikit-learn.org/stable/modules/permutation_importance.html`. [Cited on page 25]

[UNECE, 2023] UNECE (2023). The future of the forestry and wood-based industries sector in the transition to a sustainable bioeconomy unece/fao regional forest outlook study 2020-2040. [Cited on page 4]

[Wang et al., 2023] Wang, T., Ng, W. W., Zhang, M., Zhang, X., Zhang, J., and Deng, M. (2023). Moisture content prediction of sugi wood drying using deep lstm ae minimizing perturbed error. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3756–3762. IEEE. [Cited on pages 6, 16, and 17]

[Wu and Avramidis, 2006] Wu, H. and Avramidis, S. (2006). Prediction of timber kiln drying rates by neural networks. *Drying Technology*, 24(12):1541–1545. [Cited on pages 15 and 24]

[Yadav et al., 2021] Yadav, R., Omrani, A., Link, G., Vauhkonen, M., and Lähivaara, T. (2021). Microwave tomography using neural networks for its application in an industrial microwave drying system. *Sensors*, 21(20):6919. [Cited on page 16]

[Zhang et al., 2025] Zhang, J., Tan, J., Song, Q., Du, T., Hauch, J., and Brabec, C. J. (2025). Feature selection for machine learning-driven accelerated discovery and optimization in emerging photovoltaics: A review. *Advanced Intelligent Discovery*, n/a(n/a):202500022. [Cited on page 25]

[Zhang et al., 2024] Zhang, Y., Fang, Z., Li, C., and Li, C. (2024). Deep-learning-based model predictive control of an industrial-scale multistate counter-flow paddy drying process. *Foods*, 13(1). [Cited on page 16]

[ÄŘakoviÄĞ et al., 2024] ÄŘakoviÄĞ, D., KljajiÄĞ, M., MilivojeviÄĞ, N., Doder, Ä., and AnÄŚelkoviÄĞ, A. S. (2024). Review of energy-related machine learning applications in drying processes. *Energies*, 17(1). [Cited on page 16]