Enhancing Lead Time Prediction in Wind Tower Manufacturing: A ML Approach Compared to Traditional Engineering Models

Kenny-Jesús Flores-Huamán[1](✉), Antonio Lorenzo-Espejo[1], María-Luisa Muñoz-Díaz[2], and Alejandro Escudero-Santana[1,2](✉)

[1] 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
{kflores1,alorenzo,alejandroescudero}@us.es

[2] Centro de Innovación Universitario de Andalucía, Alentejo y Algarve (CIU3A), 41011 Seville, Spain
mmunozd@us.es

Abstract. The efficient estimation of lead times in manufacturing processes is crucial for optimizing production and reducing costs. In wind tower manufacturing, particularly in the longitudinal welding operation, accurate lead time prediction is essential for maintaining smooth workflows and meeting tight delivery schedules. Traditional engineering methods, which rely on analytical models and heuristics, have been widely used to estimate welding times. However, these methods often fail to account for the complexity and variability of real-world conditions, such as equipment wear, environmental factors, and production bottlenecks. In recent years, machine learning (ML) techniques have emerged as powerful tools for predictive modelling, leveraging large datasets and learning from patterns within the data. This study compares traditional engineering methods with ML approaches for lead time prediction in longitudinal welding, using data from a wind tower manufacturing plant. The results demonstrate that ML models, particularly Gradient Boosting, outperform traditional methods in accuracy and flexibility, offering significant potential for improving operational efficiency in the wind tower industry.

Keywords: Lead Time Prediction · Machine Learning models · Longitudinal welding · Manufacturing optimization · Traditional Methods

1 Introduction

The efficient estimation of lead times in manufacturing processes is crucial for optimizing production and reducing costs. In wind tower manufacturing, particularly in the longitudinal welding operation, accurate lead time prediction is essential for maintaining smooth workflows and meeting tight delivery schedules. Traditional engineering methods, which rely on analytical models and heuristics, have been widely used to estimate welding times. However, these methods often fail to account for the complexity and variability of real-world conditions, such as equipment wear, environmental factors, and production bottlenecks.

E. Barbadilla-Martín et al. (Eds.): CIO 2025, LNDECT 276, pp. 321–326, 2025.
https://doi.org/10.1007/978-3-032-10126-6_51

In recent years, machine learning (ML) techniques have emerged as powerful tools for predictive modelling, leveraging large datasets and learning from patterns within the data. Unlike traditional methods, ML models can adapt to changes in production dynamics, offering the potential for more accurate and flexible lead time estimations. Despite the promising applications of ML in manufacturing, there remains a need for a direct comparison between traditional engineering methods and machine learning approaches in real-world settings.

Currently, factories employ techniques such as direct formulation or linear programming to estimate lead times. However, these methodologies have certain limitations. Direct formulation, although widely used, may not realistically reflect the times or costs involved in processes. On the other hand, linear programming, while useful, may not be fully updatable over time, making it difficult to apply in dynamic production environments. Therefore, the ability of machine learning algorithms to analyse large sets of historical data, identify patterns, and make accurate predictions makes them invaluable for optimizing lead times and improving operational efficiency in the wind tower industry.

Notably, the only works that address wind turbine manufacturing from a production standpoint are those by Sainz [1], who describes the manufacturing process and several improvement steps based on increased automation; Park et al. [2], who analyse composite wind turbine towers from a design and manufacturing perspective; and Alorenzo (2022), who applies an ML-based approach to the bending process of wind turbine tower manufacturing, highlighting the influence of worker experience and age due to the manual nature of the operation.

In addition, Kenny et al. [3] present an ML-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, emphasizing the importance of integrating ML into production planning in the wind tower industry.

Despite these advancements, there is a notable gap in the literature regarding direct comparisons between traditional engineering methods and ML approaches in the context of wind tower manufacturing. Most existing studies focus on applying ML models without directly contrasting them with traditional methods, limiting the ability of industry professionals to make informed decisions about which approach is more suitable for their specific needs. This study aims to address this gap by comparing lead time predictions for longitudinal welding in wind tower manufacturing using traditional engineering methods and ML models developed by our team. By evaluating the performance of these two approaches, we seek to provide insights into their respective advantages and limitations, offering practical recommendations for their use in industrial applications. The remainder of this work is structured as follows: Sect. 2 describes the methodology, including the welding process, data used, ML models, and evaluation metrics; Sect. 3 presents the experimental results and discussion; and Sect. 4 concludes the study.

2 Methodology

2.1 Description of the Longitudinal Welding Process

The longitudinal welding (LW) process is a critical step in the manufacturing of wind turbine towers. It involves welding together the two edges of a bent steel plate to form a fully enclosed cylindrical or conical ferrule, ensuring its structural integrity. Depending on the plate's thickness and diameter, slight deformations such as ovalization may occur after welding. Once completed, the ferrule is ready for storage or transportation until it is required for further assembly into larger tower sections. Before final integration, these sections undergo additional welding and surface treatments to enhance their durability and prepare them for on-site installation.

2.2 Data Employed

This study relies on data collected from the manufacturing plant, where approximately 875 tower sections were produced between 2022 and 2024, comprising over 7,300 ferrules. The data were extracted from multiple company databases, including the Enterprise Resource Planning (ERP) system and the Quality Management System (QMS), and then integrated to create a comprehensive dataset containing all relevant variables for analysis.

The dataset for the longitudinal welding process includes several explanatory variables categorized into four groups: historical lead time records from upstream processes, contextual information (e.g. machine ID and operator identifiers), quality control reports, and bending lead time predictions with their absolute error values. Given the diverse data sources, ensuring data consistency was a crucial step before model development.

To address potential inconsistencies and missing values, a thorough preprocessing phase was implemented. This included outlier detection and treatment, as well as imputation of missing values where necessary. Numerical attributes were normalized to facilitate comparison across models, while categorical variables were encoded using techniques such as One-Hot Encoding. Additionally, during feature selection, columns with more than 35% missing values were removed, as they exhibited low representativity and could introduce biases, ultimately affecting the model's performance and generalizability.

2.3 Machine Learning Models and Evaluation Metrics

In this study, ten machine learning (ML) models were implemented to predict the lead time of the longitudinal welding process. These models were selected based on their effectiveness in handling structured manufacturing data and their ability to capture complex relationships among variables. The models were categorized into four main families: linear models, kernel-based methods, tree-based models, and artificial neural networks.

Linear models, including Ridge, Lasso, and Elastic Net, were incorporated as baselines due to their interpretability and ability to manage collinearity through regularization.

Support Vector Regression (SVR) was considered separately under kernel-based methods due to its ability to model nonlinear relationships by transforming data into higher-dimensional spaces.

Tree-based models were divided into single-tree methods, such as Decision Tree Regressor (DT), and ensemble methods, including Random Forest (RF), Gradient Boosting (GB), LightGBM (LGBM), and XGBoost (XGB). While DT recursively partitions the feature space, ensemble methods enhance prediction accuracy by aggregating multiple weak learners, with boosting techniques iteratively correcting prior errors.

Additionally, Artificial Neural Networks (ANNs), specifically the Multi-Layer Perceptron (MLP) Regressor, were implemented to model highly complex relationships using deep learning techniques.

To ensure a fair comparison, all models underwent the same preprocessing steps, including feature selection, categorical variable encoding, and numerical attribute normalization. Given that the engineering department calculates schedules in sections rather than individual ferrules, and each sample represents a ferrule, the training and test sets were divided using the hold-out technique with a 75%–25% ratio. This division was performed by tower section, ensuring that all ferrules from a given section were assigned to either the training or test set, thereby guaranteeing that evaluations were conducted on unseen data.

During training, models were initially evaluated without hyperparameter tuning to establish baseline performance. Once the functionality of these base models was verified, Random Search optimization combined with cross-validation was applied to fine-tune their hyperparameters. This approach was chosen due to its balance between efficiency and simplicity, outperforming Grid Search by avoiding exhaustive hyperparameter exploration and being computationally less expensive than Bayesian Optimization while still delivering competitive results.

Finally, the best-performing tuned model was retrained using the full training and validation dataset and subsequently used to predict the lead time for test set instances. If the predictive accuracy met acceptable thresholds, the model was further retrained on the complete dataset before being deployed for production use.

To evaluate model performance, Root Mean Squared Error (RMSE) was prioritized, as it clearly measures prediction error and penalizes larger deviations, making it effective for identifying outliers and simplifying interpretation. RMSE was selected for its ability to represent errors in the same units as the data and its sensitivity to large errors. In cases of tied RMSE values across models, Mean Absolute Error (MAE) was used as a secondary criterion to compare the model predictions with engineering estimates, offering a simpler measure of the average error.

3 Experimental Results and Discussion

The results in Table 1 show that Gradient Boosting (GB) outperforms all other models in every evaluation metric, achieving the lowest MAE (0.53), RMSE (0.82), and MAPE (26.98). While LightGBM and XGBoost performed similarly in some areas, GB consistently provided the best results across all key metrics. This makes Gradient Boosting the most suitable model for predicting lead times in the longitudinal welding dataset and the benchmark for comparison with engineering estimates.

Table 1. Results obtained by each model with the best configuration through random search for the longitudinal welding dataset.

Model	MAE	RMSE	MAPE	R^2
DT	0.61	0.98	30.84	0.53
Enet	0.67	0.99	34.96	0.53
GB	**0.53**	**0.82**	**26.98**	**0.67**
Lasso	0.67	0.98	34.91	0.53
LGBM	**0.53**	0.87	27.03	0.64
MLP	0.61	0.93	31.29	0.58
RF	0.58	0.89	29.77	0.62
Ridge	0.67	0.98	34.72	0.53
SVR	0.6	0.92	31.33	0.59
XGB	0.54	0.87	27.51	0.64

Once the model was obtained, it was compared to the engineering time calculations used in the factory. The engineering method calculates the total segment time based on various input characteristics, such as material properties, process parameters among other attributes. Since our data set is based on individual ferrule times, predictions were made for each ferrule within a segment and then summed to obtain the total time.

The Table 2 presents a comparison of the engineering method and machine learning predictions with actual production times, while Fig. 1 displays the absolute percentage errors relative to the real times for a random selection of 25 samples. It can be observed that the machine learning predictions closely align with the actual production times, with a smaller margin of error compared to the engineering calculations. This suggests that ML predictions could be useful not only for more accurately estimating welding times but also for detecting anomalies in case of process delays.

Table 2. Comparison of engineering method and 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

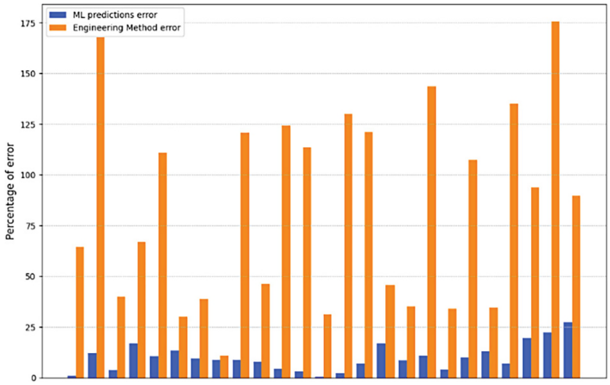

Fig. 1. Absolute percentage errors relative to the actual times for a random selection of 25 samples.

4 Conclusions

This study demonstrates the superiority of machine learning (ML) models, particularly Gradient Boosting (GB), over traditional engineering methods in predicting lead times for longitudinal welding in wind tower manufacturing. GB consistently outperforms other models across key metrics (MAE, RMSE, MAPE), with ML predictions aligning more closely with actual production times. The adaptability of ML models to dynamic environments and their ability to detect process anomalies highlight their potential to enhance operational efficiency. Future research should refine these models and explore their application to other renewable energy manufacturing processes.

Acknowledgments. Grant 2021/C005/00144319 funded by the Ministry for Digital Transformation and Public Service, Agencia Red.es and by the European Union NextGenerationEU/PRTR.

References

1. Sainz, J.A.: New wind turbine manufacturing techniques. Procedia Eng. **132**, 880–886 (2015). https://doi.org/10.1016/J.PROENG.2015.12.573
2. Park, H.: Design and manufacturing of composite tower structure for wind turbine equipment. IOP Conf. Ser.: Mater. Sci. Eng. **307**, 012065 (2018). https://doi.org/10.1088/1757-899X/307/1/012065
3. Lorenzo-Espejo, A., et al.: Machine learning-based analysis of a wind turbine manufacturing operation: a case study. Sustainability. **14**(13), 7779 (2022). https://doi.org/10.3390/su14137779
4. Flores-Huamán, K.-J., et al.: Lead-time prediction in wind tower manufacturing: a machine learning-based approach. Mathematics. **12**, 2347 (2024). https://doi.org/10.3390/math12152347