

Article

Machine Learning-Based Analysis of a Wind Turbine Manufacturing Operation: A Case Study

Antonio Lorenzo-Espejo , Alejandro Escudero-Santana * , María-Luisa Muñoz-Díaz  and Alicia Robles-Velasco

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; alorenzo@us.es (A.L.-E.); mmunozd@us.es (M.-L.M.-D.); arobles2@us.es (A.R.-V.)

* Correspondence: alejandroescudero@us.es; Tel.: +34-95-448-60-42

Abstract: This study analyzes the lead time of the bending operation in the wind turbine tower manufacturing process. Since the operation involves a significant amount of employee interaction and the parts processed are heavy and voluminous, there is considerable variability in the recorded lead times. Therefore, a machine learning regression analysis has been applied to the bending process. Two machine learning algorithms have been used: a multivariate Linear Regression and the M5P method. The goal of the analysis is to gain a better understanding of the effect of several factors (technical, organizational, and experience-related) on the bending process times, and to attempt to predict these operation times as a way to increase the planning and controlling capacity of the plant. The inclusion of the experience-related variables serves as a basis for analyzing the impact of age and experience on the time-wise efficiency of workers. The proposed approach has been applied to the case of a Spanish wind turbine tower manufacturer, using data from the operation of its plant gathered between 2018 and 2021. The results show that the trained models have a moderate predictive power. Additionally, as shown by the output of the regression analysis, there are variables that would presumably have a significant impact on lead times that have been found to be non-factors, as well as some variables that generate an unexpected degree of variability.

Keywords: machine learning; regression; process control; wind power; lead time; bending; worker experience



check for updates

Citation: Lorenzo-Espejo, A.; Escudero-Santana, A.; Muñoz-Díaz, M.-L.; Robles-Velasco, A. Machine Learning-Based Analysis of a Wind Turbine Manufacturing Operation: A Case Study. *Sustainability* **2022**, *14*, 7779. <https://doi.org/10.3390/su14137779>

Academic Editors: Marija Bogataj and Josefa Mula

Received: 9 May 2022

Accepted: 23 June 2022

Published: 26 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Wind power was the largest-growing renewable energy source in 2019, exceeding its electricity generation in the previous year by 160 TWh [1]. Considering the surge in electricity prices (reportedly, increasing 34.3% in only a year for Spanish households [2]), and as the energy demand grows globally, the current goal is to raise power output, which can be achieved through more efficient exploitation of wind currents. There are currently two strategies to do so: increasing the wind capturing area by using larger rotor blades and installing the wind generators at higher altitudes, where wind flows in a much more consistent manner. Both strategies require the manufacturing of increasingly taller towers; however, these structures endure a complex, mainly non-automated, and costly manufacturing process. Therefore, product innovations must be paralleled by improvements in the efficiency and management of their manufacturing process.

As stated above, the manufacturing of towers for wind turbines is not automated. That is, it requires a constant involvement of workers, be it for processing the pieces (usually aided by a semi-automatic tool) or for supervising the operation of a machine. Additionally, the pieces that compose a wind tower, steel cylinders referred to as ferrules, are heavy (up to 25 tons) and voluminous, which significantly complicates the transportation and processing of the parts; these two reasons pose a considerable drawback to the manufacturing of wind towers: the lead times of the processes show a tremendous variability, and this issue strongly hinders the controlling and planning of the manufacturing process, which avoids

detecting deviations in production. At the same time, production plans are hardly precise and usually just loosely followed. If the time needed to process parts was accurately defined, the time windows utilized in production planning could be significantly slimmed.

The goal of this study is to analyze and predict the completion time of the main process of wind tower manufacturing: bending. Through this operation, steel plates are inserted between rolls and transformed into cylinders, which then need to be welded in order to fully close the ferrules. In order to examine the data and predict the processing times, several machine learning algorithms have been tested using a validation set, selecting the multivariate Linear Regression and M5P methods as the most adequate techniques for this analysis. Additionally, multiple experience-related variables have been obtained and added to the dataset in a second stage; this has been done to compare the performances of the models with and without these variables and to determine if such additional information can improve their predictive power. Finally, the effect of the input variables on the bending lead time has been analyzed by examining the importance assigned to each factor in the models.

Several machine learning approaches have been conducted in the wind industry, but they usually do not involve the manufacturing process; most academic works focus on the operation of wind farms, particularly on the monitoring of the condition of the wind turbines [3]. A case study has been used to validate this methodology. The data used correspond to the operation of a Spanish wind tower manufacturer from 2018 to 2021.

Therefore, this article addresses three areas regarding machine learning and wind power that, based on a thorough review of the literature (part of which is shown in Section 3), are receiving insufficient research attention: the prediction of process lead times using machine-learning-based approaches, wind turbine manufacturing, and the effect of experience-related variables on workers' efficiency. In this work, through the use of a real-world case study, these three research gaps are addressed, using an intuitive and user-friendly machine learning workbench, WEKA. The case presented in this article could potentially serve as a basis for similar studies conducted in wind turbine manufacturing and comparable industries. With this goal, the methodology followed has been presented in a general way, and later particularized for the case study at hand. Furthermore, the use of the WEKA software makes the approach followed in this study more accessible to other researchers or industry managers, independent of their programming knowledge.

The remainder of the article is structured as follows. Next, a more comprehensive description of the bending operation and its relevance in the wind turbine tower manufacturing process is presented in Section 2, followed by a brief overview of machine learning methods and their application to the wind power industry found in Section 3. The methodology utilized in this study is thoroughly described in Section 4; Section 5 summarizes and discusses the results of the analyses; finally, Section 6 presents the conclusions of the study.

2. Wind Turbine Tower Manufacturing: The Bending Process

In order to provide a context for this study, a brief overview of the bending process in wind tower manufacturing is presented next.

The steel structures that support the generator and rotor of a wind turbine are called towers and can weigh from 60 to 300 tons. Towers of off-shore generators (wind turbines installed in large bodies of water) are generally more massive than those of on-shore generators. Wind towers are assembled on-site by joining steel cylinders or cones, called sections, together. There are at least three sections in a wind tower: the bottom, mid and top sections. If there are more than three of them, the remaining ones are considered to be mid-sections. The sections are bolted to each other using flanges that have been previously welded to their bases; these sections are the final product of a wind tower manufacturing plant. They are built using ferrules, smaller cylinders, or cones that are welded together. There are five main processes that characterize the wind turbine tower manufacturing process: bending, longitudinal welding, flange fitting, ferrule fitting and circular welding. The process starts with rectangular steel plates, which may have been processed before to

adjust their dimensions or bevels, which are bent into a ring shape (bending). In order to bend the plates, they are inserted through rollers lengthwise, slowly giving them a circular shape. Next, the gap between the edges of the bent plate is welded (longitudinal welding). The now-closed ferrules are transported to a table in which they are joined by weld spots to the flanges (flange fitting). Flanges are steel joints used to connect one tower section with another through bolts. Therefore, not all ferrules need to be fitted with flanges, only the top and bottom ones of each section. All ferrules are then taken to long hangars in which they are fitted next to each other and fixed in place through spot welding (ferrule fitting). Finally, the union between the ferrules is secured by welding throughout the perimeter of each junction (circular welding), giving way to a semi-finished section, pending surface treatment and other auxiliary processes.

Bending is a critical step of the wind turbine tower manufacturing process, for several reasons:

- The bending operation is one of the first carried out in the manufacturing process and, additionally, constitutes the bottleneck of the process. Having a bottleneck at the start of the production process can lead to non-desired idle times at downstream operations. Since the technical specifications require bending to be carried out before any other of the main operations in the manufacturing process, the focus must be placed on ensuring a continuous workflow at the bending station.
- Minor defects caused during bending can significantly slow down posterior operations; these faults derive in delays, especially during the fit-up process. In this operation, two ferrules that have already undergone the bending and longitudinal welding processes are set up next to each other and their relative position is fixed by spot welding; thus, it is at this point that imperfections caused by bending, mainly non-circularity (ovality) and lack of flatness, are easily identified.
- Reworks are costly and time-consuming. When a major bending-related fault is detected in the downstream processes, the part must be transported back to the bending station; it must be kept in mind that the layout of these types of plants is optimized with the goal of reducing the difficulty of the movements required in the normal production flow. Both the configuration of the layout and the large size of the machines explain why reworks usually involve moving the parts between buildings, using multiple transports, and occupying several employees. The direct outcome is that these reworks cause great delays in production, mainly for two reasons: (a) if the fault is detected in or after the fit-up process, the production of the entirety of the section must be stopped, since the ferrules are either partially or fully welded to each other. The affected operation cannot continue until the faulty ferrule has been reprocessed and moved back to said station; (b) whether the fault has been detected before or after the fit-up process, the defective ferrule is assigned top priority, moved to the bending station and reprocessed before any other ferrule in the backlog, in order to attempt to meet the customer's delivery deadline. Additionally, if a different product is being processed at the bending station, it is likely that the machine or tools configurations are to be modified, deriving from the corresponding set-up times.

The three reasons given above reflect the importance of the efficiency of the bending operation for the overall performance of the manufacturing process. From an organizational, as opposed to technical, standpoint, there are two major aspects that could help improve the efficiency of the bending process: strict control of the anomalies during production that generate faulty parts, and the improvement of the accuracy of the production schedules. Controlling the production anomalies would allow detecting a higher percentage of faulty parts just as the bending operation concludes, thus avoiding most defective units being passed downstream. On the other hand, designing accurate production plans would facilitate a continuous workflow at the bending station, reducing the idle times in posterior processes. Pérez-Cubero and Poler [4] highlight the importance of accounting for processing lead time variations on job-shop production scheduling. Additionally, given the relevance

of reworks, it is essential that the schedules of the bending station can be recalculated with accuracy, so as to reduce the impact on the overall lead times of the rest of the units.

3. Machine Learning and Wind Turbine Tower Manufacturing

In his seminal work, Samuel [5] coined the concept of machine learning; the author believed that programming computers to learn from experience could alleviate much effort devoted to tasks that were trivial but still involved a learning process. Samuel's work led the way for an immense amount of research on machine learning techniques and applications. Jordan and Mitchell [6] claim that the focus of this field of study is centered around two main research lines: finding out how to construct computer systems that learn from experience, and identifying the fundamental laws governing learning systems.

The use of machine learning techniques for lead time prediction has not received sufficient research attention from academia. A systematic review performed by Kang, Catal and Tekinerdogan [7] highlights quality-related problems as the most frequently addressed a problem in the literature on machine learning applications to production lines, with a clear predominance over lead-time prediction, yield improvement, waste reduction or even preventive maintenance; this could be due to the scarcity of data regarding times and other operational and process data, however, this obstacle is likely to disappear as industries advance in their digitization process, thus increasing their sensorization and the availability of data.

Most of the literature regarding lead time prediction with machine learning techniques addresses the forecasting of the total flow time or completion time, that is, the time elapsed between the arrival of a part and the fulfillment of all the operations required in its manufacturing specifications; these works utilize historical data of the process times, and focus on organizational variables, such as the precedence between operations or the dispatch rules used, to train the models and generalize for future instances. For example, Backus et al. [8] make use of three different machine learning algorithms to predict the cycle time of product lots in a factory: clustering, regression trees and the K-nearest-neighbors model; they also propose a hybrid method that first groups the lots based on their similarity using a clustering algorithm, and then generates a regression tree for each of those clusters. Similarly, Öztürk et al. [9] utilize a regression tree model to estimate the flowtimes in different make-to-order shop configurations, testing their approach through a computer simulation.

Along the same lines, Alenezi et al. [10] posit a support vector regression model for the real-time prediction of order flowtimes in a make-to-order manufacturing environment; while these works focus on organizational variables, such as the precedence between operations or the dispatch rules used, the approach presented in this article is centered on the variables affecting the process itself.

Alternatively, Wang and Jiang [11] present a deep neural network that predicts order completion times based on order information and real-time production data obtained from a RFID system embedded in the manufacturing process. The authors show that their deep-belief network approach outperforms other neural networks training methods, such as back-propagation, multi-hidden layers back-propagation and combining principal components analysis and back-propagation.

As shown, most of the research relevant to this article addresses the prediction of the completion time; this is logical, since completion time is a key parameter, necessary to provide customers with delivery deadlines and to assess the overall performance of the manufacturing process; however, some researchers have started to shift their focus towards individual process lead-time estimation. Gyulai et al. [12] present a data analytics tool for "situation aware" production control. Their tool utilizes a closed-loop control in order to deploy and update online a digital data twin [13]; this tool is based on accurate simulation models of manufacturing systems, which allow performing prospective simulations that forecast deviations in production. Their models are enabled by process lead time predictions based on machine learning algorithms, as presented in [14,15]. The results presented by the authors indicate that traditional analytical techniques are outperformed by the machine

learning methods that they had tested for the lead time predictions. After comparing the results provided by multiple machine learning algorithms, the authors selected the Random Forest method for their system.

These works make use of two groups of variables for the training and prediction of the lead times: static data regarding product features, extracted from the enterprise resource planner (ERP), and dynamic event-based logs drawn from the manufacturing execution systems (MES). While their approach allows for a real-time prediction and control, the work presented in this article is directed mainly toward the characteristics of the part being processed. There are still contextual variables, such as the personnel operating the machine, the shift in which the part is manufactured or the machine itself, but all of these variables can be previously set and tested by the manager. The goal with this is that predictions of the process lead time can be obtained with enough time in advance to serve as input for the scheduling.

Regarding the case study presented in this work, it is noticeable that wind turbine manufacturing has received very little attention in the academic literature. Most of the works regarding machine learning techniques and wind power focus on the operational phase of the wind turbines. Mainly, three research lines can be outlined: the prediction of the electrical output, which can be based only on historical data [16], on wind velocity records [17,18]—which can also be predicted using Machine Learning and Artificial Intelligence techniques as hybrid models [19], fuzzy logic [20], Deep Learning [21] or ensemble methods [22]—or on multiple environmental variables [23,24]; the creation of assistant systems for the design and control of wind turbines [25] and wind farms [26–28]; and the development of smart and knowledge-based maintenance system for the wind turbines, mainly focused towards fault classification [29,30], anomaly detection [31] and remaining useful life (RUL) estimation [32–34]. For an exhaustive review of the machine-learning-based approaches to wind turbine condition monitoring, see Ref. [3].

However, reportedly, only Sainz [35] addresses the wind turbine manufacturing process from an operational standpoint. The author describes different steps of the process and indicates several automation technologies that would give way to an increase in the production capacity of the industry. Additionally, no works have been found that focus on the manufacturing of wind turbine towers, or that apply machine learning and artificial intelligence techniques to the wind turbine manufacturing process control and planning.

There has also been a wide research stream on the effect of aging on worker productivity, but most of it has been conducted from a theoretical or even medical standpoint. Previous works suggest that workforce aging could hinder productivity, mostly due to age-related deterioration of some cognitive and physical abilities [36]. Most importantly, age has been found to increase the potential to be involved in occupational accidents in industrial settings, which is particularly notable in the case of workers aged 50 or more [37]. A method to assess worker performance is the Work Ability Index (WAI) [38], based on a questionnaire administered to the employees including data regarding absenteeism, health, and self-evaluation; however, Theppitak et al. [39] did not find a significant effect of WAI on task performance for industrial workers. In fact, the authors found a higher impact of the workers' ages on task performance than on the WAI. Additionally, Kumudini et al. [40] found that age is more detrimental to the WAI of average-performance industrial workers than to the higher-performing ones. Nevertheless, the study presented in this paper focuses on the effect of age and experience from a purely operational standpoint, that is, focusing exclusively on the influence of these input factors on the lead time of the bending operation in wind turbine tower manufacturing.

After a study of the relevant literature, it can be seen that, by creating machine learning models that are capable of analyzing high volumes of data and of identifying underlying patterns between the relevant variables, the accuracy of the lead time predictions can be significantly improved. It must be noted that in many companies, especially those that are less digitized, such as the case studied in this work, these predictions are made just by averaging the historical records, or even just based on personal experience. Therefore, in

order to design a somewhat accurate production plan, managers must use large margins of error for the lead times; in this case, these margins of error can significantly increase the idle times at the bending station and, particularly, at the downstream processes. By reducing the difference between the actual lead times and their prediction, the sought-after continuous flow at the bottleneck of the process can be achieved. Additionally, after a major fault and its mandatory rework, the recalculated schedules would be more accurate and, thus, would allow the plant to retrieve its original production goals faster. Finally, the machine-learning-based predictions can also serve as the basis of an anomaly control system. Traditionally, these systems resort to sensorization, such as recording the vibrations of the machine [41], to assess whether the equipment is working towards its expected performance or not; however, once again, in industries with low degrees of digitization, these signals are usually not available. For these cases, the mismatch between the expected lead time and the actual lead time could be used as an indicator of the performance of the machine and the quality of the part being manufactured.

To sum up, the main contribution of the article is two-fold. On the one hand, a novel, machine-learning approach to the prediction of the lead time of a process in the wind turbine manufacturing industry is presented in this study. On the other hand, a previously unexplored inclusion of age and experience-related variables into lead-time prediction modules is conducted.

4. Materials and Methods

The methodology behind this study is based on the analysis and prediction of the lead times of the wind turbine tower manufacturing process. Given that this analysis is structured around machine learning regression models, it must follow the typical steps in a data analytics pipeline: data gathering and processing; variable selection; model selection; and regression analysis and results interpretation. In this case, the interpretation of the results is two-fold: not only will the predictions of the models be analyzed, but also, given that some of the machine learning models used in this work provide deeper insight into the underlying patterns of the bending process, a study of the impact of several input factors on the lead time will be conducted.

In an attempt to provide a framework for the analysis of bending or similar operations in the wind power manufacturing industry using machine learning models, the methodology is first presented in a general manner and then particularized to the case study in Section 5.

4.1. Data Gathering and Processing

One of the most work-intensive steps of a machine learning-based approach is data acquisition. In industrial settings, most of the data come from either the company's Enterprise Resource Planner (ERP) or from a Manufacturing Execution System (MES).

Once the data are obtained, it is essential to preprocess the database in order to increase its quality; this implies removing outliers, applying filters (such as normalization) if necessary, obtaining new compound variables, or discarding unfeasible values with the help of the domain knowledge.

4.2. Variable Selection

The data utilized for the regression analyses performed in this study can be split into two categories: input variables and output variables. Input or independent variables are the factors considered to affect the lead time of the process, which is the output or dependent variable of the models. Specifically, the dependent variable represents the time required to process a ferrule in the bending station (measured in hours). It must be noted that ferrules are the smallest unit involved in the manufacturing process after the bending operation. Therefore, predicting this lead time allows a more in-depth planning of the bending operation and, as a result, of the following processes.

The variables listed below have been suggested as potential determinants of the bending time after an exploration of the data and interviews with workers of the plant analyzed later in the case study; these variables have been divided into three categories: operational variables, technical/product-related variables and experience-related variables.

Operational variables:

- Work shift during which the operation was completed.
- Personnel, i.e., the worker that finalized the operation.
- Machine used for the processing of the part.

Technical/product-related variables:

- Position of the ferrule in the section, starting from the bottom position (1) and increasing.
- Position of the section in the tower. Generally, towers have three sections (bottom, mid and top). The thickness of the ferrule, measured in millimeters (See Figure 1).
- Length of the plate that is curved into a cylinder, measured in millimeters; this dimension corresponds to the perimeter of the ferrule base.
- Width of the plate that is curved into a cylinder, measured in millimeters; this magnitude corresponds to the height of the ferrule.
- Steel plate yield strength (in N/mm^2) for a nominal thickness of 16 mm or less. Yield strength is one of the most important properties of structural steels.
- Steel plate toughness designation, measured with the Charpy impact test; this test provides the maximum impact strength absorbed by a notched test steel sample without fracturing. The test results are usually codified as subgrades [42]: JR, J0, J2, NL, K2.
- Steel plate normalization, i.e., whether the steel plate that is to be curved into a ferrule has received a normalization treatment, which consists of heating the material and allowing it to cool in order to increase the toughness of the steel.

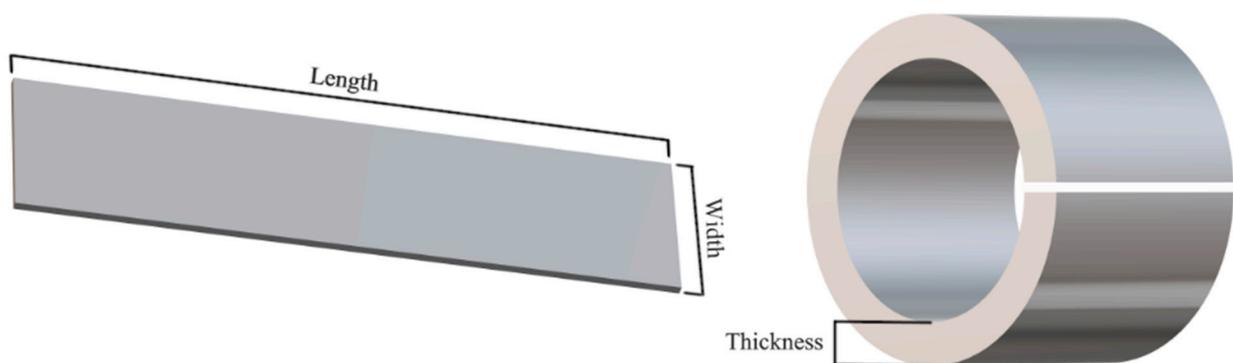


Figure 1. Length, width and thickness, indicated in a schematic drawing of the raw plate and resulting ferrule.

Experience-related variables:

- Age of the worker completing the operation.
- Experience of the worker at the bending station, that is, the number of years for which the worker has been developing the bending operation.
- Experience of the worker at the plant/industry, that is, the experience of the worker not only at the bending station itself, but in the wind turbine manufacturing sector.
- Number of bending operations previously performed by the worker, that is, the experience of the employee at the bending station measured as the number of bending operations performed before the analyzed instance.
- Worker frequency at the bending station, which measures the average number of bending operations performed daily by a worker in a certain period of time. Five time periods have been considered: the time passed since the worker's first operation,

the last 180 days (six months), the last 90 days (three months), the last 60 days (two months) and the last 30 days (one month).

By simply observing the definition of these input variables, one can expect to find significant correlations among them. Therefore, a correlation analysis has been performed in order to determine the optimal set of variables to utilize in the regression.

4.3. Regression Analysis

As stated previously, the main goal of this study is to achieve accurate predictions of the lead time of the bending process in wind tower manufacturing and to analyze correlations between these completion times and several factors that come into play in this process; this can be done by building a regression model through machine learning algorithms.

To do so, a Java-based machine learning software called WEKA [43], developed at the University of Waikato in New Zealand, has been used in its version 3.9.5. WEKA allows us to efficiently apply some of the most widely used machine learning algorithms in a very intuitive manner, owing to its graphical user interface. WEKA incorporates the Knowledge Flow tool, which facilitates the development of flowcharts depicting the machine learning approach used in the analysis.

The WEKA workbench has been previously used in the wind power research field by Mansour et al. [44] to predict wind speed in different locations in order to assess their wind energy productivity; and by Joshuva et al. [45] for the development of a vibration-based fault classification system for wind turbine blades.

There are three fundamental concepts in a machine learning approach: task, learning algorithm and model [46]. The task represents the problem that is to be solved through the approach. The model is the mathematical tool used to perform the task, and it is trained using a learning algorithm.

Regression is one of the most frequently encountered machine learning tasks. In a regression task, the goal is obtaining a model able to predict a real value based on a series of factors and their corresponding weights; these weights are determined with a machine learning algorithm, based on historical data. Therefore, regression is an example of a predictive (as opposed to descriptive) task, since the aim is ultimately to produce predictions of unknown instances based on recorded data. It is also a supervised (as opposed to unsupervised) task since it requires the data to be “labelled”, that is, the instances must have values of the target variable.

4.3.1. Model Selection

The machine learning model utilized to obtain the prediction can be just as important as the quality of the training data or the selected variables. Different models behave in diverse manners regarding training time, test time and, most importantly, predictive power. In order to find an adequate model for a given task, an initial exploration of the performance of various algorithms can be carried out using a validation set. Using a validation set implies dividing the available data set into three: a training set, a validation set and a hold-out test set; this way, the models can be trained and validated using part of the data, and their performance is then evaluated using the instances in the test set, which the model has not “seen”. A widely used 80–20% split is proposed for this work, meaning that 80% of the instances compose the training and validation set and the remaining 20% form the test set.

Furthermore, instead of just using one training set and one validation set, a tenfold cross-validation approach has been followed for the model selection; this entails dividing the training and validation dataset into ten splits and performing said number of iterations of the analysis. In each iteration, nine splits are used to train the model and the remaining one is used as the validation set. For each experiment, a different split is used to evaluate the trained model, until the ten splits have been used as validation sets; this method is a way to avoid overestimating the predictive power of a model if a “lucky” training-validation

split is produced. The described process is repeated ten times with each split, reaching a total of 100 runs.

The approach, starting from the dataset upload and up to the generation of the results, has been modeled in the WEKA Knowledge Flow tool, which facilitates the development of flowcharts depicting the structure and sequence of the machine learning analysis. The pipeline of the cross-validation analysis is shown in Figure 2.

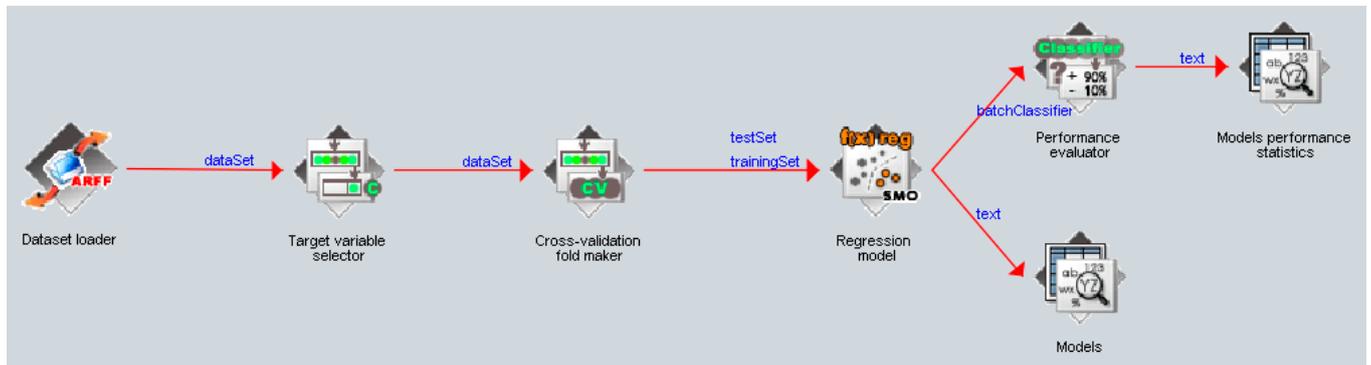


Figure 2. Cross-validation pipeline in WEKA Knowledge Flow.

An ARFF (a specific format for data input in WEKA) file is first uploaded using the *ArffLoader* module. In this case, it contains the training and validation dataset. Next, the loaded set is processed through a *ClassAssigner* module, which allows the user to select which variable of the input data is to be predicted. Consecutively, filters can be utilized before generating the training and test datasets.

For the cross-validation experiment, a *CrossValidationFoldMaker* module, which allows dividing the set into as many equal-sized splits as desired. A tenfold cross-validation configuration is selected, which feeds the following module with 10 pairs of training-validation sets.

Each regression algorithm is applied using the corresponding module in the Classifiers group (regression can be regarded as a classification task in which the target variable is a real variable). WEKA algorithms provide users with some parametrization options, which vary depending on the selected model.

The output of this analysis consists of the structure and parameters of the trained models and the performance evaluation of the model, which are saved as “.txt” files. It must be noted that, while the model information is directly produced by the classifier object, a *ClassifierPerformanceEvaluator* (included in the Evaluation group) must be introduced to generate the model performance statistics.

The output of the cross-validation process is used to determine the adequate models for the problem at hand. WEKA provides several measures of the fitness of the model: correlation coefficient, mean absolute error, root mean squared error, relative absolute error and root relative squared error. In this work, the metric used to decide between one algorithm or another is the Root Mean Squared Error (RMSE), mainly because it penalizes large errors by squaring the difference between the predicted and the actual value. In order to fulfil the applications discussed previously in this paper, the predictions produced by the model should particularly be accurate when dealing with extreme values, which can severely affect production plans if they are not predicted. RMSE is calculated as follows:

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

where N is the number of instances in the test set, $y(i)$ is the i -th observation, and $\hat{y}(i)$ is its corresponding prediction.

The tested algorithms have been chosen so as to include different sorts of models: a multivariate Linear Regression model, two geometrical instance-based models (K-nearest neighbors, IbK in WEKA; and Support Vector Machines, SMOreg in WEKA), a regression tree (REPTree), a tree ensemble (Random Forest), a model tree (M5P) and a simple neural network (Multilayer Perceptron).

4.3.2. Model Implementation

Once a model is selected, its performance can be evaluated in the WEKA workbench using a test set. Figure 3 depicts the pipeline of the application of a regression algorithm to a training-test split:

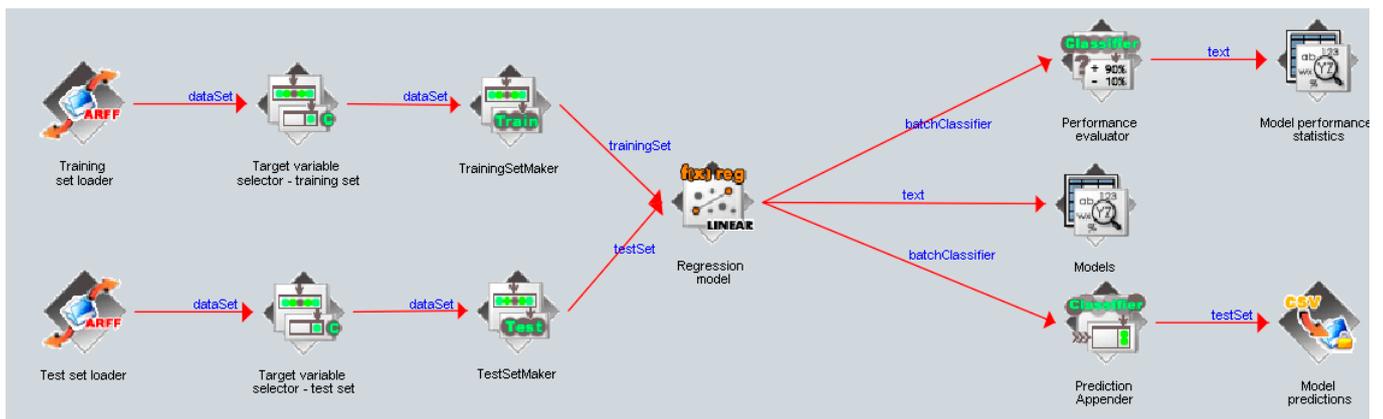


Figure 3. Regression pipeline in WEKA Knowledge Flow.

While they share a similar structure, there are several differences between the cross-validation pipeline shown in Figure 2 and the one depicted in Figure 3. Firstly, the latter includes two data sources instead of one: a source for the training set and a different one for the test set. Both sets are fed to a *TargetVariableSelector* and then to a training and test set maker, respectively. The data in the training set are used to learn the parameters that characterize the regression model. The performance of the trained model is then evaluated using the test set; this evaluation consists of predicting the target variable values for the instances in the test set, which have not been used for the training of the model, and comparing them to the actual values; this is a way to ensure that the model is not overfitting the training data and can perform accurate predictions of instances with different characteristics to those used in the learning algorithm. For this experiment, all instances in the training and validation set are used to train the model, and the remaining 20% of the total data compose the test set.

Additionally, in this case, aside from the structure and parameters of the trained models and the performance evaluation of the model, it is interesting to produce the predictions and errors of the test instances; this is done by adding a *PredictionAppender* module that converts the test results to an Excel file.

5. Results

In this section, the results of the machine learning regression analyses of the case study are presented and discussed. First, additional context regarding the case study particular situation is shown first. Next, the results of the model selection step are shown, and the resulting algorithms are applied to the case study.

5.1. Case Study Description

As previously mentioned, the proposed approach has been tested using the case of a Spanish wind turbine tower manufacturer. The data were collected from various databases of the company's ERP; these different sources were merged so the resulting database would encompass the variables discussed above. The dataset includes information from the nearly

900 tower sections manufactured in the plant from March 2018 to February 2021, which are composed of over 7400 ferrules.

In this case, most of the information regarding lead times and machine use is entered by the plant workers in the midst of the operation. In order to ensure that the data accurately depict the functioning of the plant, the dataset has been preprocessed by removing outliers and infeasible values. Instances that show bending operations performed in less than 20 min or more than 5 h have been eliminated, since they are considered to be errors in the employees' recording of the data; these instances represent only 2.23% of the dataset.

With regards to the input variables values, the following aspects must be noted:

- The possible ferrule positions in the records of the plant operation range from the bottom position (1) to the highest ferrule position recorded (16).
- As mentioned earlier, generally, towers have three sections (bottom, mid and top); however, the data contain examples of towers with up to six sections.
- The studied plant has three different work shifts: morning, afternoon, or night. Additionally, it must be noted that employees periodically rotate their assigned shifts. As a result, the variability introduced by the workers and by the shifts is not expected to be confounded.
- Eighteen different workers have performed the bending operation during the period analyzed.
- There are two bending machines at the plant, of the same model.
- Two values of the steel plate yield strength are found in this study, 355 N/mm² and 455 N/mm².
- Regarding the steel plate toughness designation, the subgrades presented in Table 1 have been found in the dataset:
- Table 2 summarizes the distribution of the experience-related variables.
- Table 3 summarizes the distribution of the thickness, length, width and lead time values of the instances of the dataset. Additionally, a histogram of the lead time distribution is shown in Figure 4.

Table 1. Charpy impact test subgrades.

Charpy Impact Test		
Subgrade	Impact Strength (J)	Test Temperature (°C)
JR	27 J	20
J0	27 J	0
J2	27 J	−20
NL	27 J	−50
K2	40 J	−20

Table 2. Experience-related variables distribution.

Variable	Minimum	Maximum	Mean	Standard Deviation
Worker age (in years)	28	57	36.47	6.39
Experience at station (in years)	0	2.88	1.13	0.80
Experience in industry/plant	0	4.458	2.17	1.09
Number of previous bending operations	0	1239	421.83	333.57
Worker frequency—global	0	6.00	1.04	0.46
Worker frequency—180 d	0	1.84	0.98	0.47
Worker frequency—90 d	0	2.34	1.09	0.45
Worker frequency—60 d	0	2.85	1.16	0.49
Worker frequency—30 d	0	3.13	1.28	0.56

Table 3. Thickness, length, width and lead time variables distribution.

Variable	Minimum	Maximum	Mean	Standard Deviation
Thickness (in mm)	11.7	70	25	10.92
Length (in mm)	7324	24,953	13,217.86	2333.61
Width (in mm)	915	3058	2833.93	347.91
Lead time (in hours)	0.334	4.94	1.62	0.56

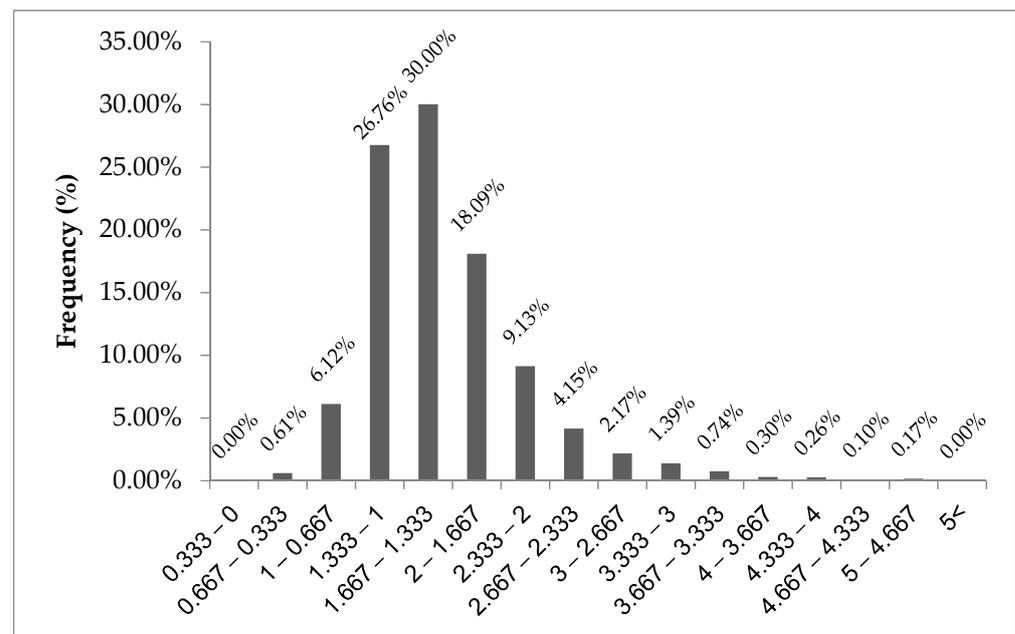
**Figure 4.** Dataset lead time histogram.

Table 3 and Figure 4 provide insight into the lead times of the bending operation. Just under 75% of the recorded bending operations lasted between one and two hours. Of the remaining 25%, only 6.63% occur in less than an hour: 18.41% of the operations are performed in more than two hours, with some lasting up to 5 h; these are the instances in which the model should prove its predictive power in order to be used as an accurate forecasting tool for production planning and control. For this reason, as mentioned in Section 4.3.1, the RMSE metric has been used as the deciding metric in the model selection step in order to minimize the prediction errors of these extreme values.

The potential correlations between the previously described input factors have been analyzed. The Pearson correlation coefficients for each pair of numeric variables are shown in Table 4. Additionally, the correlation between each input variable and the output variable (bending lead time) is analyzed.

The Pearson correlation coefficients shown in Table 4 are significant, with a 0.05 significance level, save for the coefficients written in italics. The coefficients higher than 0.3 (a standard threshold above which a correlation can be considered of moderate strength), are presented in bold. The correlations between the thickness and length variables and the thickness and width measures show the highest coefficients amongst the non-experience-related variables. Particularly, the thickness of the plate appears to be indirectly correlated with its width ($\rho = -0.439$). On the other hand, the thickness of the plate is directly correlated with its length ($\rho = 0.342$). Additionally, the thickness variable is strongly correlated to two experience-related variables. Thus, for the regression analyses presented later in this paper, the thickness variable has not been considered in order to avoid misinterpreting the outputs of the analyses, even if the predictive power of the model could have been increased with its inclusion.

Table 4. Pearson correlation coefficients for each pair of numeric variables.

Variables	Ferrule Pos.	Thickness	Length	Width	Worker Age	Exp. (Station)	Exp. (Sector)	Number of Operations	Global Freq.	180 d Freq.	90 d Freq.	60 d Freq.	30 d Freq.	Lead Time
Section pos.	−0.108	0.240	0.184	0.152	−0.031	0.235	0.181	0.251	0.092	0.102	0.043	0.016	0.013	0.074
Ferrule pos.	-	−0.232	−0.141	0.043	−0.019	−0.066	−0.044	−0.049	0.055	0.013	0.010	0.021	0.019	0.021
Thickness	-	-	0.342	−0.439	−0.066	0.390	−0.281	0.375	−0.084	0.064	0.004	0.036	−0.044	0.135
Length	-	-	-	−0.096	−0.073	0.277	−0.214	0.277	0.055	0.166	0.048	−0.011	−0.042	0.103
Width	-	-	-	-	0.023	−0.033	−0.025	−0.033	0.055	0.018	0.022	0.023	0.017	0.010
Worker age	-	-	-	-	-	−0.157	0.035	−0.121	0.027	0.030	−0.023	−0.005	0.021	0.001
Exp. (station)	-	-	-	-	-	-	0.740	0.977	−0.070	0.672	0.478	0.319	0.159	0.040
Exp. (sector)	-	-	-	-	-	-	-	0.746	0.079	0.534	0.412	0.332	0.251	−0.107
Number of Operations	-	-	-	-	-	-	-	-	0.066	0.709	0.517	0.367	0.216	0.061
Global freq.	-	-	-	-	-	-	-	-	-	0.099	0.198	0.292	0.380	0.021
180 d freq.	-	-	-	-	-	-	-	-	-	-	0.840	0.675	0.452	0.073
90 d freq.	-	-	-	-	-	-	-	-	-	-	-	0.894	0.654	−0.002
60 d freq.	-	-	-	-	-	-	-	-	-	-	-	-	0.812	−0.050
30 d freq.	-	-	-	-	-	-	-	-	-	-	-	-	-	−0.090

Regarding the experience-related variables, Table 4 shows considerably high values of the pairwise correlation coefficients; this was to be expected, given the definition of the variables, but should be taken into account in the design of the regression experiments. Thus, the following variables have also been discarded from the analyses: experience at the station, number of operations, global frequency, frequency in the previous 180 days, frequency in the previous 90 days and frequency in the previous 60 days.

Based on the results of this preliminary correlation analysis, two sets of input variables are considered for the regression analyses:

- Without experience variables: includes the section position, ferrule position, length, width, shift, personnel, machine, yield strength, toughness and normalization variables.
- With experience variables: includes the same variables as the previous set, plus the worker age, experience in the sector/plant and 30-day frequency variables.

Neither of the two variable configurations present conflicts regarding correlation: the correlation coefficients of all the pairs of input variables in each set are under 0.3.

5.2. Model Selection Results

As explained above, the selection of the applied models has been carried out based on the results of a cross-validation experiment; these results have been evaluated using the following metrics: correlation coefficient between predicted and actual values; mean absolute error; root mean squared error; relative absolute error; and training and evaluation runtime. As specified, minimizing the RMSE metric is selected as the decision criterion. The results of the cross-validation test are shown in Table 5: the average values of the metrics and their standard deviation (in brackets) across the hundred iterations are shown for each algorithm.

Table 5. Cross-validation results.

Algorithm	Variable Set	Correlation Coefficient	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Training Time	Testing Time
M5P	w/o Experience	0.3915 (0.0365)	0.3611 (0.0144)	0.5166 (0.0265)	88.4095 (1.9483)	0.3063 (0.0272)	0.0004 (0.0005)
	w/Experience	0.4600 (0.0437)	0.3470 (0.0141)	0.4987 (0.0268)	84.9607 (2.5687)	0.3616 (0.0089)	0.0003 (0.0005)
K-nearest neighbors (IbK)	w/o Experience	0.2701 (0.0429)	0.4404 (0.0177)	0.6336 (0.0310)	107.8816 (4.3267)	0.0003 (0.0004)	0.1036 (0.0093)
	w/Experience	0.3025 (0.0412)	0.4502 (0.0181)	0.6536 (0.0297)	110.2860 (4.5027)	0.0003 (0.0005)	0.0908 (0.0087)
REPTree	w/o Experience	0.3856 (0.0372)	0.3653 (0.0140)	0.5234 (0.0248)	89.4346 (2.2966)	0.0146 (0.0069)	0.0001 (0.0003)
	w/Experience	0.4468 (0.0435)	0.3497 (0.0143)	0.5062 (0.0262)	85.6314 (2.5570)	0.0149 (0.0014)	0.0001 (0.0003)
Linear Regression	w/o Experience	0.3439 (0.0358)	0.3693 (0.0136)	0.5267 (0.0259)	90.4232 (1.6455)	0.0268 (0.0042)	0.0003 (0.0005)
	w/Experience	0.3975 (0.0393)	0.3606 (0.0140)	0.5148 (0.0262)	88.2896 (1.9916)	0.0285 (0.0194)	0.0005 (0.0006)
SVM (SMOreg)	w/o Experience	0.3377 (0.0357)	0.3593 (0.0148)	0.5343 (0.0277)	87.9632 (1.8150)	49.8787 (3.4219)	0.0004 (0.0005)
	w/Experience	0.3837 (0.0393)	0.3513 (0.0153)	0.5246 (0.0282)	86.0024 (2.1449)	52.8269 (2.2717)	0.0005 (0.0007)
Random Forest	w/o Experience	0.3566 (0.0366)	0.3860 (0.0146)	0.5513 (0.0255)	94.5422 (3.0341)	1.0257 (0.0665)	0.0283 (0.0070)
	w/Experience	0.4500 (0.0382)	0.3598 (0.0140)	0.5152 (0.0252)	88.1021 (2.8555)	0.1163 (0.0196)	0.0034 (0.0012)
Multilayer Perceptron (Neural network)	w/o Experience	0.4060 (0.0209)	0.3813 (0.0401)	0.5366 (0.0254)	93.2534 (8.4230)	72.8065 (0.5206)	0.0015 (0.0047)
	w/Experience	0.4740 (0.0437)	0.3496 (0.0187)	0.5013 (0.0267)	85.5503 (3.0486)	205.4821 (3.3804)	0.0093 (0.0080)

Table 5 offers a clear conclusion regarding the performance with the two sets of input variables: the inclusion of the experience variables significantly increases the predictive power of the models. The average improvement in the correlation coefficient amounts to 6.04%. Additionally, the M5P algorithm proves to be the dominant method in all of the predictive metrics for the sets with the experience-related variables, except for the correlation coefficient. In particular, the M5P model reaches a correlation coefficient between the predicted and actual values of 39.15% and 46% without and with the experience variables, respectively. It must be noted that the Multilayer Perceptron Neural Network obtains a higher correlation coefficient than the M5P model in both cases, with correlation coefficients of 40.6% and 47.4%. In any case, in line with the decision criteria outlined previously, the M5P algorithm presents the lowest RMSE amongst all the models, using the datasets with and without the experience-related variables (0.4987 and 0.5166, respectively).

As a result of this cross-validation study, two machine learning algorithms are used for the regression task suggested in this work: a multivariate Linear Regression algorithm and the M5P method, developed by Wang and Witten [47] and based on the work by Quinlan [48], which combines decision trees with Linear Regression models. On the one hand, the M5P method shows the lowest RMSE of the tested model and, thus, can be expected to produce a lesser degree of large errors; this indicates that the model might perform better with extreme values.

On the other hand, interpreting the results of a model tree is not an easy task, given the multitude of decisions at the nodes and of models at the leaves. Therefore, and since its metrics are not excessively distant from the M5P algorithm metrics, the multivariate Linear Regression model has also been selected to ease the interpretation of the results. The Linear Regression algorithm is a fast method for obtaining models without extensive parametrizations that can produce accurate predictions.

5.3. Model Implementation Results

As explained in the previous section, the Linear Regression and M5P algorithms have been chosen for this study. The models are now tested using an 80/20 training-test split, as stated in the methodology section. In this case, there are over 1450 operations in the hold-out test set; these experiments have been conducted for the input variables set including the experience variables, and without them.

The Linear Regression algorithm is applied by selecting the *LinearRegression* module found in the Classifiers group. The WEKA workbench offers some parametrization options: the algorithm has been set to eliminate collinear attributes and to perform a selection, using the M5 method, of the attributes that are to be considered in the regression model.

Next, the M5P method has also been applied to the 80/20 split; this can be done in the Knowledge Flow models by substituting the *LinearRegression* module for the M5P module. The M5P algorithm constructs a decision tree where the leaves correspond to a set of Linear Regression models, also known as a model tree; this tree is capable of dealing with numeric attributes both on its decision nodes and on its leaves, thus using the traditional decision-model structure of classification algorithms but allowing numeric attributes to be predicted as the target variable, as required in a regression task. The model tree created by the M5P algorithm divides the dataset using a splitting method that minimizes the variation between the instances allocated to the same subset. Each leaf of the tree contains a Linear Regression model that uses the data in its subset to predict the target variable value for the evaluated instances that reach said leaf after going through the tree's decision nodes; the algorithm also includes a pruning method that simplifies the branches of the decision tree as long as the expected adjusted error at the resulting leaves decreases. When an instance is fed to the trained model, it arrives at one of the tree's leaves through attribute-based decisions at the tree's nodes. Once there, the value of the target variable is predicted utilizing the corresponding leaf's Linear Regression model.

The main performance statistics for the multivariate Linear Regression and M5P models are presented in Table 6. Additionally, the predictions of the test set have been examined and the percentages of instances for which the models have produced predictions deviating in less than 10, 15 and 30 min are included.

Table 6. Evaluation of the performance of the models.

Experiment	Input Variables Set	Correlation Coefficient	Mean Abs. Error (h)	Root Mean Squared Error	Relative Abs. Error (%)	% of Accurate Predictions		
						Error below 10' (%)	Error below 15' (%)	Error below 30' (%)
Linear Regression	w/o Experience	0.3365	0.3685	0.5262	90.31	31.91	47.11	76.89
Linear Regression	w Experience	0.3413	0.3681	0.5254	90.12	32.67	47.52	76.82
M5P	w/o Experience	0.3952	0.3619	0.5145	88.66	34.86	49.31	75.92
M5P	w Experience	0.4992	0.3363	0.4848	82.38	36.86	50.07	80.54

Overall, the results show a better performance of the M5P method over the Linear Regression model, as expected in view of the cross-validation results. There are moderate correlations between the predicted and actual times for both methods and datasets, with the M5P coefficients being higher. The mean absolute error of the predictions is less than a minute lower for the M5P method than for the Linear Regression model when using the

dataset without the experience-related variables. The inclusion of the experience-related variables does not cause a significant improvement in any of the metrics for the experiments carried out using the Linear Regression Model; however, the results show an increase of 10 percentage points of the correlation coefficient when adding the experience-related variables to the M5P model. Similarly, the mean absolute error is reduced in over 1.5 min if the complete dataset is used.

The relative absolute error represents the percentage error reduction of applying each method compared to predicting the lead time of every test instance as the mean lead time of the complete training dataset. Without the experience-related variables, the Linear Regression approach shows a 9.09% error reduction, while this value increases to 12.39% in the case of the M5P method. By adding the experience-related variables, these values are augmented to 11.3% and 17.62%, respectively.

Additionally, the M5P algorithm produces a higher percentage of “accurate” predictions than the Linear Regression approach in any of the proposed thresholds (10, 15, or 30 min). Once again, the use of the M5P method with the experience-related variables shows the highest accuracy in any of the thresholds. Nevertheless, it must be noted that the accuracy of each experiment can see changes in each threshold: for example, the M5P model without experience-related variables produces more accurate predictions with a margin of error of 10 and 15 min than the Linear Regression approach with any of the datasets; however, the latter proves more accurate than the former if the threshold is increased to 30 min.

Finally, the RMSE value is lower for the M5P model than for the multivariate Linear Regression model, as expected given the cross-validation results, in the comparison with both datasets; this suggests that the M5P produces fewer large errors than the Linear Regression model, which was the main goal of the regression analysis. To delve deeper into the fitness of each model, the predicted and actual values of each experiment have been plotted in Figure 5 (Linear Regression without experience-related variables), 6 (Linear Regression with experience-related variables), 7 (M5P without experience-related variables), and 8 (M5P with experience-related variables). The graphs plot the actual lead time values of the observation in the x -axis, while the corresponding predicted values are shown in the y -axis. The orange dashed line represents the line with slope 1, that is, where the perfect predictions would be located. The blue dotted line represents a linear trendline of the observations, showing the tendency of the predictions as the actual values change. By observing Figures 5 and 6, just small changes can be found between the Linear Regression approach without the experience variables and with them.

Figure 7 shows that the trend of the predictions moves closer to the line with slope 1, which, while not necessarily indicating a smaller error, suggests a better performance of the M5P model without experience-related variables. Furthermore, Figure 8 reveals that the predictions of the M5P model without experience-related variables seem to be even more accurate, as indicated by the results shown in Table 6.

Figures 5–7 suggest that the corresponding models tend to underestimate the predicted values of the bending lead time as it increases; this can be observed in the right-most points, which are further away from the line with slope 1 than those with lower actual values. For lead times over three hours, the models’ predictions are significantly inferior to the actual values, with over two hours of difference in the most extreme cases; this effect seems to be diminished in Figure 8, corresponding to the M5P model without experience-related variables. To further analyze the behavior of the models with this sort of instances, Table 7 shows the percentage of instances exceeding 2 hours in actual lead time, for which its predictions differ in less than 10, 15 and 30 min of the actual value, respectively. There are 252 of the 1454 instances in the hold-out test set that present an actual lead time longer than two hours.

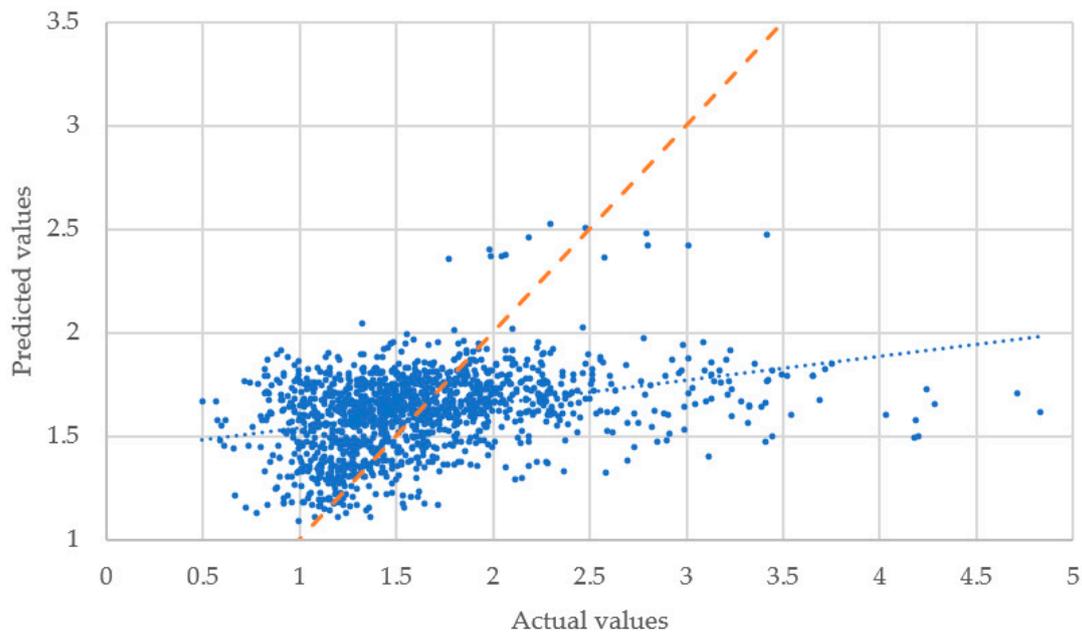


Figure 5. Actual vs. predicted values—Linear Regression without experience-related variables.

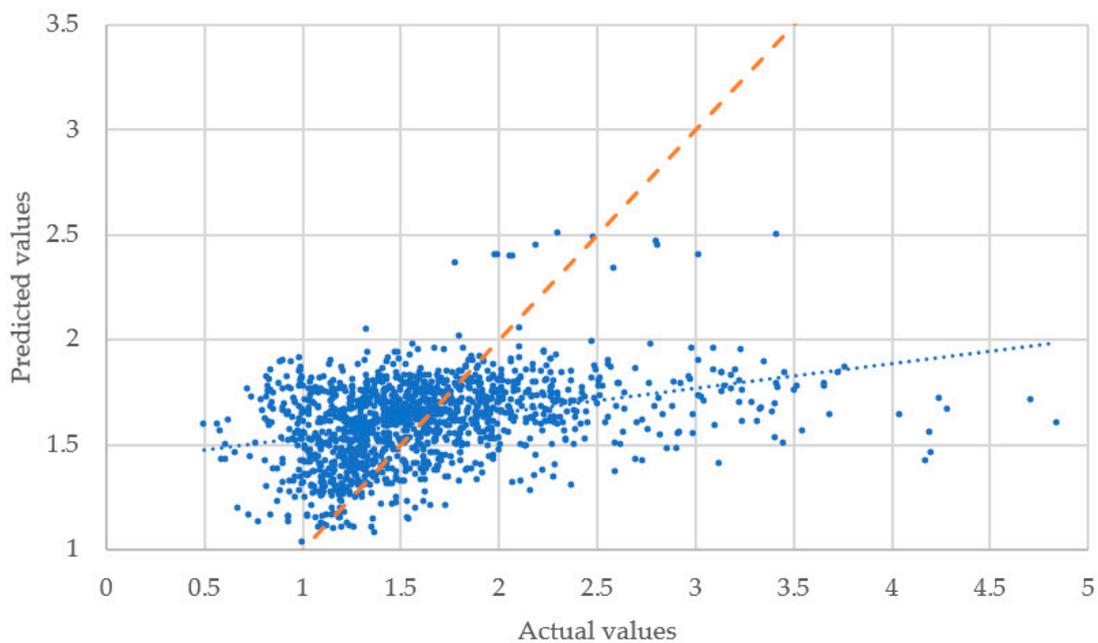


Figure 6. Actual vs. predicted values—Linear Regression with experience-related variables.

Once again, the M5P model with the experience variables shows superior performance when predicting the lead times of the extreme values; these results are encouraging, given that 41% of the instances over two hours can be predicted with less than 30 min of error. It must be remembered that almost 75% of the instances in the entire dataset range between 1 and 2 h, and thus the model should be able to predict the lead time for such instances with high accuracy; however, the critical aspect of the analysis is that the models accurately forecast the lead times for the more “uncommon” instances. Figure 9 shows a plot of the predicted and actual values for the M5P model with experience-related variables for the instances with an actual lead time over 2 h.

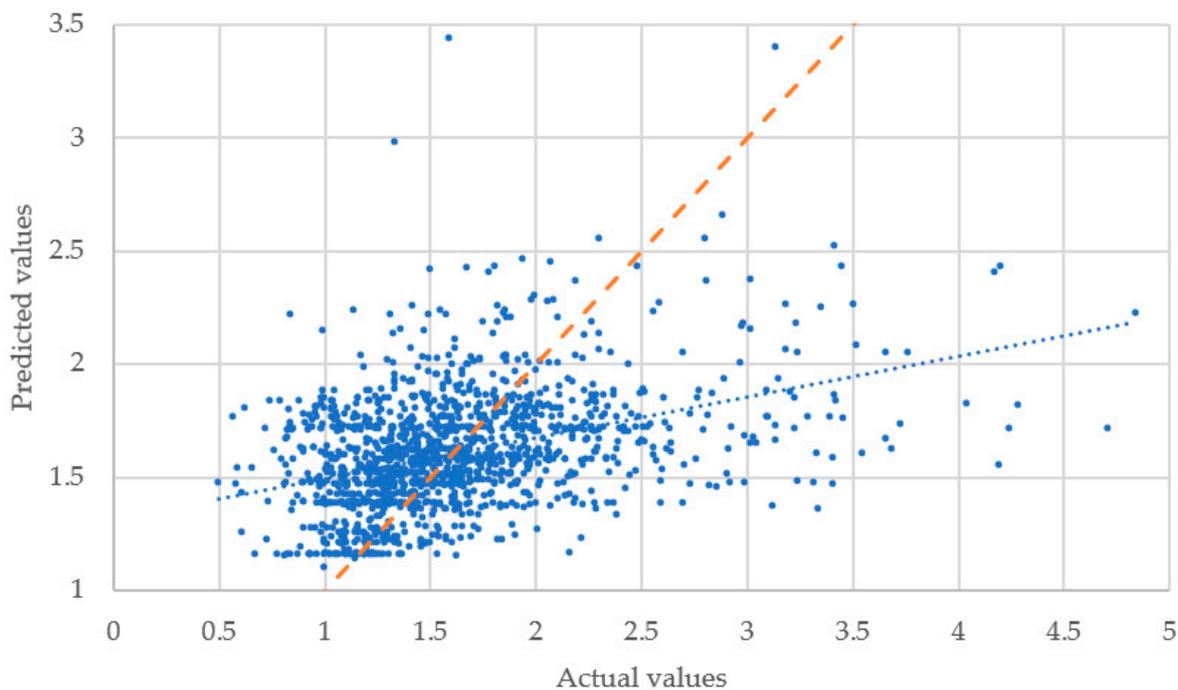


Figure 7. Actual vs. predicted values–M5P without experience-related variables.

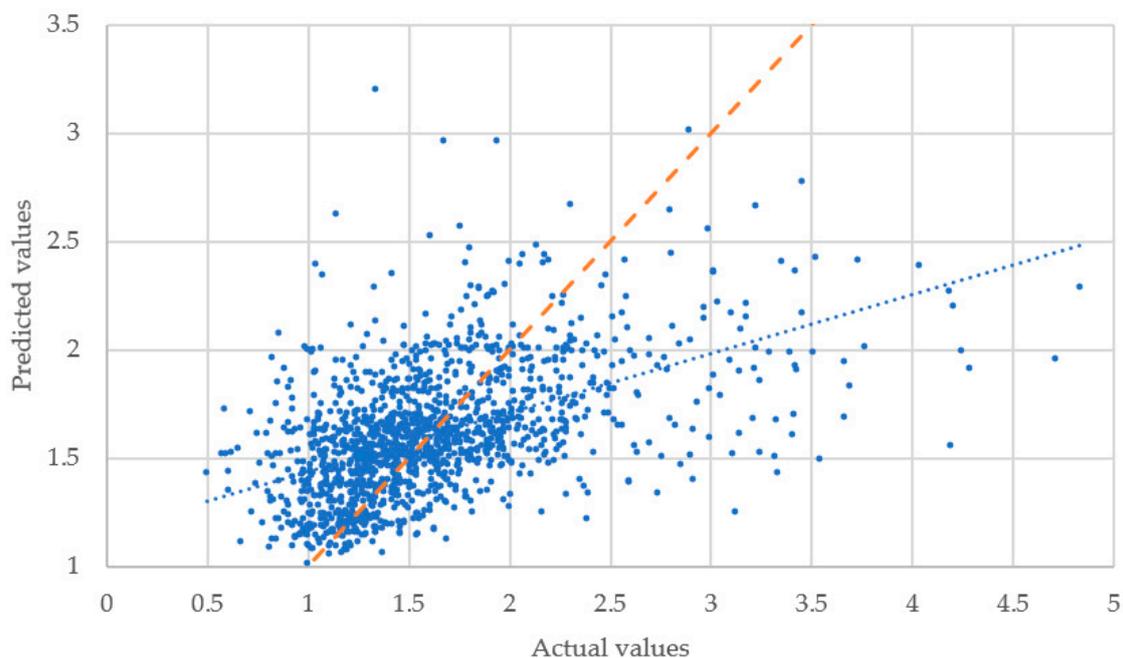


Figure 8. Actual vs. predicted values–M5P with experience-related variables.

Table 7. Evaluation of the performance of the models on instances of more than two hours.

% of Accurate Predictions for Instances over Two Hours				
Experiment	Input Variables Set	Error below 10' (%)	Error below 15' (%)	Error below 30' (%)
Linear Regression	w/o Experience	1.19	5.16	30.56
Linear Regression	w Experience	1.98	5.56	32.14
M5P	w/o Experience	3.97	10.71	34.92
M5P	w Experience	9.52	14.29	41.27

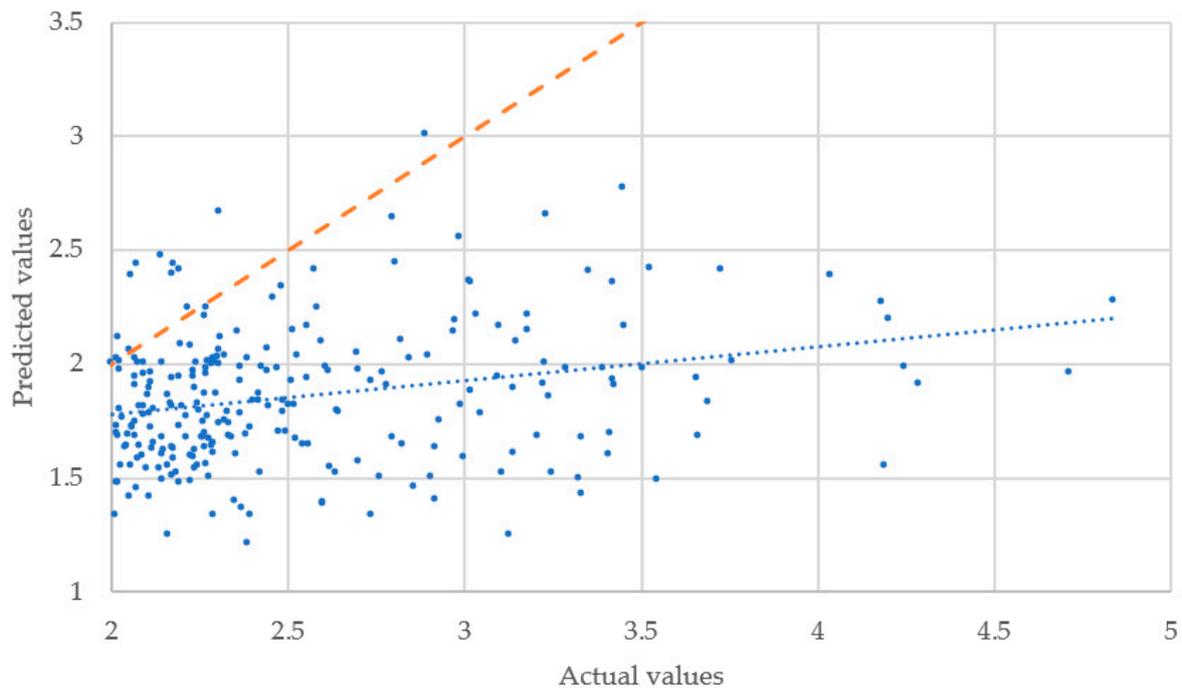


Figure 9. Actual vs. predicted values—M5P with experience-related variables for actual values over 2 h.

After examining the performance of the models, the results of multivariate Linear Regression analysis with experience-related variables are interpreted, focusing on new-found correlations between input variables and lead time. Table 8 shows the coefficients determined for each variable in this experiment, as well as their standard error and significance. WEKA does not provide the p -values, but it performs a two-tailed Student's t -test. The t -statistic values can be then converted into the significance p -values.

It must be noted that there are both numeric and nominal input variables considered in the Linear Regression model. Coefficients for numeric variables represent the estimated growth in lead times when said numeric variable increases its value in one unit; however, when dealing with nominal variables, there is a coefficient for every level of the variable. For example, regarding the shift variable, which contains three levels (corresponding to the morning, afternoon, and night shifts), two resulting coefficients can be expected. If the night shift is taken as a reference, the coefficient for the morning shift level represents the expected variation in lead time for a morning-shift operation compared to when the instance corresponds to the night shift. Similarly, the model should produce a coefficient for the afternoon shift; however, the algorithm autonomously selects a reference level, and if the difference between the reference and a certain level is not significant, it does not produce its coefficient.

The results shown in Table 8 provide interesting insight into the bending operation studied. Firstly, it must be noted that the variables included in the table are the ones chosen in the M5-based feature selection filter ran before executing the Linear Regression algorithm. The rest of the variables have not been found to provide additional information for the lead time prediction (save for the personnel variable, which will be discussed next). In particular, such input variables are the width, steel yield strength, steel normalization and worker age variables.

Table 8. Multivariate Linear Regression model coefficients and statistics.

Variable	Type	Level	Coefficient	SE	T-Statistic	p-Value
Intercept	-	-	0.985610	0.074708	13.1929	>0.0001
Shift	Nominal	M	0.026372	0.014198	1.8574	0.0633
Machine	Nominal	A	0.146304	0.017338	8.4382	>0.0001
	Nominal	K2, J2, NL	-0.054601	0.167111	-0.3267	0.7438
Toughness	Nominal	J2, NL	0.078347	0.169700	0.4617	0.6443
	Nominal	NL	0.615679	0.236723	2.6008	0.0093
Length	Numeric	-	0.000018	0.000004	4.7167	>0.0001
Ferrule pos.	Numeric	-	0.012061	0.002465	4.8923	>0.0001
Section pos.	Numeric	-	0.033489	0.005823	5.7517	>0.0001
Exp. (sector)	Numeric	-	-0.036360	0.012205	-2.9790	0.0029
30d freq.	Numeric	-	-0.046213	0.013841	-3.3389	0.0009

Secondly, it can be seen that most of the variable coefficients are significant at a 0.01 significance level; however, there are two exceptions: the shift variable coefficient, adding 0.0264 expected hours (less than two minutes) when the operation is performed in the morning shift, is only significant at a 0.1 level. Additionally, the toughness variable only has a level that is predicted to add a significant deviation, the NL toughness subgrade.

The intercept is a constant value that represents the estimated time when all the nominal variables are at their reference levels and the numeric variables are 0; this value is not of interest for this interpretation, since there are no plates with null length, width, or thickness, for example.

Regarding the nominal variables, the most noteworthy effects are those of the machine variable and, particularly, of the steel plate toughness variable. The operations performed in the bending station A are expected to take nearly 9 min more than those performed on station B. Furthermore, the steel plates with a toughness subgrade NL (the second toughest of the plates encountered in the dataset) are estimated to take 37 min longer than those with the lower subgrades JR and J0; these are significant time increases, especially when the mean bending time of all the operations in the dataset is of 1.62 h (97 min).

Regarding the numeric variables, the coefficients show slight increases as both the position of the section in the tower and the position of the ferrule in the section rise; this increase amounts to 11 min when comparing the lowest position of a ferrule (1) to the highest (16), and to 10 min when comparing ferrules from the bottom section to ferrules in the highest top section produced in the analyzed timespan (6).

As opposed to what was a priori expected, the length of the plate does not have a remarkable effect on the lead time of the process. The expected lead time increase per meter of length is only 1.08 min. The difference between processing the longest (24.953 m) and shortest (7.324 m) plates in the records is expected to be 19 min. It must be kept in mind that the steel plates are inserted through the bending machine rolls lengthwise and, therefore, it could be anticipated that a longer plate would take significantly more time to be bent, but the results suggest otherwise.

Regarding the experience variables, it can be seen that the age of the workers has been discarded from the model by the feature selection filter; however, the experience at the sector and the frequency at the station in the last 30 days pose significant effects on the bending lead time. For example, a worker with 4 and a half years of experience in the sector (the maximum value observed in the dataset) is expected to employ 9.8 min less to perform a bending operation than one with no experience, *ceteris paribus*. Furthermore, a worker that has performed 3.13 daily bending operations on average during the previous 30 days is predicted to take 8.7 min less to carry out a bending task than an employee with no operations performed in the previous 30 days, *ceteris paribus*.

It can be observed that the personnel variable has not been included in Table 8, for the sake of conciseness. Conversely, the aggregate coefficients of each of the levels of the personnel variable are shown in Table 9. There are 18 levels for the variable, representing each employee that has performed the bending operation in the recorded timespan: two

of those workers are taken as the reference level (Q and R). Another two (A and B) are expected to perform the bending operations in nearly five fewer minutes than workers Q and R. The remaining 14 employees are estimated to produce an increase in the operation time ranging from 8 min up to 69 min over the expected bending time for employees Q and R. In fact, the highest difference found in expected lead time (O and P vs. A and B) amounts to 74 min, 76.19% of the average bending lead time, a testament to the relevance of personnel for the prediction of the bending lead time.

Table 9. Aggregate Linear Regression coefficients for each level of the personnel variable.

Employee	Aggregate Coefficient
A, B	−0.080487
C	0.175264
D	0.134637
E, F	0.19485
G	0.252704
H, I	0.319486
J	0.431811
K, L	0.338771
M, N	0.493093
O, P	1.153838
Q, R	Reference

6. Discussion

The results presented in Section 5 show that the bending operation produces a significant time variability, which is somewhat expected due to it being a non-automated process, the outcome and duration of which are undoubtedly influenced by the employees' actions.

However, certain factors' influences, or lack thereof, must be discussed using the results of the Linear Regression approach. Surprisingly, the dimensional variables do not have a particularly strong effect on the lead time of the bending process. Even if width can understandably, given the configuration of the process, be a non-factor, it was a priori expected that the length of the plate would have a significantly higher effect on the duration of the process, but the results prove otherwise.

While the shift in which the process was completed, the machine used, and the position of the section, and the ferrule do have a significant yet slight effect on the lead time, the most noteworthy factors are the steel toughness and the personnel. The analysis shows a near 40-minute expected increase in the bending operation when one of the toughest steels, with subgrade NL, is processed.

The variability introduced by personnel is even greater, with 2 of the 18 workers being expected to increase the lead time of the bending operation, which takes an average of 97 min, by 74 min; this suggests the need for more in-depth analysis and for the standardization of the bending process in the plant.

Overall, the machine learning algorithms show moderate predictive power, which can be considered useful in an industrial setting such as the one in which this work is based. The proposed models will predict the lead time with less than a 10-min error in 31.9–36.9% of the occasions, which can serve as a somewhat solid base for the planning of the bending process and for the detection of production anomalies. Additionally, while the use of the multivariate Linear Regression algorithm proves useful to analyze the effects of each input variable, the performance of the M5P method is clearly superior to that of the former. The fact that both analyses require relatively low and similar runtimes justifies the use of the superior approach: the M5P method.

The performance of the M5P model proposed is especially interesting when dealing with instances containing extreme lead time values, over 2 h, which are uncommon but have a very negative effect on the manufacturing process planning and control if not detected preemptively. The model is able to predict 41.3% of these instances with an error of less than 30 min, and 14.3% of them with under 15 min of error.

The effect of the inclusion of the experience-related variables is of particular interest. As expected, as the experience of the workers in the plant or sector increases, so does their time-wise efficiency while performing the bending operation. Similarly, if a worker has had short-term experience with the bending process, he is expected to carry it out in shorter lead times. While the Linear Regression coefficients of these variables do not denote a higher effect of these input variables over the previously described ones, the improvement of the predictive power of the models with the inclusion of the workers' ages, experience at the plant/sector and 30-day frequency is noteworthy, particularly in the case of the M5P method.

7. Conclusions

In this work, the lead time of the bending process in a wind tower manufacturing plant is analyzed. Two machine learning models have been applied to the dataset corresponding to a real-world case study. One of them, the M5P model, maximizes the accuracy of the lead time predictions, based on a preliminary cross-validation-based model selection process. On the other hand, the Multivariate Linear Regression model serves as a basis to analyze the effect of each input variable in the prediction.

While the prediction results produced by the models could not be considered sufficient in other applications, they must be put into context: the lead time variability, the quality of the data, and the non-automated nature of the process hinder the predictive power of the proposed machine learning models. Nevertheless, these predictions are useful for the wind turbine tower manufacturing industry, characterized by its high competitiveness. Aside from the potential applications of the improved predictions already discussed in the article, anomaly detection and production planning and control, there are plenty of other uses of the approach presented in the paper: wind farm projects are usually awarded on a competitive basis among a small number of manufacturing companies. As a result, the selling price of the tower must be accurately specified: a high bid would exclude the company from the tender, and a low bid would jeopardize the economic viability of the production. There are two fundamental aspects of cost estimation that can be improved through the lead time prediction: a better estimation of personnel costs (which account for approximately 30% of the total production costs in this industry) and improved forecasting of the supplies and consumables used during production, as well as of the storage costs.

Moreover, as discussed throughout the text, this approach does not only provide value from a predictive point of view, but also from the perspective of a continual improvement process. The findings of this work allow identifying areas of improvement that would in turn increase the reliability of future predictions, such as employee training and standardization; it also provides, given the results of the inclusion of the experience-related variables, additional implications for the hiring, training, and scheduling fronts. First of all, the results show that employees that have already had experience at the plant or the sector are expected to be more efficient in the bending operation from a lead-time standpoint. Training and practice in the process may also increase the performance of the workers, as they accumulate experience at the bending station. Regarding manpower scheduling, the findings highlight the necessity to consider the short-term experience of the workers at the station as a means of improving their performance.

This study presents some limitations. Firstly, the size of the dataset is relatively small for a machine learning analysis. The addition of further instances could help strengthen the claims made throughout this article; it could also favor the use of more advanced techniques such as the Neural Networks, which have been proven to produce remarkable results when the dataset is large enough. An exhaustive hyperparameter tuning process has been conducted in order to optimize the performance of the Multilayer Perceptron Neural Network; however, it is believed that the limited size of the dataset has favored the performance of the M5P model over the former. Additionally, having wider ranges in the experience-related input variables could have helped to further analyze the effect of age and experience on the performance of workers. The minimum and maximum workers' ages in

the current dataset are of 28 and 57 years, respectively. Having data on workers at even later stages could have been of interest and help obtain significant effects of age, particularly as governments in western countries have increased the retirement age in recent years (and are planning to extend it even further) given the World's population aging.

Author Contributions: Conceptualization, A.L.-E., A.E.-S. and M.-L.M.-D.; Data curation, A.L.-E.; Formal analysis, A.L.-E., A.E.-S. and M.-L.M.-D.; Funding acquisition, A.E.-S.; Investigation, A.L.-E., A.E.-S. and M.-L.M.-D.; Methodology, A.L.-E., A.E.-S. and M.-L.M.-D.; Project administration, A.E.-S.; Resources, A.E.-S.; Software, A.L.-E. and A.R.-V.; Supervision, A.E.-S.; Validation, A.L.-E. and A.E.-S.; Visualization, A.L.-E.; Writing—original draft, A.L.-E. and A.E.-S.; Writing—review and editing, A.L.-E. and A.E.-S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was co-funded by the European Regional Development Fund ERDF by means of the Interreg V-A Spain-Portugal Programme (POCTEP) 2014–2020, through the CIU3A project (reference 0754_CIU3A_5_A), and by the Agency for Innovation and Development of Andalusia (IDEA), by means of “Open, Singular and Strategic Innovation Leadership” Programme, through the joint innovation unit project OFFSHOREWIND (reference 802C2000003). The research was also supported by the Ministry of Universities of Spain through a grant for the Training of University Researchers (Ayuda para la Formación del Profesorado Universitario, reference FPU20/05584).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. bp Statistical Review of World Energy 2020. Available online: <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html> (accessed on 15 January 2021).
2. Instituto Nacional de Estadística (INE). Índice de Precios de Consumo—Índices Nacionales de Subgrupos. Available online: <https://www.ine.es/up/yRDuqUS7iB> (accessed on 10 October 2021).
3. Stetco, A.; Dinmohammadi, F.; Zhao, X.; Robu, V.; Flynn, D.; Barnes, M.; Keane, J.; Nenadic, G. Machine learning methods for wind turbine condition monitoring: A review. *Renew. Energy* **2019**, *133*, 620–635. [\[CrossRef\]](#)
4. Pérez-Cubero, E.; Poler, R. Aplicación de algoritmos de aprendizaje automático a la programación de órdenes de producción en talleres de trabajo: Una revisión de la literatura reciente. *Dir. Organ.* **2020**, *72*, 82–94. [\[CrossRef\]](#)
5. Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Dev.* **1959**, *3*, 210–229. [\[CrossRef\]](#)
6. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [\[CrossRef\]](#)
7. Kang, Z.; Catal, C.; Tekinerdogan, B. Machine learning applications in production lines: A systematic literature review. *Comput. Ind. Eng.* **2020**, *149*, 106773. [\[CrossRef\]](#)
8. Backus, P.; Janakiram, M.; Mowzoon, S.; Runger, G.C.; Bhargava, A. Factory Cycle-Time Prediction with a Data-Mining Approach. *IEEE Trans. Semicond. Manuf.* **2006**, *19*, 252–258. [\[CrossRef\]](#)
9. Öztürk, A.; Kayaligil, S.; Özdemirel, N.E. Manufacturing lead time estimation using data mining. *Eur. J. Oper. Res.* **2006**, *173*, 683–700. [\[CrossRef\]](#)
10. Alenezi, A.; Moses, S.A.; Trafalis, T.B. Real-time prediction of order flowtimes using support vector regression. *Comput. Oper. Res.* **2008**, *35*, 3489–3503. [\[CrossRef\]](#)
11. Wang, C.; Jiang, P. Deep neural networks based order completion time prediction by using real-time job shop RFID data. *J. Intell. Manuf.* **2019**, *30*, 1303–1318. [\[CrossRef\]](#)
12. Gyulai, D.; Pfeiffer, A.; Bergmann, J.; Gallina, V. Online lead time prediction supporting situation-aware production control. *Procedia CIRP* **2018**, *78*, 190–195. [\[CrossRef\]](#)
13. Gyulai, D.; Pfeiffer, A.; Nick, G.; Gallina, V.; Sihni, W.; Monostori, L. Lead time prediction in a flow-shop environment with analytical and machine learning approaches. *IFAC-PapersOnLine* **2018**, *51*, 1029–1034. [\[CrossRef\]](#)
14. Pfeiffer, A.; Gyulai, D.; Kádár, B.; Monostori, L. Manufacturing Lead Time Estimation with the Combination of Simulation and Statistical Learning Methods. *Procedia CIRP* **2016**, *41*, 75–80. [\[CrossRef\]](#)
15. Lingitz, L.; Gallina, V.; Ansari, F.; Gyulai, D.; Pfeiffer, A.; Sihni, W. Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. *Procedia CIRP* **2018**, *72*, 1051–1056. [\[CrossRef\]](#)
16. Treiber, N.A.; Heinemann, J.; Kramer, O. Wind power prediction with machine learning. In *Computational Sustainability. Studies in Computational Intelligence*; Lässig, J., Kersting, K., Morik, K., Eds.; Springer: Cham, Switzerland, 2016; Volume 645, pp. 13–29, ISBN 978-331-981-138-3.

17. Demolli, H.; Dokuz, A.S.; Ecemis, A.; Gokcek, M. Wind power forecasting based on daily wind speed data using machine learning algorithms. *Energy Convers. Manag.* **2019**, *198*, 111823. [[CrossRef](#)]
18. An, G.; Jiang, Z.; Chen, L.; Cao, X.; Li, Z.; Zhao, Y.; Sun, H. Ultra Short-Term Wind Power Forecasting Based on Sparrow Search Algorithm Optimization Deep Extreme Learning Machine. *Sustainability* **2021**, *13*, 10453. [[CrossRef](#)]
19. Fogno Fotso, H.R.; Aloyem Kazé, C.V.; Djuidje Kenmoé, G. A novel hybrid model based on weather variables relationships improving applied for wind speed forecasting. *Int. J. Energy Environ. Eng.* **2022**, *13*, 43–56. [[CrossRef](#)]
20. Wang, J.; Li, H.; Wang, Y.; Lu, H. A hesitant fuzzy wind speed forecasting system with novel defuzzification method and multi-objective optimization algorithm. *Expert Syst. Appl.* **2021**, *168*, 114364. [[CrossRef](#)]
21. Neshat, M.; Majidi Nezhad, M.; Abbasnejad, E.; Mirjalili, S.; Bertling Tjernberg, L.; Astiaso Garcia, D.; Alexander, B.; Wagner, M. A deep learning-based evolutionary model for short-term wind speed forecasting: A case study of the Lillgrund offshore wind farm. *Energy Convers. Manag.* **2021**, *236*, 114002. [[CrossRef](#)]
22. Morshed-Bozorgdel, A.; Kadhodazadeh, M.; Valikhan Anaraki, M.; Farzin, S. A Novel Framework Based on the Stacking Ensemble Machine Learning (SEML) Method: Application in Wind Speed Modeling. *Atmosphere* **2022**, *13*, 758. [[CrossRef](#)]
23. Chaudhary, A.; Sharma, A.; Kumar, A.; Dikshit, K.; Kumar, N. Short term wind power forecasting using machine learning techniques. *J. Stat. Manag. Syst.* **2020**, *23*, 145–156. [[CrossRef](#)]
24. Kim, G.; Hur, J. A Short-Term Power Output Forecasting Based on Augmented Naïve Bayes Classifiers for High Wind Power Penetrations. *Sustainability* **2021**, *13*, 12723. [[CrossRef](#)]
25. Liu, Z.; Wang, Y.; Hua, X.; Zhu, H.; Zhu, Z. Optimization of wind turbine TMD under real wind distribution countering wake effects using GPU acceleration and machine learning technologies. *J. Wind Eng. Ind. Aerodyn.* **2021**, *208*, 104436. [[CrossRef](#)]
26. Richmond, M.; Sobey, A.; Pandit, R.; Kolios, A. Stochastic assessment of aerodynamics within offshore wind farms based on machine-learning. *Renew. Energy* **2020**, *161*, 650–661. [[CrossRef](#)]
27. Petrov, A.N.; Wessling, J.M. Utilization of machine-learning algorithms for wind turbine site suitability modeling in Iowa, USA. *Wind Energy* **2015**, *18*, 713–727. [[CrossRef](#)]
28. Fischetti, M.; Fraccaro, M. Machine learning meets mathematical optimization to predict the optimal production of offshore wind parks. *Comput. Oper. Res.* **2019**, *106*, 289–297. [[CrossRef](#)]
29. Yang, Z.X.; Wang, X.B.; Zhong, J.H. Representational learning for fault diagnosis of wind turbine equipment: A multi-layered extreme learning machines approach. *Energies* **2016**, *9*, 379. [[CrossRef](#)]
30. Gao, Q.; Wu, X.; Guo, J.; Zhou, H.; Ruan, W. Machine-Learning-Based Intelligent Mechanical Fault Detection and Diagnosis of Wind Turbines. *Math. Probl. Eng.* **2021**, *2021*, 9915084. [[CrossRef](#)]
31. Helbing, G.; Ritter, M. Deep Learning for fault detection in wind turbines. *Renew. Sustain. Energy Rev.* **2018**, *98*, 189–198. [[CrossRef](#)]
32. Yeh, C.H.; Lin, M.H.; Lin, C.H.; Yu, C.E.; Chen, M.J. Machine learning for long cycle maintenance prediction of wind turbine. *Sensors* **2019**, *19*, 1671. [[CrossRef](#)]
33. Elasha, F.; Shanbr, S.; Li, X.; Mba, D. Prognosis of a wind turbine gearbox bearing using supervised machine learning. *Sensors* **2019**, *19*, 3092. [[CrossRef](#)]
34. Carroll, J.; Koukoura, S.; McDonald, A.; Charalambous, A.; Weiss, S.; McArthur, S. Wind turbine gearbox failure and remaining useful life prediction using machine learning techniques. *Wind Energy* **2019**, *22*, 360–375. [[CrossRef](#)]
35. Sainz, J.A. New Wind Turbine Manufacturing Techniques. *Procedia Eng.* **2015**, *132*, 880–886. [[CrossRef](#)]
36. Skirbekk, V. Age and productivity potential: A new approach based on ability levels and industry-wide task demand. *Popul. Dev. Rev.* **2008**, *34*, 191–207.
37. Rembiasz, M. Impact of employee age on the safe performance of production tasks. *MATEC Web Conf.* **2017**, *94*, 07009. [[CrossRef](#)]
38. Ilmarinen, J. The Work Ability Index (WAI). *Occup. Med.* **2007**, *57*, 160. [[CrossRef](#)]
39. Theppitak, C.; Higuchi, Y.; Kumudini, D.V.G.; Lai, V.; Movahed, M.; Izumi, H.; Kumashiro, M. Aging and work ability: Their effect on task performance of industrial workers. In *Ergonomics in Asia: Development, Opportunities and Challenges, Proceedings of the 2nd East Asian Ergonomics Federation Symposium (EAEFS 2011), Hsinchu, Taiwan, 4–8 October 2011*; Shih, Y.-C., Liang, S.-F.M., Huang, Y.-H., Lin, Y.-C., Lin, C.-L., Eds.; CRC Press: Boca Raton, FL, USA, 2012.
40. Kumudini, D.V.G.; Higuchi, Y.; Theppitak, C.; Lai, V.; Movahed, M.; Izumi, H.; Kumashiro, M. Effects of mental capacity on work ability in middle-aged factory workers: A field study. In *Ergonomics in Asia: Development, Opportunities and Challenges, Proceedings of the 2nd East Asian Ergonomics Federation Symposium (EAEFS 2011), Hsinchu, Taiwan, 4–8 October 2011*; Shih, Y.-C., Liang, S.-F.M., Huang, Y.-H., Lin, Y.-C., Lin, C.-L., Eds.; CRC Press: Boca Raton, FL, USA, 2012.
41. Toh, G.; Park, J. Review of vibration-based structural health monitoring using deep learning. *Appl. Sci.* **2020**, *10*, 1680. [[CrossRef](#)]
42. UNE-EN 10025-2:2020; Hot Rolled Products of Structural Steels—Part 2: Technical Delivery Conditions for Non-Alloy Structural Steels. UNE: Madrid, Spain, 2020. Available online: <https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0064323> (accessed on 15 January 2021).
43. Frank, E.; Hall, M.A.; Witten, I.H. The WEKA Workbench (Online Appendix). In *Data Mining: Practical Machine Learning Tools and Techniques*; Morgan Kaufmann: Burlington, MA, USA, 2016; ISBN 978-012-804-291-5.
44. Mansour, A.M.; Almutairi, A.; Alyami, S.; Obeidat, M.A.; Almkahles, D.; Sathik, J. A Unique Unified Wind Speed Approach to Decision-Making for Dispersed Locations. *Sustainability* **2021**, *13*, 9340. [[CrossRef](#)]

45. Joshuva, A.; Vishnuvardhan, R.; Deenadayalan, G.; Sathishkumar, R.; Sivakumar, S. Implementation of Rule based Classifiers for Wind Turbine Blade Fault Diagnosis Using Vibration Signals. *Int. J. Recent Technol. Eng.* **2019**, *8*, 320–331. [[CrossRef](#)]
46. Flach, P. *Machine Learning: The Art and Science of Algorithms That Make Sense of Data*; Cambridge University Press: Cambridge, UK, 2012; ISBN 978-110-742-222-3.
47. Wang, Y.; Witten, I.H. Inducing Model Trees for Continuous Classes. In Proceedings of the 9th European Conference on Machine Learning, Prague, Czech Republic, 23–25 April 1997; pp. 128–137.
48. Quinlan, J.R. Learning with Continuous Classes. In *AI '92, Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, Hobart, Tasmania, 16–18 November 1992*; Adams, A., Sterling, L., Eds.; World Scientific: Singapore, 1992; pp. 343–348.