Computers & Industrial Engineering 209 (2025) 111410

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Enhancing process lead time forecasting with machine learning and upstream process data: A case study in wind tower manufacturing

Kenny-Jesús Flores-Huamán [a], Antonio Lorenzo-Espejo [a,*], María-Luisa Muñoz-Díaz [b], Alejandro Escudero-Santana [a,b,**]

[a] Departamento de Organización Industrial y Gestión de Empresas II, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Cm. de los Descubrimientos, s/n, 41092 Seville, Spain

[b] Centro de Innovación Universitario de Andalucía, Alentejo y Algarve (CIU3A), 41011 Seville, Spain

ARTICLE INFO

Keywords:
Lead time prediction
Machine learning
Production planning and control
SHAP
Model interpretability
Wind tower manufacturing
Comparative analysis

ABSTRACT

Accurate lead time prediction is critical for optimizing sequential manufacturing processes, particularly in industries with high variability such as wind turbine tower production. This paper proposes a machine learning-based system to estimate lead times for two pivotal sequential operations: bending and longitudinal welding (LW). A distinctive feature of this system is its innovative integration strategy, where the predictive output from the bending model, specifically, the predicted bending lead time and its associated error, is leveraged as an input feature for the LW lead time estimation model. This approach explicitly models and enhances the representation of inter-process dependencies. While bending predictions show moderate performance, their inclusion as inputs demonstrably and significantly improves LW lead time estimation accuracy. A key contribution of this work is the comparative analysis between the ML-based LW predictions and traditional engineering methods. Our results demonstrate that the integrated ML model for LW achieves a 54% reduction in MAE (from 11.36 to 2.03 h) and a 74% lower RMSE (from 12.01 to 3.13 h) compared to engineering estimates, validating its superior accuracy. To enhance interpretability, SHAP (SHapley Additive Explanations) identifies critical factors such as sheet thickness, personnel experience, and upstream process quality, including the impact of the integrated bending predictions. The system's low execution time enables real-time scheduling adjustments, offering a practical solution for production planning. These findings highlight the transformative potential of ML, particularly through such sequential predictive integration, in replacing outdated engineering heuristics and providing actionable insights for complex manufacturing environments.

1. Introduction

According to the Council (2024), a record 117 GW of new capacity was installed in 2023, marking the best year ever for new wind power. Furthermore, it was a year of continued global growth, with 54 countries across all continents contributing to the expansion of wind power. Spanish households will face significant electricity price increases in 2025, driven by VAT adjustments and fixed cost hikes. The VAT (Value Added Tax) increased to 21% in January 2025, increased the temporary reductions implemented during the energy crisis in 2021 (Menendez-Roche, 2025). Additionally, the ever-growing global energy demand has spurred technical development in the wind power industry, which has focused on increasing power output through better exploitation of wind currents.

To achieve this, wind turbine manufacturers are exploring two main strategies: (a) raising turbines to higher altitudes, where wind currents are more consistent, and (b) using larger rotor blades to capture more wind energy. However, both approaches require larger and taller wind turbine towers, posing significant technical challenges. The manufacturing of these towers is already a complex, labour-intensive process, and the increasing size requirements further complicate operations and supply chain management.

In this context, enhancing the predictive capabilities of production planning systems is fundamental, particularly in manufacturing environments characterized by high variability and strong operational interdependencies. Within any supply chain, the production system comprises a series of interdependent operations whose variability in execution and level of coordination directly affect overall efficiency

* Corresponding author.
** Corresponding author at: Departamento de Organización Industrial y Gestión de Empresas II, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Cm. de los Descubrimientos, s/n, 41092 Seville, Spain.
E-mail addresses: alorenzo@us.es (A. Lorenzo-Espejo), alejandroescudero@us.es (A. Escudero-Santana).

https://doi.org/10.1016/j.cie.2025.111410
Received 19 March 2025; Received in revised form 17 July 2025; Accepted 19 July 2025
Available online 31 July 2025

and delivery times. These operations often include both sequential and parallel activities, where delays or quality deviations in the early stages can propagate to subsequent ones, amplifying their impact in later phases.

In wind turbine tower manufacturing, sequential processes such as bending and longitudinal welding (LW) are particularly critical, and their inherent variability significantly impacts overall production planning. A review of the literature on lead time prediction using machine learning in manufacturing contexts reveals a predominant trend: most studies focus on modelling and predicting the lead time of individual processes in isolation (Bender et al., 2022; Flores-Huamán et al., 2024; Gyulai, Pfeiffer, Bergmann, & Gallina, 2018; Gyulai et al., 2018; Kang et al., 2020; Lingitz et al., 2018; Lorenzo-Espejo et al., 2022; Onaran & Yanı k, 2020; Pfeiffer et al., 2016). While such methods provide additional sorts of explanatory variables across different stages of the process, the dynamic integration of sequential predictive models – in which the output of a model for an earlier stage (e.g., bending), together with its associated uncertainty or error, is directly incorporated as input to improve predictions in a subsequent stage (such as longitudinal welding) – remains a significantly under-explored area.

In this regard, the work by Lorenzo-Espejo et al. (2024) stands out as one of the few efforts to explicitly model the sequential nature of production processes. Nevertheless, even in that study, advanced interpretability techniques were not employed to clarify the rationale behind the model's predictions, an essential aspect for its effective adoption in industrial decision-making. Additionally, the absence of a systematic comparison with traditional approaches limits the ability to fully evaluate the tangible benefits of the proposed sequential modelling framework.

To address this gap and enhance the accuracy of lead time prediction in sequential processes, this article presents a novel machine learning (ML) based system for estimating the lead times of bending and longitudinal welding operations in wind turbine tower manufacturing. The effectiveness of the proposed approach is evaluated through a case study at a Spanish wind turbine tower manufacturing plant, utilizing production records collected between 2022 and 2024.

The main contribution and novelty of this work lie in its architecture of sequential prediction integration: a system is designed and implemented where the predicted lead time from the bending process, as well as an estimate of its prediction error, can be employed as input features for the prediction model of the longitudinal welding process. The specific contributions of this study are:

- Development and evaluation of a sequentially integrated predictive system: we propose and evaluate a framework in which ML models for consecutive processes are interconnected, allowing predictive information, not just historical data, to flow from early to later stages to improve overall estimation.
- Empirical evidence of improved downstream accuracy: we provide quantitative evidence that incorporating predictions from the bending model significantly enhances the accuracy of the longitudinal welding model. This improvement is observed both when compared to models that exclude this integrated predictive information and relative to traditional engineering estimation methods used at the studied plant.
- Interpretability analysis: we utilize SHAP (SHapley Additive Explanations) to identify the most influential factors in LW predictions, including the specific impact of the integrated bending process predictions.

The system comprises two main ML regression models, one for each operation (bending and LW). These models are designed to capture complex, non-linear relationships between input variables, which traditional regression techniques often fail to identify. The bending model is fed with: historical lead time data for each process and its upstream operations; context information such as personnel, machines, and product types; and quality control data for raw materials and semi-processed parts. The LW model incorporates the output of the previous bending model along with additional relevant data. Additionally, we employ SHAP (SHapley Additive Explanations) to enhance model interpretability, allowing us to identify key factors influencing lead time predictions and providing actionable insights for decision-making (Lundberg & Lee, 2017).

The approach is assessed with different configurations of the LW ML regression model. These include using the predicted lead time values from the bending model, incorporating the prediction error, using only the actual bending lead time observations, and excluding any information about the bending lead time. The alternatives are compared and evaluated, with the goal of minimizing prediction errors for LW times, particularly aiming to improve the accuracy of extreme lead time values. Our results demonstrate that while bending lead time predictions exhibit moderate accuracy, incorporating them as inputs significantly improves LW lead time estimation. This highlights the importance of leveraging upstream process data to enhance downstream predictions.

The remainder of this paper is structured as follows: Section 2 presents a description of the wind turbine tower manufacturing process, along with the distinctive constraints and characteristics of its production planning and control, as well as a brief review of the relevant literature. Section 3 details the proposed system and methodology, including data preprocessing, feature selection, and model implementation. Section 4 presents and discusses the results of applying the proposed approach to the case study of a Spanish wind turbine tower manufacturer, with a focus on model performance and interpretability. Section 5 provides a comprehensive discussion of these results, contextualizing them with the existing literature, exploring their industrial implications and limitations, and proposing directions for future research. Section 6 summarizes the conclusions of the study, highlighting the practical implications of our findings for production planning and control, and, finally, the references in this paper are listed.

2. Wind turbine tower manufacturing: Background and applications of machine learning

Wind turbines are large-scale devices that convert the kinetic energy of the wind into electrical energy. The most common types are installed either onshore or offshore and consist of four main components: the rotor, the generator, the yaw system, and the tower. The rotor spins due to the wind's forces acting on its blades, and the kinetic energy from this motion is converted into electrical energy by the generator. The yaw system rotates the generator and rotor around a vertical axis to face the wind direction. Finally, the towers, which are the focus of this work, are steel structures that support the other three components.

Wind turbine towers are assembled on-site by joining large steel cylinders or conical frustums (sections) together. These sections are bolted to each other using flanges that have been previously attached to their top and bottom ends. The top flange of section n is bolted to the bottom flange of section n+1. Wind towers are composed of at least three sections: a bottom, a mid and a top section. When higher towers are required, more mid sections are installed.

These sections are built in wind tower manufacturing plants and transported to the wind-farm location. The sections are assembled in the plant using ferrules, smaller cylinders or conical frustums that are welded together. Previously, the ferrules are formed by bending steel plates into rings, which are then welded together to form a closed conical frustum or cylinder. The production process of a wind tower involves several stages, as shown in Fig. 1, which illustrates the different states of the tower assembly. Succinctly, the operations involved in the process are the following:

1. **Plate cutting and bevelling:** the plate cutting process involves cutting raw steel to obtain sheets of the required size to form

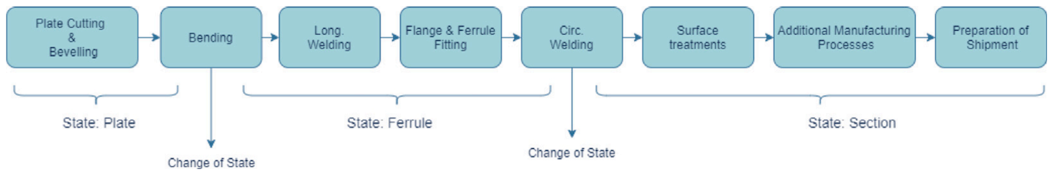


Fig. 1. Production process of a wind turbine tower, distinguishing between its various states.

the cylinders, using techniques such as plasma cutting or oxy-fuel cutting. Then, the edges are bevelled at an angle other than 90 degrees to ensure a stronger and higher-quality weld before joining the pieces.

2. **Bending:** rectangular steel plates are bent into cylinders or conical frustums.
3. **LW:** the edges of the bent plate are welded to one another in order to form a fully closed ferrule.
4. **Flange fitting:** flanges are fitted to the inferior and superior ferrules of the sections and given several weld spots so that they hold their position.
5. **Ferrule fitting:** the ferrules that compose a section are fitted to each other and given multiple weld spots to ensure that they hold their position.
6. **Circular welding:** the fitted ferrules and flanges are finally welded together, following the weld spots given in the fitting process.
7. **Surface treatment:** the sections then go through a series of processes that prepare the internal and external surfaces for the conditions they must endure during service.

A more exhaustive depiction of the wind turbine manufacturing processes and their recent advancements is provided by Sainz (2015). However, in this article the focus is set on the bending and LW processes. These two operations are amongst the most intricate of the manufacturing process. Firstly, they are the two initial major processes in the manufacturing system (for technical reasons). If, owing to the configuration of the system, any of these two operations constitute the bottleneck of the process, a great planning effort must be performed in order to ensure that there is a continuous flow in the corresponding workstations. In the case that this is not achieved, non-desirable idle times could be expected in the many posterior processes.

Also owing to the position of the processes in the workflow, reworks due to major faults occurred during bending and LW are time-consuming and costly. If a bending- or LW-related defect is identified in any of the downstream processes, the part must be carried back to the start of the production layout. This is not an easy endeavour, since, due to the size and weight of the parts, the layout is optimized in order to allow the process to be completed with as little movement of the part as possible but, evidently, in the natural flow of production. Adding to that commented above, bending and LW-related faults cause costly reworks also for the following reasons: firstly, in order to resume production of the impending tower as soon as possible (towers can deteriorate and deform inside the production line due to their weight), the faulty ferrule is assigned the highest priority at the bending or LW station. If a different model of ferrule is currently in production in the stations, a setup time for tool or configuration modifications can be expected. Furthermore, the defective ferrule cannot simply be replaced by another ferrule, as most of them have different product specifications and designs.

Moreover, if the defect is found after the fit-up process, the whole production of the tower must be set to a stop, since by then the ferrules are welded to one another. Of course, this also implies an extra rework time, as the defective ferrule must be separated from the rest of the tower.

Finally, minor defects, which are less frequently detected on time by employees, can significantly reduce the performance of the downstream workstations. It must be borne in mind that the fit-up process unites two ferrules, which must match with very low tolerances in order to ensure the structural integrity of the tower. If a ferrule is not perfectly curved or if its lips do not come to a perfect union, the complexity of the fit-up process is severely increased. Additionally, the lead time for the circular welding process may increase, as larger welds are often needed to compensate for these imperfections.

2.1. Production planning and control in wind turbine tower manufacturing

Wind turbine tower manufacturing is a challenging production process from a production and control standpoints, for several reasons: (a) the raw materials and products are voluminous and heavy; (b) as a consequence of the volume and weight of the parts, it is a mainly non-automated manufacturing process; (c) despite being a low-volume production process, there is a strong variability between client orders; and (d) in spite of the size of the parts produced, wind turbine towers are subject to very strict regulations and small tolerances.

In this context, recent research has explored advanced scheduling methods for unrelated parallel machines, where processing times depend on both the machine and the job. A novel approach introduces support machines, which, despite having reduced capacity, can perform partial tasks before transferring them to main machines for completion. Muñoz-Díaz et al. (2024) formulated a Mixed-Integer Linear Programming (MILP) model to optimize this problem and evaluated Tabu Search, Simulated Annealing, and a Constructive Heuristic. Their results indicate that support machines can improve production efficiency, with Tabu Search achieving the best performance, while the Constructive Heuristic offers a faster alternative for real-time applications.

However, the implementation of these advanced scheduling methods in wind turbine tower manufacturing faces significant obstacles due to the lack of sensorization and digitization in many plants. This challenge is also evident in the manufacturing plant studied in this paper, where manual data collection has several implications for process planning and control. Firstly, a considerable amount of employee effort is required to record production data, which is often of poorer quality than sensor-generated data. The absence of a standardized protocol, or non-adherence to it, introduces bias and errors into the manufacturing records. It must be noted that workers often view data recording as a secondary task, sometimes performing it under less-than-ideal conditions.

In particular, the process lead time variable is likely the most affected by errors in manual recording. In the case of the plant studied in this paper, employees had to move from their workstations to fill in the completion time of a part and then return to their post to resume the operation. This led to them forgetting to fill in these records or even waiting until the end of their shift. Therefore, these circumstances undoubtedly affect the quality of the lead time data, which, in turn, has a significant effect on lead time forecasting accuracy. Simply using the averages of the lead times for these processes is bound to produce inaccurate predictions. Pérez-Cubero and Poler (2020) emphasized the importance of considering lead time variability in job-shop production scheduling. Thus, other determinant factors of the lead time must be utilized in order to generate precise estimations that, if good enough, may serve as input for the production planning and control of the manufacturing process.

Accurate lead time predictions are particularly crucial for two main applications: job scheduling and anomaly control. Regarding job scheduling, if both efficient and attainable schedules are to be produced, it is essential that the lead times of each job are accurately represented. If the time slots allocated to a job are lengthier than what is actually needed, the workstation will most likely experience inefficient idle time. On the other hand, if the schedule includes less time than required to complete the process, upstream stock levels are likely to increase, and more importantly, there is a risk that product delivery dates will not be fulfilled. In addition, accurate lead time predictions can enable anomaly control systems in cases where sensor data (such as vibration, temperature, or noise records) are not available. In these instances, comparing the expected lead time with the actual processing time can serve as a warning of potential machine failures or defective parts.

2.2. ML applications to lead time prediction and wind power

A review of the production management literature reveals that there is only a limited number of works addressing the use of machine learning techniques for the prediction of process lead times. This stands in contrast to the more extensively studied problem of job-shop scheduling (JSSP), where the application of machine learning, particularly reinforcement learning (RL) and deep learning (DL), has received significant academic attention (Pérez-Cubero & Poler, 2020). While JSSP focuses on optimizing the sequence of operations, our work addresses the prerequisite challenge of accurately estimating the duration of those operations, a critical input for any effective scheduling system. This highlights a practical gap in the literature that our research aims to fill.

Along these lines, Kang et al. (2020) produce a systematic literature review in which they identify quality-related problems as those most frequently approached using ML techniques out of other less researched managerial aspects regarding production lines, such as yield improvement, preventive maintenance, waste management and, the topic of this article, lead time prediction. On their part, Bertolini et al. (2021) literature review of ML industrial applications does not even consider lead time prediction as a unique research topic inside production planning and control (PPC), but rather as an intermediate step of Performance Prediction and Optimization, Scheduling or Process Control solutions.

Usuga Cadavid et al. (2020) present an exhaustive literature review of industrial applications of ML-aided production planning and control (ML-PPC). The authors identify "time estimation" as an additional use case of ML-PPC, which was not previously considered in the revision of data-driven smart manufacturing applications made by Tao et al. (2018) .

Burggräf et al. (2020) conduct a systematic literature review of the approaches to lead time estimation in Engineer-To-Order (ETO) environments. The authors find that, in the sample of academic works used in their review, material and employee-related data are seldom used to produce the predictions: 5% and 0% of the 42 studies that they analyse include material and employee-related data, respectively.

Most of the articles found on this research line focus on completion time estimation (Alenezi et al., 2008; Backus et al., 2006; Kramer et al., 2020; Öztürk et al., 2006; Ruschel et al., 2021; Wang & Jiang, 2019). This trend is understandable, as the completion or total lead time, which in a manufacturing environment can be thought of as the interval between the arrival of a part and the fulfilment of all the operations required in its manufacturing specifications, is a key performance metric for many companies. This is particularly true in Make-To-Order (MTO) systems, as delivery dates must be previously agreed upon and then fulfilled to maintain customer satisfaction, trust, and loyalty. This also applies to resource-sharing departments and entities (Szaller & Kádár, 2021).

Several different approaches have been used to address completion time prediction through ML. They mostly vary in the methods and data sources used to generate the estimations. Mohsen et al. (2022) utilize diverse ML algorithms, namely linear regression, K-nearest neighbour, random forest and neural networks, to estimate the cycle time of an industrialized building manufacturer, using three groups of input variables: product specifications, real-time tracking data using RFID acquisition technologies and engineered features representing workload conditions. Modesti et al. (2022) compare the performance of empirical methods and artificial neural networks at the prediction of manufacturing flowtimes and completion due dates in job-shop settings.

Instead of focusing on manufacturing lead times, Steinberg et al. (2022) predict the possibility of manufactured parts experiencing delays at their arrival at an assembly station in an MTO environment. The authors evaluate six different ML models for classification with and without a set of variables corresponding to the design of the material. They conclude that the performance of the models with the material-related information is higher, but by a relatively low margin. This fact is attributed by the authors to the scarce variability of material designs. Along these lines, Lim et al. (2019) utilize Support Vector Machines to address completion time prediction as a classification task, discretizing lead time into multiple classes.

In comparison to completion time prediction, forecasting the lead times of individual processes adds the complexity of a lesser number of data sources from which to draw useful knowledge. In this article, this obstacle is tackled by gathering data from previous processes and connecting the prediction modules of sequential processes.

Other authors go a step beyond the completion time, focusing on transition or waiting times, which are, essentially, the periods that a part spends waiting or being transported between processes. For example, Schuh et al. (2018) posit a framework for the determination of transition times using data mining techniques with the goal of improving the adherence to delivery dates. According to the authors, transition times are often the cause of unsatisfied delivery times due to the lack of standardization, their high variability, and the simplification of its calculation. Additionally, Schuh, Gützlaff, Sauermann, and Theunissen (2020) present an approach to transition time prediction using time series data mining (TSDM), combining product specifications and organizational variables with historical data. Similarly, Gützlaff, Sauermann, Kaul, and Klein (2020), Schuh et al. (2019) utilize Regression Trees and Random Forests to forecast transition times, also determining the influence of several production-knowledge variables on the predictive power of the models.

Recently, the prediction of specific process lead time has started to gain attention from researchers. Unlike completion time prediction, being able to estimate the lead times of one or multiple processes can be directly applied to production planning and control and to scheduling. Along these lines, Gyulai, Pfeiffer, Bergmann, and Gallina (2018) develop a data analytics system that implements what they coin as "situation aware" production control. In their system, a closed-loop control is used to provide online updates for a digital data twin Gyulai et al. (2018). The digital twin is supported by process lead time predictions conducted using ML algorithms, which, according to Pfeiffer et al. (2016) and Lingitz et al. (2018), outperform traditional analytical techniques. Specifically, in the case study in which these works are supported (a semiconductor manufacturing process), the random forest method stands out among other algorithms for its performance. The system proposed by the authors is focused on real-time control of the lead times by using information about dynamic events occurring simultaneously (as well as product-specific data). Instead, the system proposed in this article focuses on short-term prediction, as the input variables are set in advance. Depending on the variables chosen for the model, which are discussed later in the article, the lead time predictions can be produced with varying levels of anticipation.

Bender et al. (2022) present two practical cases of application of three different automated ML (AutoML) frameworks and compare them to simple lead time prediction approaches used in the enterprises under study. AutoML aims to automate the complete ML pipeline, from data preprocessing to deployment. The authors address two Make-To-Order (MTO) manufacturing processes composed of many operations, for which they estimate the distinct process, but the results are shown aggregated in their study. The authors employ product and organizational-related variables as predictors, and the proposed systems only outperform the simple mean-based predictions in one of the two companies. While Bender, Trat and Ovtcharova endorse the value of AutoML, they highlight the need for holistic solutions that are able to fully support users in labour-intensive processes such as data understanding, transformation, filtering, preprocessing and feature engineering. Bender and Ovtcharova (2021) also present a prototype that integrates an Enterprise Resource Planner (ERP) to provide data, the AutoML software to produce lead time predictions and a Manufacturing Execution System (MES) to control the operation in the plant. Sousa et al. (2022) also utilize AutoML packages, but for order completion time prediction.

Zhu and Woo (2021) combine a new self-organizing hierarchical particle swarm algorithm (PSO) with a Support Vector Machine (SVM) prediction model in order to forecast the lead times of two production processes in the shipbuilding industry. Rizzuto et al. (2021) present a case study of the application of multiple ML models to predict the lead times of the tooling, placing and execution operations in a drilling factory. The results show that the random forest algorithm outperforms the rest of the methods used in their comparison for each of the three operations.

Finally, Onaran and Yanı k (2020) utilize the Multilayer Perceptron, one of the most frequently used neural networks, to predict the lead time of a manual-labour-intensive operation in a textile-manufacturer's production line. The authors employ product and order-related variables, as well as employee data and a measure of the efficiency of the complete line.

The studies mentioned above all present different systems or approaches to the prediction of the lead times of specific processes. However, their proposals do not collate the times of the operations with each other to evaluate potential improvements in the predictions, as posited in this article. Furthermore, another contribution of this article, based on the review of the extant literature, is utilizing the predictions of the lead times of a process to feed other process prediction systems.

Concerning the use case of the proposed prediction system, it must be noted that wind power has received significant research attention, but not regarding its manufacturing stage. A review of the literature presenting ML approaches to wind power settings reveals that most studies address the operational stage of wind power. Three main research fields can be identified in the literature:

- Smart maintenance systems for wind turbines, specifically condition-based monitoring. The three most common research lines on this topic are anomaly detection (Helbing & Ritter, 2018), fault classification (Gao et al., 2021) and remaining useful life (RUL) estimation (Carroll et al., 2019) – see Stetco et al. (2019) for an exhaustive review on this topic.
- Expert systems for wind turbine and wind farm design and control (Fischetti & Fraccaro, 2019; Petrov & Wessling, 2015).
- Prediction of power output, which can be based on multiple different input variables, such as historical output records (Treiber et al., 2016) or wind and weather conditions (Kim & Hur, 2021).

Noticeably, the only works that address wind turbines from a manufacturing standpoint are those by Sainz (2015), who describes the manufacturing process and several improvement steps based on an increased automation; Park (2018), who analyses composite wind turbine towers from a design and manufacturing standpoint; and Lorenzo-Espejo et al. (2022). In the latter, a machine learning-based approach to the bending process of wind turbine tower manufacturing is conducted, which highlights the influence of worker experience and age, given the manual character of the operation.

In addition, Flores-Huamán et al. (2024) present a machine learning-based approach to predict lead times for different operations in wind tower manufacturing. Their study, based on data collected from facilities in Spain and Brazil, evaluates nine regression algorithms, including Random Forest, XGBoost and LightGBM, as well as deep learning models such as TabNet and NODE. The results indicate that models based on Gradient Boosting are the most effective in predicting processing times and optimizing resource allocation, highlighting the importance of integrating ML into production planning in the wind tower industry.

Similarly, Rocha-Jácome et al. (2025) propose a non-contact measurement system using LiDAR sensors and ML techniques to predict geometric parameters in large-scale industrial components, specifically in wind tower manufacturing. Their approach combines geometric analysis with digital filtering and ML models to improve the accuracy of curvature radius measurements. Their results validate the system's effectiveness in real production environments, emphasizing its potential for optimizing manufacturing processes through ML.

Recent advancements in lead time prediction for manufacturing processes have been explored by Lorenzo-Espejo et al. (2024), who developed a machine learning-based system for predicting lead times in wind turbine tower manufacturing. Their system utilized sequential process data and achieved notable improvements in prediction accuracy, particularly for the longitudinal welding process. However, their approach faced limitations in the accuracy of bending process predictions, which showed moderate performance due to the high variability and manual nature of the operation. Additionally, while their system provided valuable insights, it lacked advanced interpretability techniques to explain the model's predictions, which is crucial for decision-making in industrial settings.

In this study, we build upon these developments by utilizing a different dataset, collected from a more recent production period (2022–2024), which includes updated operational parameters and a larger sample size. This allows us to validate and extend their findings under current production conditions. Unlike the previous work, which primarily relied on GB for predictions, we explored a broader range of machine learning approaches, including XGBoost, LightGBM, and neural network architectures such as MLP. This broader evaluation enables us to identify the most effective model for each process, significantly improving the accuracy of bending predictions, which was a key limitation in the previous study.

Furthermore, we address the lack of interpretability in the previous system by incorporating SHAP analysis. This technique provides detailed insights into the contribution of each input variable to the predictions, allowing production managers to understand the factors driving lead times and make more informed decisions.

These enhancements not only lead to more robust and accurate predictions but also provide actionable insights for production planning and control, particularly in optimizing resource allocation and identifying potential anomalies in the manufacturing process. By integrating these improvements, our system represents a significant advancement over the previous approach, offering a more comprehensive and interpretable solution for lead time prediction in wind turbine tower manufacturing.

Apart from the cited studies, no other contributions on wind turbine manufacturing and ML applications to such process are found in the literature.

3. Methodology

The methodology followed in this study is presented in this section. For conciseness, the steps are directly outlined as applied to the case study at hand. In particular, the system shown includes the bending and LW processes. However, the conceptual design of the system is applicable to any sequence of manufacturing processes, provided that a correlation between their lead times is expected. There are five main steps in the proposed regression analysis: data gathering; exploratory data analysis; data preprocessing; system design; and model implementation.

3.1. Data gathering

In this study, data are gathered from the manufacturing of nearly 900 tower sections produced between 2022 and 2024, each consisting of over 7,300 ferrules. The data are collected using the company's ERP (Enterprise Resource Planning) and QMS (Quality Management System) databases, capturing various aspects of the production process. To create a comprehensive dataset for analysis, information from these databases is carefully integrated. This collation process ensures that the final database encompasses the necessary variables for further exploration and study.

The data collection pipeline relies on manual entry by plant operators into the ERP and QMS systems at the conclusion of each manufacturing operation (e.g., bending, welding). This introduces a variable time lag between the actual completion of a task and its digital registration, typically ranging from a few hours to the end of a work shift. Consequently, the data frequency is tied to the completion rate of individual ferrules rather than a fixed time interval. While this process provides essential operational data, its manual nature is a known source of potential inaccuracies and delays, a challenge that this study's modelling approach is designed to accommodate.

The explanatory variables are selected based on an initial data exploration phase and subsequent discussions with plant personnel, aimed at identifying the factors that potentially influence operation completion times. This selection process results in the classification of variables into four main categories: historical lead time records from upstream processes, contextual information, quality control reports, and predictions generated by machine learning regression models.

Within these categories, the inclusion of specific contextual variables
–such as product specifications (e.g., nominal thickness, plate dimensions) and organizational attributes like the personnel assigned to the immediate process–is grounded in prior research on manufacturing lead times (Flores-Huamán et al., 2024; Lorenzo-Espejo et al., 2022). However, this study offers several novel contributions regarding the variables posited as potential determinants of process lead time:

- Explicit incorporation of historical lead time records from multiple upstream processes (e.g., sheet cutting, bevelling, bevel cleaning) as direct predictors of downstream operations (bending and longitudinal welding). While the influence of the immediately preceding step is sometimes considered in existing models, our approach systematically integrates a broader set of upstream performance data.
- Extension of organizational variables – including operator and machine identifiers – to also cover upstream processes. This is particularly innovative, as traditional models typically restrict input data to the process whose lead time is being predicted. Our rationale is that specific personnel or equipment used during upstream stages, such as bevelling, may directly affect the quality and characteristics of the intermediate product, thus influencing the lead time of subsequent processes like bending and welding. This allows us to capture critical inter-process dependencies that are often overlooked.
- Comprehensive integration of quality control reports from various inspection points, enabling the linkage of specific quality metrics to lead time variability.

The first three categories of variables are described in detail below. The fourth category consists of predictions generated by machine learning models trained on the bending stage, which are then used as inputs for predicting longitudinal welding times. This approach represents a key architectural innovation and is discussed in Section 4.

3.1.1. Historical lead time records of up-stream processes

The lead times of processes taking place before the bending and LW operations may serve as contributing predictors of the corresponding bending and LW process times. There are three main processes that precede the bending operation: sheet cutting, bevelling and bevel cleaning. There are not further significant operations between the bending and LW processes. Three hypotheses that could explain the potential correlation between an operation and its preceding processes can be posited: (a) since the dimensions of the parts are expected to greatly influence the lead time of the processes, it should be expected that taking longer to process a part at the, for example, bevelling station, could be correlated with a longer bending lead time; (b) long process times may be indicators of production anomalies or defective units/equipment. If undetected, these could extend downstream, increasing the lead times of coming processes; and (c) on the other hand, excessively short process times may be indicators of a poor-quality work. While this may not result in immediate defective units, it can show later along the production process. Therefore, while it is difficult to pinpoint a specific reason a priori, the correlation between the lead times of different processes seems reasonable and worth studying.

3.1.2. Context information

As previously discussed, when accurate sensor-based data are unavailable and the only information available is that recorded manually by the workers, it can be ill-advised to rely simply on historical lead times for the prediction. However, there are other variables referring to aspects of the process that are usually set in advance and involve less uncertainty. This category of variables is again split into two groups: product specifications and organizational variables. There are eight variables related to the product specifications a priori relevant to the lead time:

- The position of the section that contains the processed ferrule in the tower, a numeric variable ranging from 1 (bottom section) to 6 (highest section produced).
- The position of the ferrule in the section in which it is to be included, a numeric variable which can take a value from 1 (bottom ferrule of the section) to 16 (highest ferrule position).
- The yield strength of the steel with which the plate was manufactured, for a nominal thickness of 16 mm or less (355 N/mm^2 or 455 N/mm^2).
- The toughness subgrade of the steel with which the plate was formed, measured with the Charpy test (JR: 27 J of impact strength at 20 ° C; J0: 27 J at 0 ° C; J2: 27 J at −20 ° C; NL: 27 J at −50 ° C; and K2: 40 J at −20 ° C).
- Whether the steel plate has received a normalization treatment in order to increase its toughness or not.
- Nominal thickness, length, and width of the plate.

These variables are common for every process since they refer to product specifications. On the other hand, the organizational variables, the personnel, and machine variables, are particular to each of the processes. In a previous analysis (Lorenzo-Espejo et al., 2022), the bending lead time has been found to be significantly affected by which worker performed the operation. Therefore, it seems reasonable that the personnel and machine variables could also impact the lead times of the downstream operations. Thus, the models include this information not only for the bending and LW operations but also for the sheet cutting, bevelling, and bevel cleaning discussed above.

3.1.3. Quality control reports

The QMS module of the system provides information regarding the several quality inspections performed throughout the process. The quality reports available when the parts reach the bending and LW processes refer to the sheet inspections made at the receiving warehouse and after the sheet cutting, bevelling and bevel cleaning operations are

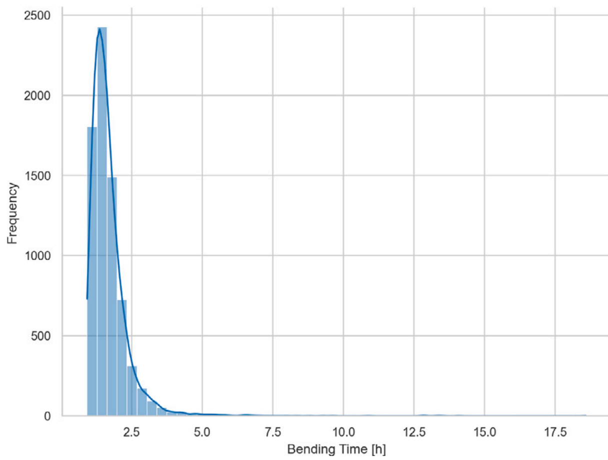


Fig. 2. Histogram of bending lead time.

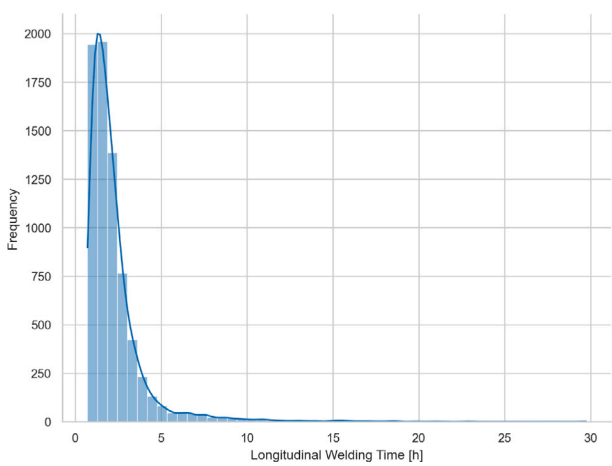


Fig. 3. Histogram of longitudinal welding lead time.

completed. The variables recorded in these quality inspections refer mostly to additional measures of the dimensions of the sheets. These are far more detailed and accurate than the nominal dimensions obtained from the ERP system. Furthermore, the conformity of the bevels with the product specifications is checked, as well as the sheet dimensions after the cutting and bevelling processes.

3.2. Exploratory data analysis

Following data gathering, an exploratory data analysis is conducted to understand the underlying data distributions and the explanatory power of selected variables in estimating manufacturing lead times. The analysis focuses on the characteristics of the target variables (Bending and LW lead times) and their relationship with key process and product features.

3.2.1. Lead time distribution analysis

Figs. 2 and 3 illustrate the distribution of lead times for the bending and longitudinal welding operations. As shown in the figures, both processes exhibit a right-skewed distribution, indicating that most operations are completed in a relatively short time, but there is a non-negligible proportion of cases where the required time is significantly longer. This asymmetry may be associated with variability in sheet thickness, process interruptions, or operational inefficiencies.

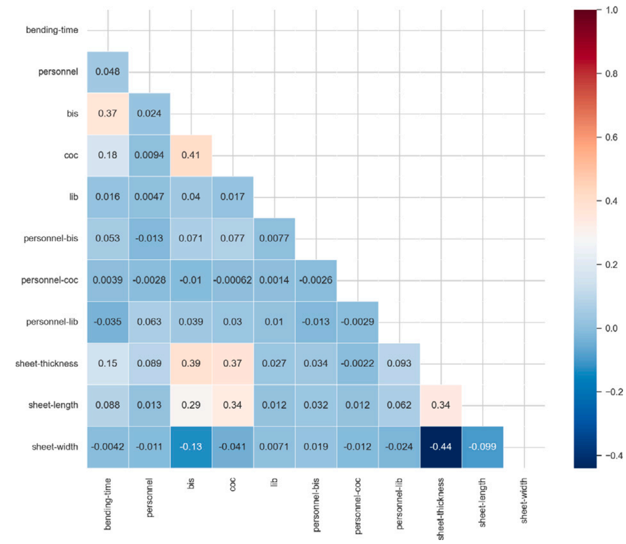


Fig. 4. Pearson correlation matrix for the bending process. The analysis includes only the most relevant numerical attributes affecting bending lead time, such as product specifications and upstream durations.

3.2.2. Correlation and feature importance analysis

To assess the initial predictive power of the numerical features, Pearson correlation matrices are computed for both the Bending and Longitudinal Welding (LW) processes, as shown in Figs. 4 and 5. This analysis is structured to mirror our sequential modelling approach: the first matrix examines the Bending process in isolation, while the second incorporates predictive features from the Bending stage to analyse their impact on the subsequent LW process. For clarity, both visualizations focus on a curated set of the most relevant numerical features identified through preliminary analysis and domain knowledge.

The correlation analysis for the Bending process, presented in Fig. 4, reveals that the strongest correlation is observed with the bevelling time (bis), with a value of 0.37, followed by sheet cutting time (coc) with a correlation of 0.18. This suggests that the duration of upstream operations directly influences the complexity and time required for the bending task.

Positive correlations are also found with material properties such as sheet thickness (0.15) and sheet length (0.088). These results indicate that bending time is affected not only by upstream operations but also by the geometric characteristics of the raw material. In contrast, features related to workforce allocation, bevel cleaning, or sheet width show very low or even negative correlations, implying a limited or negligible impact on bending lead time.

The correlation analysis for the longitudinal welding process (Fig. 5) shows that sheet thickness has the highest correlation with total lead time, with a value of 0.70. This suggests that thicker sheets generally require longer processing times, which is consistent with the operational logic of industrial processes.

In addition to material properties, the analysis confirms the value of incorporating data from the preceding stage. Upstream process times, such as bevelling, sheet cutting, and the historical bending-time itself, all show moderate positive correlations with the LW lead time. This demonstrates that the outcomes of prior operations have a cascading effect on subsequent tasks. In contrast, features like bevel cleaning (lib) and the number of assigned personnel show near-zero correlations, indicating a limited direct linear impact on the total welding time within this dataset.

3.2.3. Impact of human factors on process performance

Given the significant manual component of the operations, the influence of operator experience is analysed separately for each process.

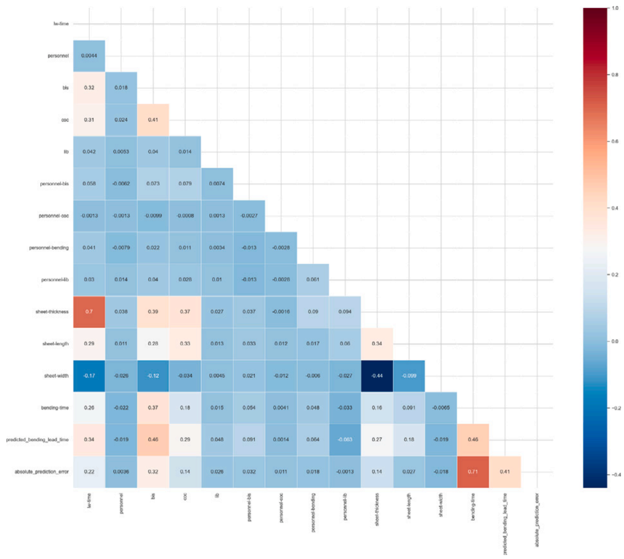


Fig. 5. Pearson correlation matrix for the longitudinal welding process. In addition to welding-specific features, this analysis incorporates attributes from the preceding bending process to capture potential interdependencies.

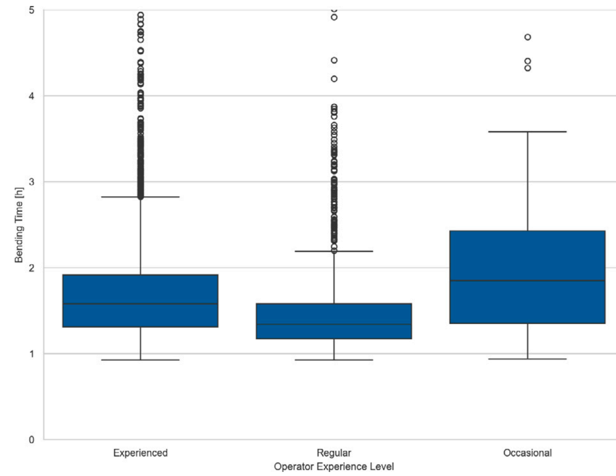


Fig. 6. Box Plot of lead times for the bending process, segmented by operator experience. The y-axis has been limited to the range [0, 5] hours to facilitate visual comparison.

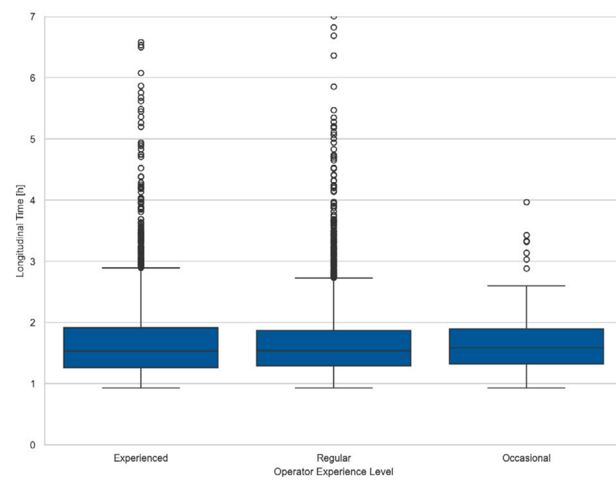


Fig. 7. Box Plot of lead times for the longitudinal welding process across experience levels. The y-axis has been limited to the range [0, 7] hours to facilitate visual comparison.

Using the total number of recorded operations as a proxy for experience, operators are categorized into 'Experienced' (> 500 operations), 'Regular' (100–500 operations), and 'Occasional' (<100 operations). Figs. 6 and 7 illustrate the resulting lead time distributions for the Bending and Longitudinal Welding processes, respectively.

Fig. 6 reveals a clear relationship between operator experience and performance in the bending process. 'Occasional' operators exhibit a higher median lead time (approx. 1.8 h) and, more significantly, much greater variability, as shown by the taller box and wider whisker range. This indicates a less predictable performance. In contrast, 'Regular' and 'Experienced' operators show a lower median time (approx. 1.5 h) and exceptional consistency. However, the most critical insight comes from the high frequency of outliers in these experienced groups. Given that the task is identical for all, these outliers do not represent more complex assignments. Instead, they likely represent the 'hidden work' of troubleshooting. When faced with a process disruption, experienced operators are expected to diagnose and resolve the issue, with this time being captured in their lead time. Less experienced operators, by contrast, would typically escalate the problem, thus externalizing the resolution time.

The analysis of the longitudinal welding process, shown in Fig. 7, strongly corroborates these findings. The same pattern emerges: 'Occasional' operators are slightly slower and less consistent, while 'Regular' and 'Experienced' operators perform at a higher and more predictable level. Crucially, the paradoxical pattern of outliers is also present, with the most experienced groups showing a much higher incidence of exceptionally long cycle times. This consistency across two different processes reinforces the hypothesis that these outliers are not indicators of inefficiency but are quantitative evidence of the additional responsibilities–such as on-the-spot problem-solving–handled by senior operators.

3.3. Data preprocessing

A significant challenge in this stage is ensuring the consistency and quality of the data. The information regarding lead times and machine usage, was manually entered by plant workers at the end of each operation. As a result, the data are susceptible to human error and missing values, which could negatively impact their quality. To mitigate these issues, a rigorous data preprocessing is applied. This includes the treatment of outliers, normalization, and the creation of composite variables, all of which help to improve the reliability and usefulness of the dataset for its subsequent use in machine learning models.

3.3.1. Data cleaning

In the data cleaning phase, which aims to enhance data quality through various techniques, the first step involves removing any duplicate elements. Subsequently, outliers in process delivery times are addressed, as some recorded values are unrealistic. For instance, a bending process was recorded as taking only 10 s, which is physically impossible for a human operator. Since such values do not reflect actual variability in the process, they are classified as outliers. To determine the threshold beyond which a value would be considered an outlier, the Isolation Forest algorithm, a widely used method for anomaly detection, is applied.

This algorithm focuses on isolating anomalies rather than modelling normal instances. It leverages the quantitative properties of anomalies, which are "few and different", making them more susceptible to isolation compared to normal data points. The algorithm constructs a set of isolation trees (iTrees) that recursively isolate instances. Anomalies tend to be separated closer to the root of the tree, whereas normal

Table 1
Ranges of non-outliers' values identified using the isolation forest algorithm for the bending and longitudinal welding.

Operation	Minimum (Non-outlier)	Maximum (Non-outlier)
Bending	0.9229 h	2.140 h
Longitudinal Welding	0.7163 h	1.771 h

Table 2
Dimensionality expansion after one-hot encoding.

Dataset	Initial attributes	Final attributes	New attributes added
Bending	29	114	85
LW	32	128	96

points are isolated at deeper levels. This approach enables efficient anomaly detection with linear time complexity and low memory requirements, making it well-suited for large-scale datasets (Liu et al., 2008).

The ranges considered as non-outliers are presented in Table 1. Additionally, the following percentages of values are identified as outliers:

- **Bending dataset:** 5.58% of values are below 0.9229 h, and 9.56% of values are above 2.140 h.
- **Longitudinal Welding dataset**: 4.75% of values are below 0.7163 h, and 11.97% of values are above 1.771 h.

Once the percentages are found, values above the maximum are retained because they represent errors that can occur in production, and therefore, they may help prepare the models for anomalous cases. However, in the case of the minimum values, it has been decided to eliminate them because they were deemed unrepresentative of typical production conditions and could potentially introduce bias or distort the model's ability to generalize to real-world scenarios. Regarding missing values, these are imputed using the mean (numeric variables) or the mode (nominal variables). It must be noted that most of the variables missing a significant percentage of data are quality control variables.

3.3.2. Data transform

The data transformation phase focuses on preparing the dataset for machine learning models by addressing differences in scale and converting categorical variables into numerical formats. Numerical data are scaled using the Standard Scaler method, which standardizes features by centring them around a mean of zero and a standard deviation of one, ensuring a consistent scale that is particularly beneficial for models sensitive to data magnitudes, such as neural networks and support vector machines.

For categorical variables, the One-Hot Encoding technique was applied, creating binary columns for each category to represent its presence or absence. This approach avoids imposing any ordinal relationships between categories, ensuring compatibility with machine learning algorithms while managing the resulting increase in dimensionality. As shown in Table 2, one-hot encoding significantly increases the dimensionality of the datasets: the bending dataset expands from 29 to 114 attributes (adding 85 columns), while the longitudinal welding dataset grows from 32 to 128 attributes (96 new columns). Given the final sample sizes after data cleaning, the feature-to-instance ratio remains sufficiently low to avoid the problem of dimensionality, ensuring robust model training.

3.3.3. Feature selection

Feature selection is the process of identifying the most relevant and representative variables in a dataset to enhance precision and efficiency. It is a crucial step in data preprocessing, aimed at reducing dimensionality by eliminating uninformative or noisy features. In this study, many quality-related features are removed, as 99% of the values in those columns are null, likely due to the limited number of tests conducted on that characteristic. Consequently, all columns with more than 35% null values are excluded to improve data reliability and model performance.

3.4. System design

The proposed system, illustrated in Fig. 8, extracts data from two primary sources: the ERP system and the QM system. The ERP system provides historical lead time records of upstream processes, contextual information, and the dependent variables–the bending lead time (LT) and longitudinal welding (LW) lead time. Meanwhile, the QM system contributes data from quality inspection reports, including raw materials and process-related quality data. Both datasets undergo a preprocessing phase, where data are cleaned, encoded for compatibility with machine learning models, and subjected to a feature selection process to retain the most relevant variables.

Once preprocessed, the data are used to train machine learning regression models, with one model dedicated to each operation. With this setup, two independent forecasting systems could be created. However, by linking the bending LT and LW LT prediction modules, a more integrated forecasting system is achieved. This approach follows the same rationale as the inclusion of lead times from previous processes such as sheet cutting, bevelling, and bevel cleaning. The integration is particularly relevant due to the expected high correlation between the bending quality and the LW process lead time. Specifically, if poor-quality bending causes misalignment in the sheet edges that are to be welded, the LW process can be significantly delayed.

The system is structured into two main stages: training and evaluation, followed by deployment in production. In the training and evaluation stage, separate machine learning models are developed for predicting bending LT and LW LT. This process involves hyperparameter tuning, model selection, and model evaluation to identify the most effective model for each task. In the case of the bending LT prediction model, different configurations are tested, including using only historical data, incorporating predicted values of bending LT, and considering the prediction error as an additional feature. The LW LT model, in turn, integrates the outputs from the bending LT model to improve its predictive accuracy. The best models are selected based on their performance, and a SHAP value analysis is conducted to interpret the contribution of different input variables.

Once trained, the models are deployed in the production stage to generate real-time lead time predictions. The bending LT prediction model produces an estimate, which, along with historical or predicted values, is used as input for the LW LT prediction model. The output of the LW LT model is then fed into the production planning and control module, where it assists in optimizing job scheduling and anomaly detection. The bending LT prediction module generates two key outputs: the bending lead time prediction and the actual bending lead time, from which the prediction error can be computed. These outputs enable different configurations for linking the bending and LW LT models, and the system's performance under these various setups is tested to determine the most effective configuration.

Finally, the predictions obtained with these models serve as inputs for other production planning and control systems, supporting job scheduling and anomaly detection. The comparative results of different system configurations and their impact on predictive performance are discussed in Section 4.

3.5. Lead time prediction modules implementation

As described in the previous subsection, the proposed system integrates two lead time prediction modules based on regression models, one for the bending process and another for the longitudinal welding

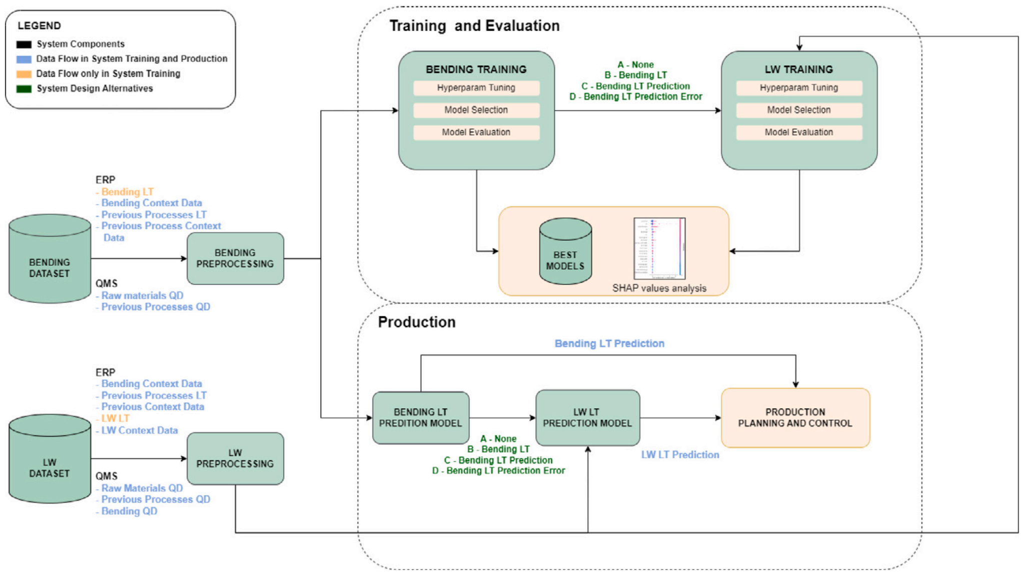


Fig. 8. System design diagram.

process. Both share a structured methodology that includes data partitioning, model selection, and hyperparameter tuning. However, the LW module has a particular feature: it can incorporate not only the actual bending lead time but also its estimated values and associated errors, allowing for the evaluation of different input configurations. The following section details the methodology used for the implementation of each module.

3.5.1. Bending module

First, the models were trained and evaluated without adjusting any hyperparameters, to use them as baseline models for comparison with those where hyperparameters are tuned. The bending dataset is divided into a training and validation set and a test set following the holdout method. Specifically, 20% of the instances in the dataset are randomly selected to form the test set, ensuring that these instances are not used during training or validation. This is essential to guarantee that the model performance metrics accurately reflect its generalization capability once deployed in production. The criterion to select a model is the minimization of the average across the five iterations of the Root Mean Square Error (RMSE) metric, which can be calculated as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} |y(i) - \hat{y}(i)|^2}{N}}$$

where N is the number of instances in the test set, $y(i)$ is the ith observation, and $\hat{y}(i)$ is its corresponding prediction. By squaring the residuals, this metric penalizes the larger errors more severely than others such as the Mean Absolute Error (MAE). This is particularly desirable for the bending lead time prediction module, the output of which will be employed in the forecast of the LW lead time. Thus, the model should perform well with extreme instances in order to pass any learnt information to the subsequent module.

For the implementation of the lead time prediction modules, ten regression models are evaluated, organized into different families. First, linear, and regularized regression models, such as Ridge, Lasso, and Elastic Net (ENET), are included. These models assume a linear relationship between variables and apply regularization to avoid overfitting, making them useful in problems with high collinearity. Next, Support Vector Regression (SVR), an extension of Support Vector Machines (SVM) for regression tasks, is evaluated. SVR captures non-linear relationships through the use of kernel functions, which implicitly map data into higher-dimensional feature spaces. This makes it particularly effective for small datasets and problems with high-dimensional feature spaces.

Regarding tree-based models, they are divided into individual trees, such as Decision Tree Regressor (DT), which recursively partitions the feature space, and ensemble tree methods, such as Random Forest (RF), Gradient Boosting (GB), LightGBM (LGBM), and XGBoost (XGB). The latter combine multiple trees to improve performance and balance bias and variance. Boosting methods build trees sequentially, correcting the errors of the previous ones.

Finally, models based on artificial neural networks are included, such as MLP Regressor, which uses a fully connected neural network to model complex relationships.

Once the correct functioning of the base models is verified, Random Search optimization combined with cross-validation is applied to fine-tune their hyperparameters and improve performance. This technique is selected due to its balance between efficiency and simplicity, outperforming Grid Search by avoiding exhaustive exploration of large hyperparameter spaces (Bergstra & Bengio, 2012) and being less computationally expensive than Bayesian Optimization while still achieving competitive results. The specific hyperparameters used in each algorithm are detailed in the Appendix A.

After that, the tuned model is finalized by training it on the full training and validation dataset, and is used to predict the bending lead time for the instances in the test set. If the accuracy of these predictions is acceptable, the model is retrained with the full dataset and deployed for production. This would conclude the process in the bending lead time prediction module.

However, if the predictions or prediction errors of the bending lead time are to be used in the LW lead time forecasting module, predicted bending lead time values are needed for as many instances in the dataset as possible. To do so, the dataset containing the historical bending lead time records and the corresponding independent variables is split into four subsets with the same number of instances. Four iterations of the ML algorithm are executed. Each one of them uses three subsets as training data. The instances in the remaining one are used to generate the bending lead time predictions. After completing the four iterations, all instances in the dataset contain an actual bending lead time value, a prediction, and the prediction error. This way, each of these values can be fed to the LW module.

3.5.2. Longitudinal welding module

Regarding the prediction of the LW lead time, the dataset is equally divided into a random 80%–20% training-test split. Again, a five-fold cross-validation analysis must be performed in order to identify the

Table 3
Hardware and software specifications used in the experiment.

Component	Details
CPU	Intel Core i9-12900K, 16 cores
RAM	32 GB
Operative System	Windows 11
Programming Language	Python 3.10
Packages	numpy (Charles et al., 2020), pandas (McKinney, 2010), matplotlib (Hunter, 2007), seaborn (Waskom, 2021), scikit-learn (Pedregosa et al., 2011)
IDE	VS Code

optimal regression model, also with the goal of minimizing the RMSE. The same candidate models as in the bending lead time prediction module have been used. This is done for each of the configurations of input variables that can be used. As a reminder, the input variables can be classified into four categories: lead time records of the upstream processes, context information, quality control reports and ML regression model predictions. The variables of the first three groups are included in all of the configurations, which vary from one another with the inclusion of the actual bending lead time values, the predictions of such, the error of the predictions, the absolute error of the predictions and combinations of the above.

Finally, the selected model from each configuration is tuned using an analogous hyperparameter Random search and trained with the entirety of the training and validation dataset. The trained model is then used to produce predictions for the instances in the test set. If the results of the evaluation are acceptable, the model is retrained using the complete dataset and deployed for production.

4. Experimental results

In this section, the results obtained using the methodology for predicting lead times in bending and longitudinal welding operations in wind turbine tower manufacturing will be discussed. Additionally, the experimental setup employed, the deployment feasibility analysis, as well as the interpretability of the models – including the collinearity assessment of our variables – will be described.

4.1. Experimental environment

For the development of this paper, all code has been written using the Python programming language. Additionally, the details of the environment in which the experiments are conducted, including the Python version, libraries, and other relevant aspects, are presented in Table 3.

4.2. Bending results

In Table 4, the results of various models evaluated on the Bending dataset using Random Search optimization are presented. LightGBM (LGBM) achieves the lowest MAE and MAPE values, indicating good performance in terms of absolute error. On the other hand, GB stands out by achieving the lowest RMSE (0.603) and the highest R^2 coefficient (0.293), suggesting a better balance between precision and explanatory power. The optimal hyperparameters for the GB model, which contribute to these results, are detailed in Table 5. In terms of efficiency, linear models such as Ridge and Lasso are the fastest in training and optimization, as they only require tuning a single hyperparameter. However, their accuracy is lower compared to the other models. In contrast, XGBoost (XGB) requires the longest tuning time (383.851 s), which could be a limiting factor in time-constrained scenarios.

The results identify GB as the optimal model, based not only on the RMSE metric but also on its superior R^2, making it the most appropriate choice for this case. The GB, developed by Friedman (2001), is an ensemble method that trains multiple decision trees iteratively using a greedy algorithm, and it is applicable to both classification and regression tasks.

The metrics indicate that the predictive power of the GB Regressor model is relatively weak. While previous studies consider MAPE values between 20% and 50% as indicative of "reasonable forecasting" (Lewis, 1982), the results show a notable improvement over mean-based predictions, with the MAPE decreasing by more than 4 percentage points. However, this moderate performance is likely attributable to the inherent high variability of the manual bending process and the limitations of manually recorded lead time data. Despite these challenges, the estimations provided by the GB model can still offer useful insights in an industrial context where inefficiencies and inaccuracies are common. Furthermore, as elaborated in Section 5, potential avenues for enhancing the accuracy of the model include advanced feature engineering, incorporating alternative data sources, and exploring other machine learning approaches. Notably, these estimations contribute to improved performance in predicting LW lead times later in this study.

4.3. Longitudinal welding results

Regarding the lead time prediction module in LW, it is important to recall that there are multiple feasible system configurations, depending on the variables linking the bending and LW modules. Twelve system configurations have been tested, and the performance evaluation metrics of the selected models, after conducting a random search for hyperparameters for each configuration, are shown in Table 6. For the sake of brevity, results for each configuration are not displayed; only the best model found for each case is presented. Additionally, performance metrics of a simple mean-based prediction are included as a reference. This model always predicts the mean value of the target variable, without using any additional information, serving as a baseline to assess whether the more complex models truly improve performance.

Appendix B summarizes the best model and its associated configurations identified through the experiments conducted for the longitudinal welding (LW) process. The results presented in Table 6 offer very interesting conclusions about the prediction system. In general, the GB model and its derivative models, such as LightGBM (LGBM) and XGBoost (XGB), have proven to be highly effective across all analysed input variable configurations. Additionally, Random Forest (RF), another ensemble method, has also demonstrated strong performance in certain scenarios.

XGBoost and LightGBM are advanced algorithms also based on boosting but with optimizations that further enhance their performance. XGB, as described in Chen and Guestrin (2016), is optimized through the regularization of the generated models and stands out for its high computational efficiency, ability to avoid overfitting, and advanced optimization techniques, such as handling missing data and support for both classification and regression tasks. Meanwhile, LightGBM, developed by Ke et al. (2017), introduces an algorithm that samples data instances, retaining those that contribute the most to information gain, i.e., those with larger gradients, and also includes a novel algorithm for feature bundling.

Random Forest (RF), first proposed by Breiman (2001), is another powerful ensemble method that constructs multiple decision trees and combines their predictions through averaging in regression problems. Unlike boosting methods, RF builds trees independently using bootstrap samples and random feature subsets, which contributes to its robustness and versatility in handling both classification and regression tasks.

Furthermore, using the actual value, the predicted value and the absolute prediction error of the bending lead time as input for the Longitudinal Welding lead time prediction module offers the best performance out of every other configuration in terms of RMSE (the decision criterion) (0.821 h) and R^2 (0.672). By using only the predicted value, the highest value of MAE is achieved (0.524 h) and

Table 4
Results obtained for each model with the best configuration through Random optimization on the Bending dataset.

Model	MAE [h]	RMSE [h]	MAPE	R^2	Tuning time [s]	Training time [s]	Testing time [s]
Baseline: Mean-based prediction	0.450	0.717	26.461	−0.001	–	–	–
Ridge	0.433	0.657	25.083	0.161	1.476	**0.006**	0.00139
Lasso	0.437	0.661	25.476	0.149	2.402	0.184	**0.00122**
ENET	0.436	0.660	25.387	0.154	19.140	0.642	0.00126
SVR	0.419	0.633	24.846	0.221	17.238	0.553	0.161
DT	0.397	0.634	22.535	0.219	**1.324**	0.057	0.00130
RF	0.388	0.607	22.527	0.283	13.884	0.565	0.0174
GB	0.383	**0.603**	22.188	**0.293**	178.981	1.668	0.0222
LGBM	**0.377**	0.620	**21.276**	0.252	46.202	0.079	0.00188
XGB	0.380	0.653	21.521	0.170	383.851	0.079	0.0150
MLP	0.413	0.632	23.759	0.223	88.931	2.586	0.00208

Table 5
Hyperparameters settings of the best GB model obtained for the Bending dataset.

Parameter	Parameter settings
learning_rate	0.0103
max_depth	8
max_features	sqrt
n_estimators	401
subsample	0.5599

produces the lowest MAPE (26.524%). These can be considered good predictions, particularly when compared to the values obtained without any input from the bending module and, even more, to those obtained using only the mean lead time as the prediction. The MAPE can be reduced by almost 50% using only the prediction error when compared to the mean-based predictions.

4.4. Model interpretability

The interpretability of machine learning models is crucial in real-world applications, particularly when decisions need to be understood by domain experts. Some models, like decision trees, are inherently interpretable (white-box models), while others, such as neural networks or support vector machines, are more opaque (black-box models), requiring additional techniques for explanation (Loyola-González, 2019).

White-box models, like decision trees, provide transparency through feature importance, making it easy for experts to understand how individual features influence predictions. This direct interpretability allows users to identify which features are most important for the model's decision-making process. For example, in decision trees, the structure itself provides a clear path of decision rules based on the feature values, enabling an intuitive understanding of the model's predictions.

In contrast, for black-box models, such as neural networks or support vector machines, interpretability is not straightforward. These models do not provide direct insights into how features influence their decisions. However, techniques like SHAP (Lundberg & Lee, 2017) or LIME (Ribeiro et al., 2016) can help by approximating the influence of each feature on individual predictions, even for complex models. These techniques aim to enhance the interpretability of otherwise opaque models by providing a clear explanation of how input features contribute to the output.

In this study, SHAP is used due to its ability to provide both global and local explanations of model predictions. SHAP (SHapley Additive exPlanations), developed by Shapley (1953) is based on cooperative game theory, and calculates the contribution of each feature using **Shapley values**. These values measure how much each feature contributes to the prediction, considering all possible combinations of features and assigning a fair weight to each one. The Shapley value for a feature i is mathematically defined as the weighted sum of its marginal contributions across all possible subsets of features. Formally, it is expressed as:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} \left(f(S \cup \{i\}) - f(S)\right) \quad (1)$$

Where:

- N represents the complete set of features in the model.
- S denotes a subset of features excluding i.
- $f(S)$ is the model's prediction when only the features in S are used.
- $f(S \cup \{i\})$ is the prediction when feature i is added to the subset S.
- ϕ_i is the Shapley value for feature i, quantifying its contribution to the model's output.

Shapley values are derived from the idea of fairly distributing the 'payout' (in this case, the model's prediction) among the 'players' (the features). To achieve this, all possible combinations of features are considered, and the marginal impact of including a specific feature in each combination is measured.

For the prediction of models in each experiment, the model and hyperparameters that demonstrate the best performance in terms of evaluation metrics are used. The results of these models are visualized through graphs depicting the 20 most relevant features, as presented in Figs. 9 and 10.

In these graphs, the X-axis represents the SHAP value, which quantifies the impact of each feature on the model's prediction, expressed as a change in log-odds. A positive SHAP value indicates an increase in the prediction, whereas a negative value suggests a decrease. The Y-axis displays the features ranked by their global importance, from highest to lowest, as determined by the absolute magnitude of their SHAP values. Additionally, the colour of each point encodes the feature value for each instance in the dataset: red tones represent high values, while blue tones indicate low values. Each point corresponds to an observation (row) in the original dataset, allowing for the identification of patterns and trends in the influence of features on the model's predictions.

Fig. 9 provides a detailed view of the contribution of each feature to the predictions of the GB model applied to the Bending dataset. Among the most influential variables, personnel, bevelling time, sheet thickness, and bevel cleaning time stand out, suggesting that factors related to labour and material properties have a significant impact on the prediction. In particular, sheet thickness exhibits a predominantly positive correlation, indicating that an increase in sheet thickness tends to raise the model's prediction.

Furthermore, Fig. 10 provides a detailed insight into the contribution of each feature to the predictions of each model across the experiments conducted on the longitudinal welding dataset. Among the most influential variables, sheet thickness, sheet width, and personnel stand out, suggesting that both metal properties and human resources have a significant impact on the predictions. As observed in bending, sheet thickness exhibits a predominantly positive correlation. Additionally, the personnel involved in the bending process (personnel-bending)

Table 6
Performance evaluation metrics for the LW lead time prediction module with different system configurations.

Experiment	Model	MAE [h]	RMSE [h]	MAPE	R^2	Tuning time [s]	Training time [s]	Testing time [s]
Baseline: Mean-based prediction	–	0.981	1.437	55.597	−0.003	–	–	–
None	GB	0.535	0.857	27.272	0.643	175.619	19.388	0.016
Actual value	GB	0.540	0.873	27.618	0.630	159.194	24.003	0.027
Predicted value	GB	**0.524**	0.833	**26.524**	0.663	221.555	35.468	0.029
Prediction error	LGBM	0.538	0.877	27.411	0.627	38.836	0.177	**0.002**
Absolute prediction error	GB	0.532	0.869	27.117	0.633	295.295	18.639	0.015
Predicted value + Prediction error	GB	0.541	0.843	27.371	0.655	191.546	11.252	0.008
Predicted value + Absolute prediction error	RF	0.572	0.877	29.682	0.626	**15.769**	1.271	0.034
Actual value + Predicted value	GB	0.533	0.847	26.725	0.651	342.586	16.057	0.012
Actual value + Prediction error	XGB	0.544	0.878	27.830	0.626	444.824	**0.088**	0.017
Actual value + Absolute prediction error	GB	0.543	0.880	28.038	0.624	352.550	40.457	0.041
Actual value + Predicted value + Prediction error	GB	0.528	0.835	26.871	0.661	240.571	27.035	0.026
Actual value + Predicted value + Absolute prediction error	GB	0.535	**0.821**	26.984	**0.672**	322.383	49.167	0.038

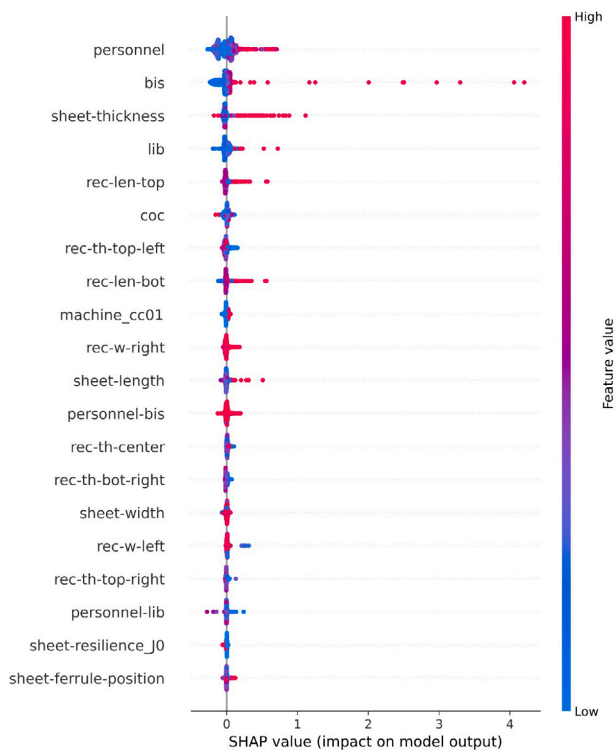


Fig. 9. SHAP Values for the Bending Lead Time Prediction Model. The plot illustrates the importance and impact of the top 20 features. Each point represents a single prediction instance. The feature's importance is ranked on the y-axis. The SHAP value on the x-axis indicates the feature's impact on the model's output. Positive values increase the predicted lead time. The colour indicates the feature's value for that instance, with red for high values and blue for low values.

and machine-related features also play a crucial role, highlighting the importance of operators and specific equipment in the longitudinal welding process. The inclusion of prediction error metrics has enabled a better understanding of the model's accuracy and reliability, illustrating how these metrics influence the interpretation of results. Taken together, these findings emphasize the importance of a careful feature selection process and the need to consider both material properties and operational and human factors to optimize model predictions.

4.5. Collinearity assessment

While SHAP analysis provides powerful insights into feature contributions, its interpretations can be sensitive to multicollinearity, where input features are highly correlated. To ensure the robustness of our feature importance rankings, a collinearity assessment is performed using the Variance Inflation Factor (VIF). The VIF quantifies how much the variance of an estimated regression coefficient is increased due to collinearity. A common rule of thumb considers a VIF value greater than 5 or 10 as an indicator of potentially problematic multicollinearity.

Table 7
Variance Inflation Factor (VIF) Results for key predictors in bending and longitudinal welding models.

Bending model		Longitudinal welding model	
Feature	VIF	Feature	VIF
sheet-thickness	1.72	cur	7.96
bis	1.39	sheet-thickness	2.82
coc	1.39	absolute_prediction_error	2.07
sheet-width	1.37	predicted_bending_lead_time	1.54
sheet-section-position	1.22	bis	1.50
sheet-length	1.20	sheet-width	1.44
sheet-ferrule-position	1.07	coc	1.40
personnel-bis	1.02	sheet-section-position	1.24
personnel-lib	1.02	sheet-length	1.21
personnel	1.02	sheet-ferrule-position	1.08
lib	1.00	–	–
personnel-coc	1.00	–	–

The analysis reveals two distinct patterns in our dataset. A specific group of features – namely the multiple, detailed measurements of sheet thickness, width, and length – exhibit extremely high VIF values, significantly exceeding the threshold of 10. This is an expected result, as these variables inherently capture redundant information about the same physical part.

Conversely, most of the other features, including the key drivers identified in our SHAP analysis, show low to moderate VIF values. Table 7 presents the VIF scores for the main predictive variables in the bending and LW model, excluding the highly inter-correlated dimensional measurements for clarity.

Overall, the GB model shows robustness to multicollinearity by selecting individual features for decision splits, thereby limiting the influence of redundant variables. Nevertheless, interpretability can be affected, especially for groups of correlated features whose importance should be assessed collectively rather than in isolation. The fact that top-ranked variables – such as staff experience and overall sheet thickness – present low VIF values supports the validity of their predictive power and reinforces the robustness of our study's main findings.

4.6. Deployment feasibility analysis

To validate the suitability of the system for production environments and assess its real-time operational capabilities, a simulated deployment experiment is conducted. This analysis aims to measure the end-to-end prediction latency by replicating the operational workflow, where on-demand predictions are required for individual manufacturing operations.

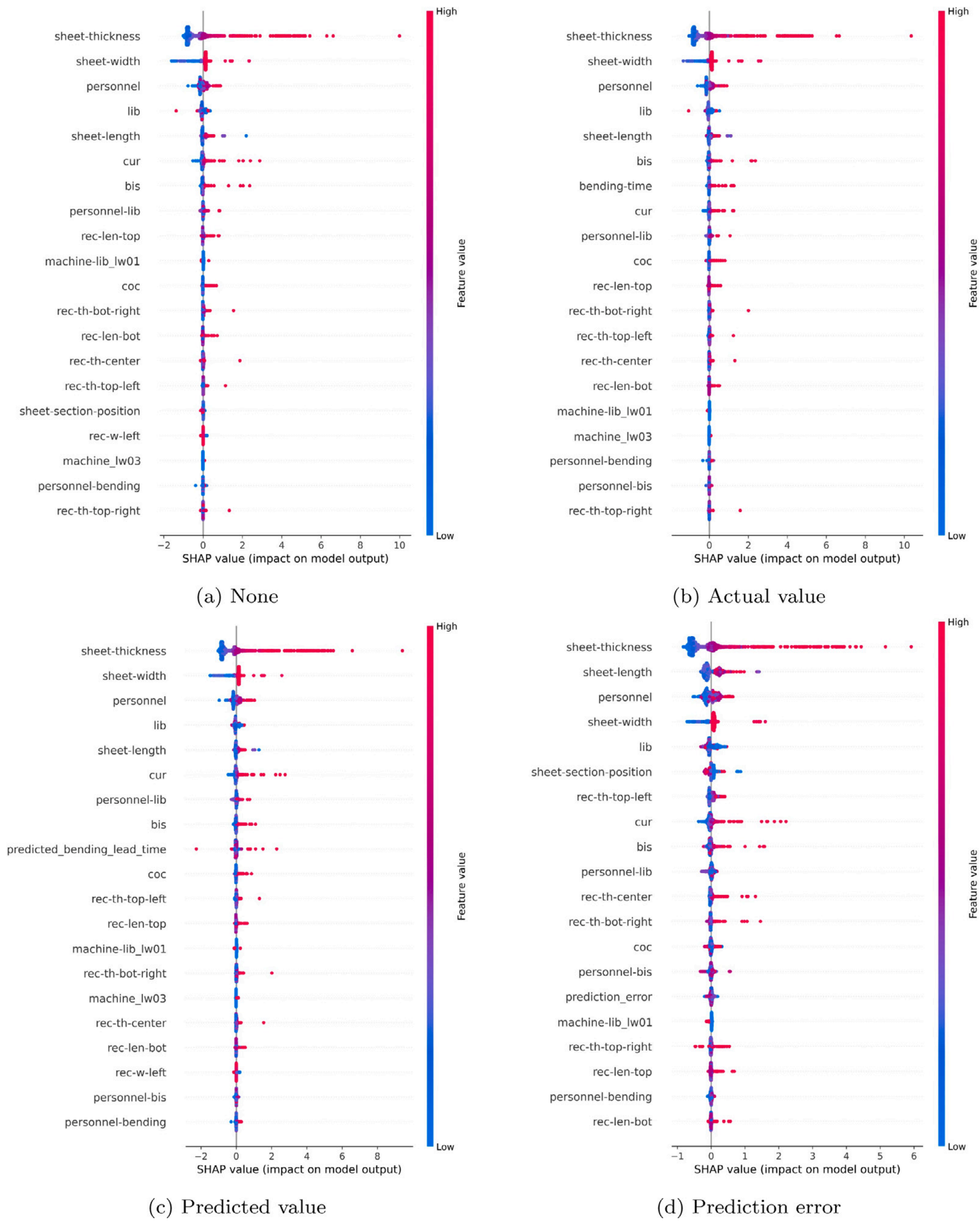


Fig. 10. SHAP values for different experiments on the LW dataset (part 1). This figure compares the top 20 feature importances across 12 LW lead time prediction models, each trained with a different set of bending-related inputs. In each subplot (a–l), features are ranked by importance (y-axis), and SHAP values (x-axis) indicate their effect on lead time predictions (positive = longer lead time). Colour represents the feature value (red = high, blue = low).

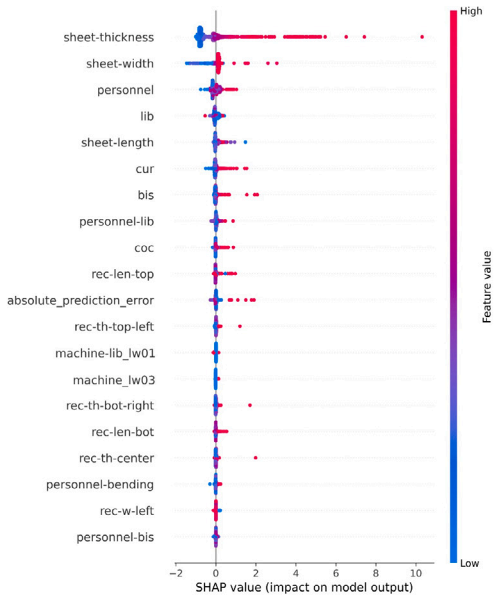
The methodology consists of timing the entire prediction pipeline for single instances from the test set. This pipeline includes loading the pre-trained model along with its preprocessing artifacts, applying real-time transformations to raw input data, and performing model inference to generate a prediction. The experiment is carried out for the best models associated with the Bending and Longitudinal Welding (LW) processes, in order to obtain a comprehensive view of the system's computational performance.

As shown in Table 8, the average prediction latency remains below 1 millisecond for both models, with the 95th percentile latency not exceeding 1.11 ms. This performance translates into a high throughput of over 1437 predictions per second for the Bending model and more than 1136 predictions per second for the LW model, all executed on standard hardware as detailed in Table 3.

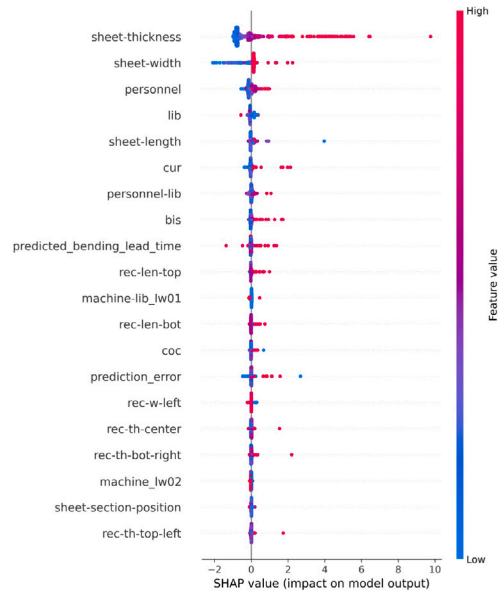
Table 8
End-to-end prediction performance in a simulated deployment environment.

Metrics	Bending model	LW model
Mean Latency (ms)	0.70	0.88
Std Dev Latency (ms)	0.09	0.12
Min Latency (ms)	0.60	0.73
Max Latency (ms)	1.35	1.95
P95 Latency (ms)	0.88	1.10
Throughput (pred/s)	1437.49	1136.04

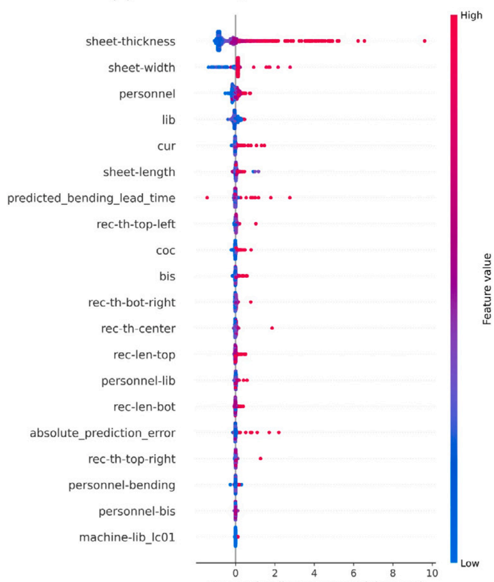
Such computational efficiency indicates that the system imposes a negligible processing load in the context of its intended use for production planning and control, where decision cycles typically operate



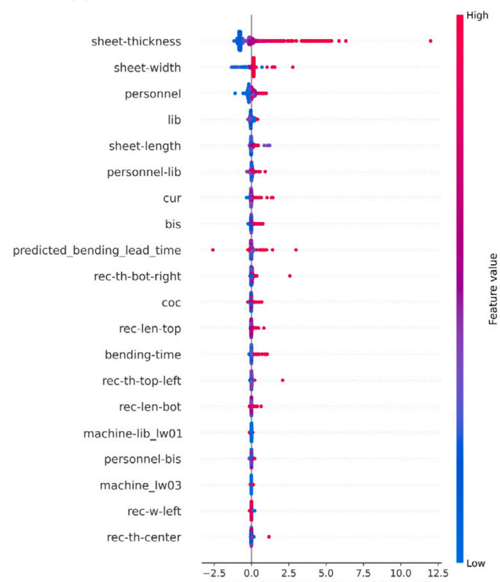
(e) Absolute prediction error



(f) Predicted value + Prediction error



(g) Predicted value + Absolute prediction error



(h) Actual value + Predicted value

Fig. 10. (*continued*).

on the scale of minutes or hours. These quantitative results not only confirm the accuracy of the system but also demonstrate its strong suitability for real-world industrial deployment, providing decision support in near real-time without compromising computational resources.

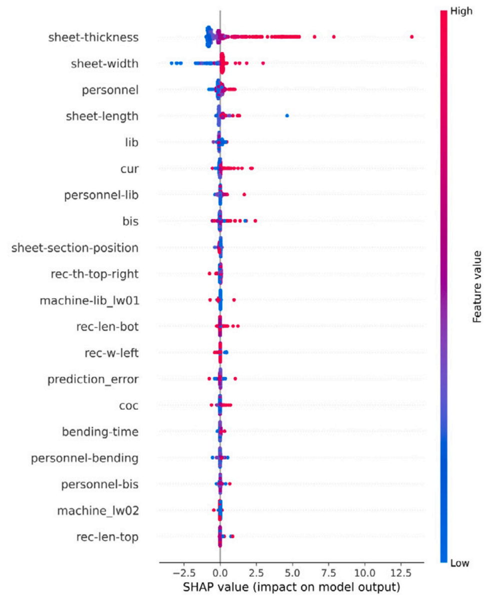
5. Discussion

This section presents a comprehensive discussion of the results obtained in this study, contextualizing them both in the academic literature and in the industrial realities of this field of study. We analyse the comparative performance of machine learning models against traditional engineering methods, explore the broader industrial implications and limitations of the findings, and propose strategies for improving prediction accuracy in particularly challenging stages of the process. Through this discussion, we aim to provide insights not only into the technical contributions of this work but also into its practical relevance and future research directions.

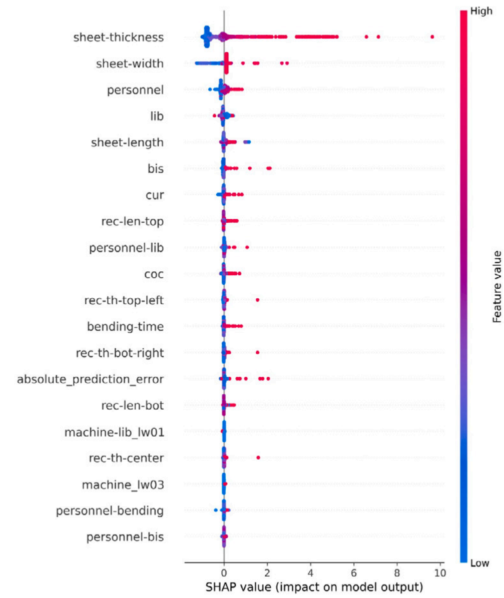
5.1. Comparison between ML models and traditional engineering methods

The results obtained in this study demonstrate that machine learning (ML) models are capable of providing more accurate estimates of production times compared to traditional engineering methods, particularly in the context of the longitudinal welding process. This improvement in accuracy is reflected in the error metrics used, where the ML model with the best performance presents significantly lower MAE and RMSE values compared to those obtained using traditional engineering estimates.

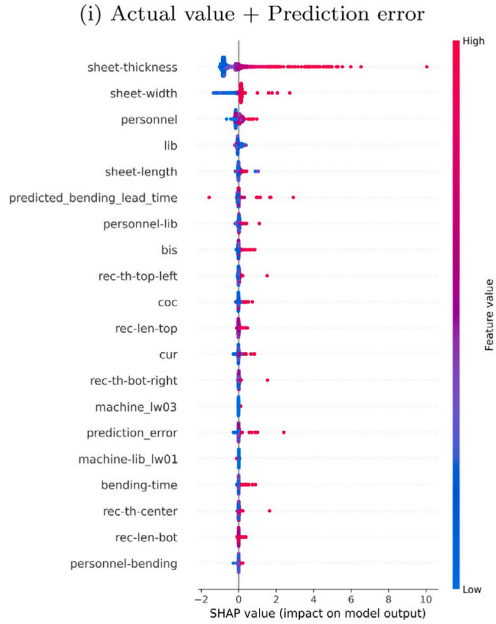
One of the key advantages of ML models is their ability to incorporate interpretability, which allows for the identification of the attributes that truly influence the predictions. This approach not only facilitates model optimization but also allows for the elimination of variables that, although initially considered relevant, do not have a significant impact on the predictions. Furthermore, ML models can be continuously updated as new data are incorporated, ensuring that



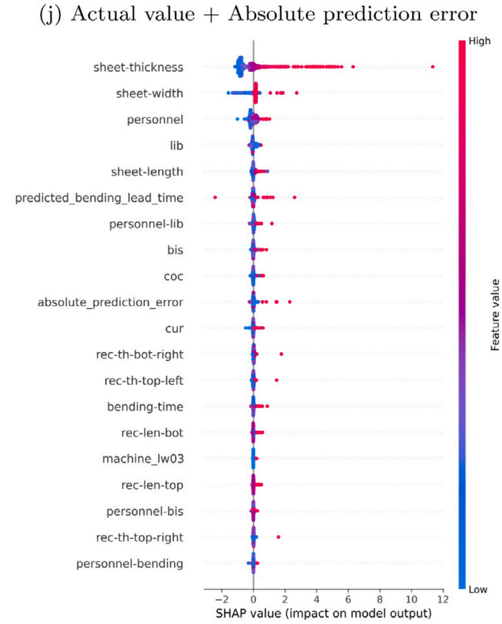
(i) Actual value + Prediction error



(j) Actual value + Absolute prediction error



(k) Actual value + Predicted value + Prediction error



(l) Actual value + Predicted value + Absolute prediction error

Fig. 10. (*continued*).

the predictions remain aligned with the current conditions of the manufacturing process. This characteristic sets ML models apart from approaches such as direct formulation, which can be unrealistic, or linear programming, which does not allow for dynamic updates.

However, the majority of previous studies on the estimation of lead times in the manufacture of wind turbine towers have focused on the implementation of ML models without a comparative evaluation against traditional engineering methods. This lack of comparison can hinder informed adoption of these technologies by industry professionals. This study seeks to address this gap by conducting a comparative evaluation of delivery time predictions for the longitudinal welding process, using both traditional engineering methods and the ML models developed in our study.

In this context, we compare the predictions of the ML model with the best performance for the longitudinal welding process with the time calculations employed in the studied factory. Traditional engineering methods estimate the total time for each tower section based on various input characteristics, such as structural features, weldable internal elements, and surface treatment schemes. However, one of the main drawbacks of these methods is that the formulas used may be based on outdated experiences, as they are not continuously updated.

In the approach adopted in this study, individual ferrule times are used to make predictions for each ferrule within a tower section, which are then summed to obtain the total time for the section. This approach differs from that used in engineering, which works with the full section.

Table 9 presents a comparison between the times estimated using the engineering method and the predictions made by the ML model, relative to the actual times recorded in the factory. It is evident that the mean absolute error (MAE) of the ML model is 2.03, significantly lower than the value of 11.36 obtained using the traditional method. Similarly, the root mean square error (RMSE) decreases from 12.01 with the engineering method to 3.13 with the ML model. Additionally, the maximum deviation decreases from 28.56 to 21.59, indicating a lower dispersion of errors in the predictions made by the ML model.

Table 9
Comparison of engineering method and longitudinal welding machine learning prediction relative to the actual times obtained in the factory.

Method	Max deviation	Min deviation	MAE	RMSE
Engineering	28.56	0.21	11.36	12.01
ML Prediction	21.59	0.00	2.03	3.13

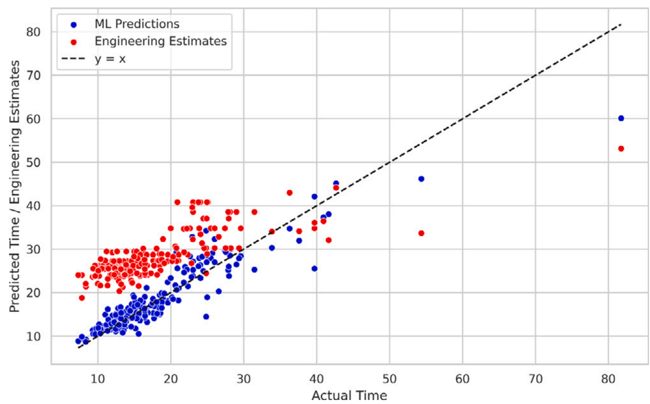


Fig. 11. Comparison of Machine Learning predictions and engineering estimates with actual manufacturing lead times. The scatter plot contrasts the accuracy of ML predictions (blue) and traditional engineering estimates (red) relative to the actual lead times (x-axis). The closer alignment of the ML points to the identity line ($y = x$) indicates a higher predictive accuracy of the ML model.

The significant improvement achieved by the machine learning model is further illustrated in Fig. 11. This visualization reinforces the findings presented in Table 9, showing that the ML model's predictions align more closely with the actual observed lead times than the traditional engineering estimates. The plot provides clear visual evidence of the model's capacity to capture the underlying complexities of the manufacturing process, resulting in more reliable and accurate forecasts.

5.2. Error propagation and reliability analysis of the model

To assess the model's robustness beyond overall accuracy, we investigate how prediction errors propagate across sequential manufacturing stages. We hypothesize that instances with high prediction errors in an early stage (Bending) would also exhibit high errors in a subsequent stage (Longitudinal Welding, LW). To test this, we employ a stratified sampling approach on the test set, creating a representative sample of instances with low, medium, and high absolute prediction errors from the Bending stage.

As shown in Fig. 12, the analysis reveals a moderate and statistically significant positive correlation between the absolute prediction errors of the two stages (Pearson's $r = 0.5082$, $p = 0.00016$). The consistency in prediction uncertainty suggests that a high error in early stages can serve as an early warning signal. This enables planners to identify potentially problematic towers and apply proactive mitigation strategies to improve the reliability of the production schedule.

5.3. Industrial implications and limitations

The findings of this study have significant industrial implications. The ability to predict longitudinal welding times with greater accuracy can enhance production planning, optimize resource allocation, and reduce costs associated with downtime or inaccurate estimates. However, it is important to acknowledge that the performance of the ML model heavily depends on the quality and quantity of available data. In environments where historical data are limited or biased, traditional methods may still offer a more stable and reliable alternative.

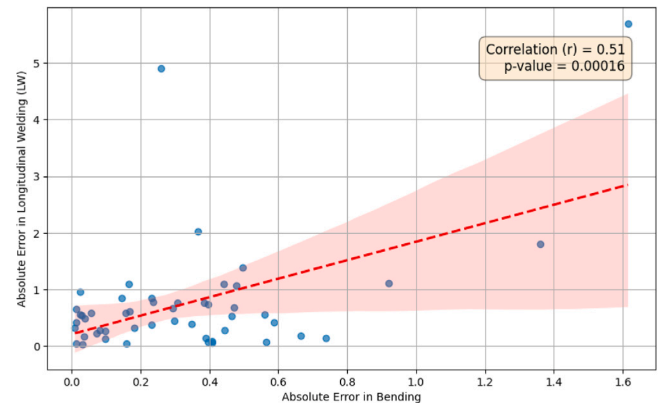


Fig. 12. Correlation between absolute prediction errors in the bending and longitudinal welding (LW) stages.

A limitation of this study is that it is based on historical production data and has not been validated in a real-time production environment. Future research could explore the implementation of these models in live monitoring systems, as well as the integration of online learning techniques to improve the model's ability to adapt to changes in manufacturing conditions.

However, to mitigate concerns about deployment feasibility, a simulated performance analysis is conducted (see Section 4.6), which confirms that the system's end-to-end latency is minimal, making it highly suitable for integration into a live production environment.

5.4. Strategies for continuous performance enhancement of predictive models

The modular nature of the proposed lead-time forecasting system allows iterative refinement and improvement of individual component models. Although the overall system has significant advantages, some individual process models (such as the folding lead time model of this study) may show moderate performance due to factors such as high process variability, reliance on manual data, or complexity of the specific operation. In this framework, several general strategies can be systematically applied to improve the accuracy of any such prediction model:

- Advanced feature engineering: the quality of input features is paramount. Continuously exploring and engineering new features based on domain expertise is crucial for improving model performance. This includes creating interaction terms between existing variables (e.g., in the longitudinal welding process, one might consider interaction terms such as the number of passes combined with the operator's experience, since multi-pass weld quality often depends heavily on the welder's skill), adding polynomial features to capture non-linear relationships, or deriving time-based features such as rolling averages of past cycle times or operator fatigue indicators (if measurable). For example, if a particular welding model is underperforming, it might be valuable to investigate features related to ambient temperature or humidity that were not previously considered.
- Alternative data sources: reducing dependence on manual data entry is a crucial objective to minimize errors and biases. For any model, the integration of data from alternative sources should be considered. This may include incorporating sensor data (such as vibration, temperature, energy consumption or image-based quality assessments), detailed operational logs capturing precise machine settings, tooling configurations, and minor process interruptions, as well as upstream or downstream quality metrics that exert a measurable influence on process outcomes or lead times.

Table A.10
Hyperparameter tuning ranges for different models.

Model	Parameter	Tuning range	Description
Ridge/ Lasso	alpha	[0.001,2]	Regularization parameter
ENET	alpha	[1×10^{-15},1]	Regularization parameter
	l1_ratio	[1×10^{-15}, 1]	L1 regularization coefficient
DT	max_depth	[1,32]	Maximum depth of each tree
	min_samples_split	[2,20]	Minimum number of samples required to split a node
	min_samples_leaf	[1,20]	Minimum number of samples required to be at a leaf node
RF	max_features	{sqrt, log2}	Number of features to consider for each split
	n_estimators	[10,200]	Number of trees in the forest
	max_depth	[4,20]	Maximum depth of each tree
	min_samples_split	[2,20]	Minimum number of samples required to split a node
	min_samples_leaf	[1,10]	Minimum number of samples required at each leaf
GB	n_estimators	[50,1000], step=50	Number of estimators
	learning_rate	[1×10^{-4}, 0.3]	Learning rate
	max_depth	[3,9]	Maximum depth of trees
	subsample	[0.5,1.0], step=0.1	Proportion of samples used to train each tree
	max_features	{sqrt, log2, None}	Number of features to consider for each split
XGB	booster	{gbtree, gblinear, dart}	Type of booster
	reg_lambda	[1×10^{-8}, 1.0]	L2 regularization coefficient (lambda)
	reg_alpha	[1×10^{-8}, 1.0]	L1 regularization coefficient (alpha)
	max_depth	[3,11]	Maximum depth of trees
LGBM	boosting	gbdt	Type of boosting algorithms
	lambda_l1	[1×10^{-8}, 10.0]	L1 regularization coefficient
	lambda_l2	[1×10^{-8}, 10.0]	L2 regularization coefficient
	num_leaves	[2,256]	Maximum number of leaves per tree
	feature_fraction	[0.4,1.0]	Proportion of features used per iteration
	bagging_fraction	[0.4,1.0]	Proportion of samples used for bagging
	bagging_freq	[1,7]	Frequency of bagging (0 disables bagging)
	min_child_samples	[1,100]	Minimum number of data points in a leaf
	min_split_gain	[0.0,0.2]	Minimum loss reduction required to make a split
SVR	kernel	linear, rbf, poly	Type of kernel
	C	[0.1,10]	Regularization parameter
	gamma	{scale, auto}	Kernel coefficient
	epsilon	[[1×10^{-3},10]	Error tolerance
MLP	hidden_layer_sizes	(50,), (100,), (50, 50), (100, 50)	Hidden layer sizes
	activation	relu, tanh	Activation function
	solver	ADAM, SDG	Algorithm for weight optimization
	learning_rate	constant, adaptive	Learning rate used by the solver

- Model re-evaluation and hyperparameter optimization: periodically re-evaluating the choice of algorithm for a specific process, or conducting more advanced hyperparameter tuning (e.g., moving beyond standard random search to Bayesian optimization or other sophisticated methods) as more data becomes available or features are refined, is a standard practise.

5.5. Economic impact and strategies for cost reduction

While the technical performance of the ML models is the primary focus of this study, their ultimate industrial value is determined by their ability to reduce production costs and enhance operational efficiency. The moderate performance of the bending model, for instance, is not just a statistical metric but represents a tangible source of economic uncertainty. This uncertainty translates into direct and indirect costs, such as:

- Increased buffer times: planners must account for prediction variability by allocating extra time between sequential operations (bending and LW), leading to planned idle time for downstream workstations and personnel, which is a direct cost.
- Reduced throughput: unexpected delays in the bending process can create bottlenecks, disrupting the production flow and potentially reducing the overall output of the plant.
- Suboptimal resource allocation: without reliable time estimates, assigning operators or machines to tasks becomes less efficient, missing opportunities to align personnel experience with process complexity.

Conversely, the proposed system already provides significant economic value. The 54% reduction in MAE for the LW process, when compared to traditional engineering estimates, allows for tighter production schedules, reduced work-in-progress (WIP) inventory, and a lower risk of incurring penalties for late deliveries. This demonstrates that even with moderate upstream predictions, the integrated ML approach is a powerful tool for cost optimization.

From a cost-benefit perspective, strategies for enhancing model accuracy should be evaluated as business investments. Improving the

Table B.11
Configuration of the hyperparameters of the different experiments obtained for the Longitudinal Welding dataset.

Experiment	Model	Parameter	Parameter settings
None	GB	learning_rate	0.0146
		max_depth	5
		n_estimators	597
		subsample	0.8548
Actual value	GB	max_depth	7
		n_estimators	649
		learning_rate	0.0072
		subsample	0.6681
Predicted value	GB	max_depth	6
		n_estimators	982
		subsample	0.7597
		learning_rate	0.0143
Prediction error	LGBM	bagging_freq	3
		lambda_l2	0.3155
		min_child_samples	21
		lambda_l1	7.6414
		num_leaves	210
		bagging_fraction	0.9266
		feature_fraction	0.6820
Absolute prediction error	GB	learning_rate	0.0221
		max_depth	6
		subsample	0.8131
		n_estimators	479
Predicted value + Prediction error	GB	max_depth	4
		subsample	0.9626
		learning_rate	0.0491
		n_estimators	337
Predicted value + Absolute prediction error	RF	n_estimators	143
		min_samples_leaf	1
		max_features	sqrt
		max_depth	19
		min_samples_split	5
Actual value + Predicted value	GB	learning_rate	0.0482
		max_depth	7
		subsample	0.9470
		n_estimators	292
Actual value + Prediction error	XGB	alpha	0.0408
		booster	gbtree
		lambda	0.0001
		max_depth	3
Actual value + Absolute prediction error	GB	n_estimators	976
		learning_rate	0.0040
		max_depth	7
		subsample	0.7136
		max_features	
Actual value + Predicted value + Prediction error	GB	n_estimators	520
		subsample	0.7315
		learning_rate	0.0109
		max_depth	8
Actual value + Predicted value + Absolute prediction error	GB	subsample	0.8213
		n_estimators	834
		max_depth	8
		learning_rate	0.0345

model through advanced feature engineering or hyperparameter re-optimization represents a low-cost initiative with a potentially high return, primarily requiring computational resources and data science expertise.

However, the most significant leap in performance and cost reduction would likely come from addressing the root cause of the bending model's moderate accuracy: its reliance on manual data. Investing in automated data collection systems (such as sensors or machine logs) involves a higher initial capital expenditure. Nevertheless, the potential return on investment (ROI) is substantial. Such an investment would not only drastically improve prediction accuracy, thereby minimizing the previously mentioned costs associated with uncertainty, but could also enable real-time process monitoring, facilitate early fault detection, and ultimately pave the way for a more resilient and cost-effective production control system. Therefore, this study provides a quantitative baseline to support and justify future investments in the digitalization of the shop floor.

6. Conclusions and future work

This work introduces a machine learning-based system for lead time prediction in wind turbine tower manufacturing, focusing on bending and longitudinal welding (LW) operations. The results underscore three pivotal contributions:

1. Superior performance over engineering methods: the ML model for LW achieves a 54% reduction in MAE (2.03 vs. 11.36 h) and 74% lower RMSE (3.13 vs. 12.01 h) compared to traditional engineering estimates. This stark contrast highlights the limitations of conventional approaches, which rely on static formulas and outdated assumptions, and validates Ml's ability to capture complex, dynamic relationships in production data.
2. Explicit modelling of sequential dependency: while bending predictions exhibit moderate accuracy due to high manual process variability, their integration as inputs – specifically the predicted lead time and its associated error – significantly enhances LW lead time estimation. This demonstrates the critical importance of modelling inter-process dependencies to improve downstream predictive accuracy in sequential manufacturing.
3. Actionable insights and practical viability: the system's interpretability, enabled by SHAP analysis, moves beyond a "black box" approach by identifying that factors such as sheet thickness, operator experience, and upstream quality are critical drivers of LW lead times. These insights provide a clear, evidence -based guide for decision-makers to optimize resource allocation and prioritize process improvements. This, combined with the system's high computational efficiency (<1 ms per prediction), confirms its practical viability for near-real-time production control.

Beyond the specific application, this research presents a methodologically generalizable framework. The core concept of sequential predictive integration is transferable to other industries characterized by multi-stage production and high variability, such as aerospace component manufacturing, shipbuilding, or engineer-to-order (ETO) systems. By leveraging standard enterprise data sources (ERP, QMS), this approach provides a template for modelling operational interdependencies and improving planning accuracy in diverse and complex manufacturing environments.

In summary, this research provides empirical evidence that a sequential ML approach not only outperforms traditional engineering methods but also offers a scalable and interpretable framework for adaptive production control. The findings lay a quantitative foundation for industrial decision-makers, offering a clear roadmap to transition from static, heuristic-based planning to a dynamic, data-driven strategy that can unlock significant gains in efficiency and predictability in complex industrial settings.

Building upon these results, future research should aim to overcome the limitations identified throughout this study and further enhance the proposed framework. A primary avenue lies in the integration of automated data sources, particularly through the incorporation of IoT sensor data from machinery, such as energy consumption and vibration metrics, as well as computer vision systems for quality control. These technologies would not only reduce human error and bias in data collection but also enable real-time anomaly detection, which is particularly valuable in highly variable processes like bending. This evolution aligns with the principles of Industry 4.0 and the emerging, more human-centric vision of Industry 5.0. The latter seeks to create resilient and sustainable production systems by harmonizing technology with human expertise and environmental concerns (Guerrero et al., 2025). By enhancing data quality and enabling real-time monitoring, our framework could become a foundational component of such an advanced, sustainable operational management system.

Beyond data acquisition, the current framework could be expanded to model the entire production line, incorporating downstream operations such as flange fitting and circular welding. Developing a holistic, end-to-end predictive model would support global production flow optimization, improve delivery time forecasting, and enhance the management of bottlenecks, ultimately leading to a more resilient and responsive manufacturing system. Furthermore, the proposal is not applicable solely to wind turbine tower manufacturing, but rather to multiple other industrial setting that show significant interdependence between their processes.

Finally, there is substantial potential in advancing the predictive models through improved feature and model engineering. This includes the creation of interaction features – for example, linking operator experience with material complexity – and exploring advanced machine learning techniques such as semi-supervised or self-supervised learning. These approaches would allow the framework to leverage large amounts of unlabelled data, thereby increasing robustness and generalizability in scenarios with limited labelled datasets.

Funding

This research was co-funded by the European Regional Development Fund ERDF by means of the Interreg V-A Spain-Portugal Programme (POCTEP) 20142020, through the CIU3A project (reference 0754_CIU3A_5_A) and by the Ministry for Digital Transformation and Public Service, Agencia Red.es through the project INARI (2021/C005/00144319).

CRediT authorship contribution statement

Kenny-Jesús Flores-Huamán: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Antonio Lorenzo-Espejo:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **María-Luisa Muñoz-Díaz:** Validation, Supervision, Methodology, Investigation, Data curation. **Alejandro Escudero-Santana:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Appendix A. Hyperparameters of machine learning models

The table in this section showcases the hyperparameter configurations used during the training stage of the machine learning and deep learning models. Table A.10 presents the hyperparameters for decision trees and random forests, while additional hyperparameter configurations for other models will be detailed in subsequent sections.

Appendix B. Hyperparameter configuration for LW experiments

Table B.11 summarizes the best model and its associated configurations identified through the experiments conducted for the longitudinal welding process.

Data availability

The data that has been used is confidential.

References

Alenezi, A., Moses, S. A., & Trafalis, T. B. (2008). Real-time prediction of order flowtimes using support vector regression. *Computers & Operations Research*, *35*(11), 3489–3503. http://dx.doi.org/10.1016/j.cor.2007.01.026, URL: https://www.sciencedirect.com/science/article/pii/S0305054807000330.

Backus, P., Janakiram, M., Mowzoon, S., Runger, C., & Bhargava, A. (2006). Factory cycle-time prediction with a data-mining approach. *IEEE Transactions on Semiconductor Manufacturing*, *19*(2), 252–258. http://dx.doi.org/10.1109/TSM.2006.873400, URL: https://ieeexplore.ieee.org/document/1628987.

Bender, J., & Ovtcharova, J. (2021). Prototyping machine-learning-supported lead time prediction using AutoML. *Procedia Computer Science*, *180*, 649–655. http://dx.doi.org/10.1016/j.procs.2021.01.287, URL: https://linkinghub.elsevier.com/retrieve/pii/S1877050921003367.

Bender, J., Trat, M., & Ovtcharova, J. (2022). Benchmarking AutoML-supported lead time prediction. *Procedia Computer Science*, *200*, 482–494. http://dx.doi.org/10.1016/j.procs.2022.01.246, URL: https://linkinghub.elsevier.com/retrieve/pii/S1877050922002551.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*(10), 281–305, URL: http://jmlr.org/papers/v13/bergstra12a.html.

Bertolini, M., Mezzogori, D., Neroni, M., & Zammori, F. (2021). Machine learning for industrial applications: A comprehensive literature review. *Expert Systems with Applications*, *175*, Article 114820. http://dx.doi.org/10.1016/j.eswa.2021.114820, URL: https://linkinghub.elsevier.com/retrieve/pii/S095741742100261X.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. http://dx.doi.org/10.1023/A:1010933404324.

Burggräf, P., Wagner, J., Koke, B., & Steinberg, F. (2020). Approaches for the prediction of lead times in an engineer to order environment—A systematic review. *IEEE Access*, *8*, 142434–142445. http://dx.doi.org/10.1109/ACCESS.2020.3010050.

Carroll, J., Koukoura, S., McDonald, A., Charalambous, A., Weiss, S., & McArthur, S. (2019). Wind turbine gearbox failure and remaining useful life prediction using machine learning techniques. *Wind Energy*, *22*(3), 360–375. http://dx.doi.org/10.1002/we.2290, URL: https://onlinelibrary.wiley.com/doi/10.1002/we.2290.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/2939672.2939785, URL: https://doi.org/10.1145/2939672.2939785.

Council, G. W. E. (2024). *Global wind report 2024*. Brussels, Belgium: GWEC, URL: https://www.gwec.net/reports.

Fischetti, M., & Fraccaro, M. (2019). Machine learning meets mathematical optimization to predict the optimal production of offshore wind parks. *Computers & Operations Research*, *106*, 289–297. http://dx.doi.org/10.1016/j.cor.2018.04.006, URL: https://linkinghub.elsevier.com/retrieve/pii/S0305054818300893.

Flores-Huamán, K.-J., Escudero-Santana, A., Muñoz-Díaz, M.-L., & Cortés, P. (2024). Lead-time prediction in wind tower manufacturing: A machine learning-based approach. *Mathematics*, *12*(15), http://dx.doi.org/10.3390/math12152347, URL: https://www.mdpi.com/2227-7390/12/15/2347.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232, URL: http://www.jstor.org/stable/2699986.

Gao, Q., Wu, X., Guo, J., Zhou, H., & Ruan, W. (2021). Machine-learning-based intelligent mechanical fault detection and diagnosis of wind turbines. In B. Yang (Ed.), *Mathematical Problems in Engineering, 2021*, 1–11. http://dx.doi.org/10.1155/2021/9915084, URL: https://www.hindawi.com/journals/mpe/2021/9915084/.

Guerrero, B., Mula, J., & Poler, R. (2025). Sustainable operations management towards industry 5.0. *Dirección Y Organización*, 85–92. http://dx.doi.org/10.37610/85.692, URL: https://revistadyo.es/DyO/index.php/dyo/article/view/692.

Gyulai, D., Pfeiffer, A., Bergmann, J., & Gallina, V. (2018). Online lead time prediction supporting situation-aware production control. *Procedia CIRP*, *78*, 190–195. http://dx.doi.org/10.1016/j.procir.2018.09.071, URL: https://linkinghub.elsevier.com/retrieve/pii/S2212827118312563.

Gyulai, D., Pfeiffer, A., Nick, G., Gallina, V., Sihn, W., & Monostori, L. (2018). Lead time prediction in a flow-shop environment with analytical and machine learning approaches. *IFAC-PapersOnLine*, *51*(11), 1029–1034. http://dx.doi.org/10.1016/j.ifacol.2018.08.472, URL: https://linkinghub.elsevier.com/retrieve/pii/S2405896318316008.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with numpy. *Nature*, *585*(7825), 357–362. http://dx.doi.org/10.1038/s41586-020-2649-2, URL: https://doi.org/10.1038/s41586-020-2649-2.

Helbing, G., & Ritter, M. (2018). Deep learning for fault detection in wind turbines. *Renewable and Sustainable Energy Reviews*, *98*, 189–198. http://dx.doi.org/10.1016/j.rser.2018.09.012, URL: https://www.sciencedirect.com/science/article/pii/S1364032118306610.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. http://dx.doi.org/10.1109/MCSE.2007.55.

Kang, Z., Catal, C., & Tekinerdogan, B. (2020). Machine learning applications in production lines: A systematic literature review. *Computers & Industrial Engineering*, *149*, Article 106773. http://dx.doi.org/10.1016/j.cie.2020.106773, URL: https://linkinghub.elsevier.com/retrieve/pii/S036083522030485X.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems 30: annual conference on neural information processing systems 2017, December 4-9, 2017, long beach, CA, USA* (pp. 3146–3154). URL:.

Kim, G., & Hur, J. (2021). A short-term power output forecasting based on augmented naïve Bayes classifiers for high wind power penetrations. *Sustainability*, *13*(22), 12723. http://dx.doi.org/10.3390/su132212723, URL: https://www.mdpi.com/2071-1050/13/22/12723.

Kramer, K. J., Wagner, C., & Schmidt, M. (2020). Machine learning-supported planning of lead times in job shop manufacturing. In B. Lalic, V. Majstorovic, U. Marjanovic, G. von Cieminski, & D. Romero (Eds.), *Advances in production management systems. the path to digital transformation and innovation of production management systems* (pp. 363–370). Cham: Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-57993-7_41.

Lewis, C. (1982). Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. In *Butterworth scientific*, Butterworth Scientific, URL: https://books.google.es/books?id=t8W4AAAAIAAJ.

Lim, Z. H., Yusof, U. K., & Shamsudin, H. (2019). Manufacturing lead time classification using support vector machine. In H. Badioze Zaman, A. F. Smeaton, T. K. Shih, S. Velastin, T. Terutoshi, N. Mohamad Ali, & M. N. Ahmad (Eds.), *Advances in visual informatics* (pp. 268–278). Cham: Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-34032-2_25.

Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W., & Monostori, L. (2018). Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. *Procedia CIRP*, *72*, 1051–1056. http://dx.doi.org/10.1016/j.procir.2018.03.148, URL: https://linkinghub.elsevier.com/retrieve/pii/S2212827118303056.

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In *2008 eighth IEEE international conference on data mining* (pp. 413–422). http://dx.doi.org/10.1109/ICDM.2008.17.

Lorenzo-Espejo, A., Escudero-Santana, A., Muñoz-Díaz, M.-L., & Guadix, J. (2024). A machine learning-based system for the prediction of the lead times of sequential processes. In R. Rodríguez-Rodríguez, Y. Ducq, R.-D. Leon, & D. Romero (Eds.), *Enterprise interoperability x: enterprise interoperability through connected digital twins* (pp. 25–35). Cham: Springer International Publishing, http://dx.doi.org/10.1007/978-3-031-24771-2_3, URL: https://doi.org/10.1007/978-3-031-24771-2_3.

Lorenzo-Espejo, A., Escudero-Santana, A., Muñoz-Díaz, M.-L., & Robles-Velasco, A. (2022). Machine learning-based analysis of a wind turbine manufacturing operation: A case study. *Sustainability*, *14*(13), 7779. http://dx.doi.org/10.3390/su14137779, URL: https://www.mdpi.com/2071-1050/14/13/7779.

Loyola-González, O. (2019). Black-box vs. White-box: Understanding their advantages and weaknesses from a practical point of view. *IEEE Access*, *7*, 154096–154113. http://dx.doi.org/10.1109/ACCESS.2019.2949286.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 4768–4777). Red Hook, NY, USA: Curran Associates Inc..

McKinney, W. (2010). Data structures for statistical computing in python. In S. van der Walt, & J. Millman (Eds.), *Proceedings of the 9th python in science conference* (pp. 56–61). http://dx.doi.org/10.25080/Majora-92bf1922-00a.

Menendez-Roche, M. (2025). Spain electricity prices go up 2025. *Euro Weekly News*, URL: https://euroweeklynews.com/2025/01/11/spain-electricity-prices-go-up-2025/.

Modesti, P., Ribeiro, J. K., & Borsato, M. (2022). Artificial intelligence-based method for forecasting flowtime in job shops. *VINE Journal of Information and Knowledge Management Systems*, *54*(2), 452–472. http://dx.doi.org/10.1108/VJIKMS-08-2021-0146, URL: https://doi.org/10.1108/VJIKMS-08-2021-0146.

Mohsen, O., Mohamed, Y., & Al-Hussein, M. (2022). A machine learning approach to predict production time using real-time RFID data in industrialized building construction. *Advanced Engineering Informatics*, *52*, Article 101631. http://dx.doi.org/10.1016/j.aei.2022.101631, URL: https://linkinghub.elsevier.com/retrieve/pii/S1474034622000969.

Muñoz-Díaz, M.-L., Escudero-Santana, A., & Lorenzo-Espejo, A. (2024). Solving an unrelated parallel machines scheduling problem with machine- and job-dependent setups and precedence constraints considering support machines. *Computers & Operations Research*, *163*, Article 106511. http://dx.doi.org/10.1016/j.cor.2023.106511, URL: https://www.sciencedirect.com/science/article/pii/S0305054823003751.

Onaran, E., & Yanı k, S. (2020). Predicting cycle times in textile manufacturing using artificial neural network. In C. Kahraman, S. Cebi, S. Cevik Onar, B. Oztaysi, A. C. Tolga, & I. U. Sari (Eds.), *Intelligent and fuzzy techniques in big data analytics and decision making* (pp. 305–312). Cham: Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-23756-1_38.

Öztürk, A., Kayalı gil, S., & Özdemirel, N. E. (2006). Manufacturing lead time estimation using data mining. *European Journal of Operational Research*, *173*(2), 683–700. http://dx.doi.org/10.1016/j.ejor.2005.03.015, URL: https://www.sciencedirect.com/science/article/pii/S0377221705003358.

Park, H. (2018). Design and manufacturing of composite tower structure for wind turbine equipment. *IOP Conference Series: Materials Science and Engineering*, *307*, Article 012065. http://dx.doi.org/10.1088/1757-899X/307/1/012065, URL: https://iopscience.iop.org/article/10.1088/1757-899X/307/1/012065.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pérez-Cubero, E., & Poler, R. (2020). Aplicación de algoritmos de aprendizaje automático a la programación de órdenes de producción en talleres de trabajo: una revisión de la literatura reciente. *Dirección Organización*, (72), 82–94. http://dx.doi.org/10.37610/dyo.v0i72.588, URL: https://www.revistadyo.es/DyO/index.php/dyo/article/view/588.

Petrov, A. N., & Wessling, J. M. (2015). Utilization of machine-learning algorithms for wind turbine site suitability modeling in iowa, USA. *Wind Energy*, *18*(4), 713–727. http://dx.doi.org/10.1002/we.1723, URL: https://onlinelibrary.wiley.com/doi/10.1002/we.1723.

Pfeiffer, A., Gyulai, D., Kádár, B., & Monostori, L. (2016). Manufacturing lead time estimation with the combination of simulation and statistical learning methods. *Procedia CIRP*, *41*, 75–80. http://dx.doi.org/10.1016/j.procir.2015.12.018, URL: https://linkinghub.elsevier.com/retrieve/pii/S2212827115010975.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135–1144).

Rizzuto, A., Govi, D., Schipani, F., & Lazzeri, A. (2021). Lead time estimation of a drilling factory with machine and deep learning algorithms: A case study. In *Proceedings of the 2nd international conference on innovative intelligent industrial production and logistics - IN4PL* (pp. 84–92). INSTICC. SciTePress, http://dx.doi.org/10.5220/0010655000003062.

Rocha-Jácome, C., Hinojo-Montero, J. M., Guerrero-Morejón, K., Muñoz-Chavero, F., & González-Carvajal, R. (2025). High-precision non-contact online measurement and predictive analysis of geometric parameters in large industrial components. *Measurement*, *242*, Article 116126. http://dx.doi.org/10.1016/j.measurement.2024.116126, URL: https://www.sciencedirect.com/science/article/pii/S0263224124020116.

Ruschel, E., Rocha Loures, E. D. F., & Santos, E. A. P. (2021). Performance analysis and time prediction in manufacturing systems. *Computers & Industrial Engineering*, *151*, Article 106972. http://dx.doi.org/10.1016/j.cie.2020.106972, URL: https://linkinghub.elsevier.com/retrieve/pii/S0360835220306446.

Sainz, J. A. (2015). New wind turbine manufacturing techniques. *Procedia Engineering*, *132*, 880–886. http://dx.doi.org/10.1016/J.PROENG.2015.12.573.

Schuh, G., Gützlaff, A., Sauermann, F., Kaul, O., & Klein, N. (2020). Databased prediction and planning of order-specific transition times. *Procedia CIRP*, *93*, 885–890. http://dx.doi.org/10.1016/j.procir.2020.04.026, URL: https://linkinghub.elsevier.com/retrieve/pii/S221282712030576X.

Schuh, G., Gützlaff, A., Sauermann, F., & Theunissen, T. (2020). Application of time series data mining for the prediction of transition times in production. *Procedia CIRP*, *93*, 897–902. http://dx.doi.org/10.1016/j.procir.2020.04.054, URL: https://linkinghub.elsevier.com/retrieve/pii/S2212827120306272.

Schuh, G., Prote, J.-P., Hünnekes, P., Sauermann, F., & Stratmann, L. (2019). Impact of modeling production knowledge for a data based prediction of transition times. In F. Ameri, K. E. Stecke, G. von Cieminski, & D. Kiritsis (Eds.), *Advances in production management systems. production management for the factory of the future* (pp. 341–348). Cham: Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-30000-5_43.

Schuh, G., Prote, J.-P., Luckert, M., & Sauermann, F. (2018). Determination of order specific transition times for improving the adherence to delivery dates by using data mining algorithms. *Procedia CIRP*, *72*, 169–173. http://dx.doi.org/10.1016/j.procir.2018.03.236, URL: https://linkinghub.elsevier.com/retrieve/pii/S2212827118304050.

Shapley, L. S. (1953). 17. a value for n-person games. In H. W. Kuhn, & A. W. Tucker (Eds.), *Contributions to the theory of games, volume II* (pp. 307–318). Princeton: Princeton University Press, http://dx.doi.org/10.1515/9781400881970-018.

Sousa, A., Ferreira, L., Ribeiro, R., Xavier, J., Pilastri, A., & Cortez, P. (2022). Production time prediction for contract manufacturing industries using automated machine learning. In I. Maglogiannis, L. Iliadis, J. Macintyre, & P. Cortez (Eds.), *Artificial intelligence applications and innovations*: *Vol. 647*, (pp. 262–273). Cham: Springer International Publishing, http://dx.doi.org/10.1007/978-3-031-08337-2_22, URL: https://link.springer.com/10.1007/978-3-031-08337-2_22.

Steinberg, F., Burggaef, P., Wagner, J., & Heinbach, B. (2022). Impact of material data in assembly delay prediction—a machine learning-based case study in machinery industry. *International Journal of Advanced Manufacturing Technology*, *120*(1), 1333–1346. http://dx.doi.org/10.1007/s00170-022-08767-3.

Stetco, A., Dinmohammadi, F., Zhao, X., Robu, V., Flynn, D., Barnes, M., Keane, J., & Nenadic, G. (2019). Machine learning methods for wind turbine condition monitoring: A review. *Renewable Energy*, *133*, 620–635. http://dx.doi.org/10.1016/j.renene.2018.10.047, URL: https://linkinghub.elsevier.com/retrieve/pii/S096014811831231X.

Szaller, Á., & Kádár, B. (2021). Effect of lead time prediction accuracy in trust-based resource sharing. *IFAC PapersOnLine*, *54*(1), 1126–1131. http://dx.doi.org/10.1016/j.ifacol.2021.08.132, URL: https://linkinghub.elsevier.com/retrieve/pii/S2405896321008934.

Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. In *Special issue on smart manufacturing*: *Journal of Manufacturing Systems*, In *Special issue on smart manufacturing*: *48*, 157–169. http://dx.doi.org/10.1016/j.jmsy.2018.01.006.URL: https://www.sciencedirect.com/science/article/pii/S0278612518300062,

Treiber, N., Heinermann, J., & Kramer, O. (2016). Wind power prediction with machine learning. In J. Lässig, K. Kersting, & K. Morik (Eds.), *Computational sustainability. studies in computational intelligence*: *Vol. 645*, (pp. 13–29). Cham, Switzerland: Springer.

Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R., & Fortin, A. (2020). Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. *Journal of Intelligent Manufacturing*, *31*(6), 1531–1558. http://dx.doi.org/10.1007/s10845-019-01531-7.

Wang, C., & Jiang, P. (2019). Deep neural networks based order completion time prediction by using real-time job shop RFID data. *Journal of Intelligent Manufacturing*, *30*(3), 1303–1318. http://dx.doi.org/10.1007/s10845-017-1325-3.

Waskom, M. L. (2021). Seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. http://dx.doi.org/10.21105/joss.03021.

Zhu, H., & Woo, J. H. (2021). Hybrid NHPSO-JTVAC-SVM model to predict production lead time. *Applied Sciences*, *11*(14), 6369. http://dx.doi.org/10.3390/app11146369, URL: https://www.mdpi.com/2076-3417/11/14/6369.