Journal of Mechanical Engineering and Manufacturing
https://www.sciltp.com/journals/jmem

Article

AI-Powered Data Management to Optimize Data Collection and Processing in a Painting Laboratory

Maria Teresa Ribeiro Pereira [1,2]*, Marisa João Guerra Pereira [1,2], Miguel Guedes Tavares [2], André Martins Guimarães [3,4] and Hermilio Vilarinho [5]

[1] Associate Laboratory for Energy, Transports and Aerospace (LAETA-INEGI, ISEP), Instituto Superior de Engenharia do Porto, Polytechnic of Porto, Dr. António Bernardino de Almeida, 431, 4249-015 Porto, Portugal
[2] Instituto Superior de Engenharia do Porto (ISEP), Polytechnic of Porto, Dr. António Bernardino de Almeida, 431, 4249-015 Porto, Portugal
[3] CISE—Electromechatronic Systems Research Centre, Department of Electromechanical Engineering, University of Beira Interior, 6201-001 Covilhã, Portugal
[4] CISeD—Research Centre for Digital Services, Polytechnic of Viseu, Campus Politécnico Santa Maria, Av. Cor. José Maria Vale de Andrade, 3504-510 Viseu, Portugal
[5] Associate Laboratory for Energy, Transports and Aerospace (LAETA-INEGI), Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
* Correspondence: mtp@isep.ipp.pt

How To Cite: Pereira, M.T.R.; Pereira, M.J.G.; Tavares, M.G.; et al. AI-Powered Data Management to Optimize Data Collection and Processing in a Painting Laboratory. *Journal of Mechanical Engineering and Manufacturing* **2026**. https://doi.org/10.53941/jmem.2026.100011

Received: 4 November 2025
Revised: 20 January 2026
Accepted: 23 January 2026
Published: 10 February 2026

Abstract: Industrial laboratories often remain under-digitized compared to production lines, creating a gap between data acquisition and analytical intelligence, critical for advanced quality control. This study addresses this gap by proposing and validating a novel framework that combines Low-Code digitalisation tools with Machine Learning (ML) and Causal Inference to optimise data collection and analysis in an automotive painting laboratory. A Microsoft Power Apps-based platform was developed in order to digitalise all measurement records, eliminating manual transcription errors (previously ≈ 40.01%) and reducing data-handling time by up to 34% of an operator's shift, while enabling centralised, traceable storage and Power BI integration. Four datasets were used to assess predictive capacity with Random Forest, XGBoost and Neural Networks; Random Forest consistently provided the most stable results—Mean Absolute Error (MAE) of 0.972, Mean Absolute Percentage Error (MAPE) of 16.45%, and Root Mean Square Error (RMSE) of 1.307. Causal models (Linear Regression, DoWhy, Causal Forest, Double Machine Learning) consistently identified ultrafiltrate I solid content of the electrodeposition process as a dominant causal factor for defects. This study provides a novel framework that bridges digitalisation and ML-based causal reasoning in laboratory settings, offering a scalable approach that can be extended and replicated in other industrial sectors, aiming to develop smart, data-driven quality control systems.

Keywords: digitalisation; laboratory; electrodeposition; machine learning; causal inference

1. Introduction

The automotive industry faces increasing pressure to improve quality, reduce waste, and enhance process efficiency, particularly under the paradigm of Industry 4.0 [1]. Within this framework, digitalisation plays a pivotal role, transforming manual data collection and disconnected systems into integrated, traceable, and data-driven environments capable of supporting advanced analytics and decision-making [2–4]. While significant progress has been achieved in production lines—through automation, sensorisation and real-time monitoring—supporting laboratory operations have often lagged behind this transformation, remaining partially manual and weakly integrated into the broader digital ecosystem.

In many industrial laboratories, measurement data are still collected manually and subsequently transcribed into spreadsheets stored across separate files and systems. Evidence from the automotive factory used to validate this study shows that such practices typically consume between 40 min and 2 h of an 8-h shift, while introducing transcription errors in approximately 40.01% of historical records. Beyond their operational inefficiency, these practices compromise data integrity, limit traceability and significantly reduce the analytical value of laboratory data. As a result, laboratories—despite generating critical process information—are frequently unable to support proactive quality control or data-driven root cause analysis. This mismatch represents a clear scientific and industrial gap: the lack of integrated frameworks that connect laboratory digitalisation with advanced analytical intelligence.

From a scientific perspective, the literature shows extensive adoption of machine learning (ML) techniques for quality control and defect detection in automotive production lines. However, most studies focus on large-scale, sensor-rich production environments, with limited attention given to laboratory-scale processes where data are scarcer, heterogeneous and often manually collected. Moreover, existing works tend to emphasise predictive accuracy, while the causal understanding of defect formation—essential for process optimisation—remains comparatively underexplored in laboratory contexts. Consequently, the potential of combining digitalisation, predictive ML and causal inference within industrial laboratories remains insufficiently addressed.

This study aims to address this gap by proposing a novel framework integrating digitalisation mechanisms with exploratory ML and causal modelling to improve data collection, data quality and analytical insight in an automotive painting laboratory, with particular emphasis on the electrodeposition (ED) process. The scientific contribution of this work is twofold. First, it empirically demonstrates how low-code digitalisation tools—specifically Microsoft Lists, Power Apps, and Power Automate—can substantially enhance data integrity, traceability, and analytical reliability in real industrial laboratory environments. Second, it evaluates the predictive and causal capabilities of ML algorithms applied to historical ED process data, assessing not only their forecasting performance but also their ability to identify key variables driving defect formation.

Therefore, the main research question guiding this work is: How can data collection and analysis efficiency in an automotive painting laboratory be improved through digitalisation mechanisms?

To answer this question, the general objective of the study is to develop and validate a framework for optimising data collection and analysis in industrial laboratories. The specific objectives include: (i) conducting a comprehensive literature review on digitalisation, ML and causal inference in industrial contexts; (ii) mapping and critically analysing the existing laboratory data collection process; (iii) designing and implementing a digital data acquisition and management tool; (iv) performing exploratory predictive and causal ML analyses to assess defect prediction capability and identify influential process parameters; and (v) validating the proposed solution through quantitative performance metrics and user feedback.

By combining Design Science Research Methodology (DSRM) with the CRISP-DM framework, this work bridges engineering practice and data science, demonstrating that even incremental digitalisation of laboratory workflows can produce measurable gains in productivity, data quality, and analytical insight. Furthermore, while validated in an automotive context, the proposed framework is inherently scalable and transferable to other industrial sectors—such as chemical, pharmaceutical, and surface treatment industries—where laboratory data play a critical role in quality control and process optimisation.

2. Literature Review

The literature reviewed in this study focuses on digitalisation in industrial contexts, machine learning (ML) for prediction and causal inference, and a special emphasis on the electrodeposition (ED) coating process. Digitalisation of processes refers to the conversion of analogue processes and data into digital formats, enabling automation, improved data accessibility, and faster decision making [5]. Within manufacturing, digitalisation is considered a core enabler of Industry 4.0, supporting real-time analytics, IoT integration, and data-driven decision-making [6,7]. Low-code development platforms such as Microsoft Power Apps and Power Automate simplify the creation of tailored applications without the need of advanced programming knowledge, reducing development time and cost [7,8]. These tools lower implementation barriers while maintaining scalability and compliance with corporate IT systems. Similarly, Enterprise Resource Planning (ERP) systems like SAP S/ 4HANA and Microsoft Dynamics 365 centralise organizational data and improve traceability, though their adoption may involve high costs and possible resistance to change on the part of some employees [9]. The COVID-19 pandemic also accelerated digital adoption, highlighting how organisations with robust digital infrastructures were more resilient and agile in responding to operational challenges [4]. Empirical studies demonstrate the economic impact of digitalisation: Pop [2] and Mehta and Rastogi [3] reported reductions in operational costs of up to 30% and significant gains in traceability and quality control accuracy. However, successful digital transformation requires cultural and structural adaptation, as employees must transition from manual to analytical workflows [5,9].

While digitalisation provides the necessary data foundation, the literature also highlights limitations associated with classical forecasting and analytical models. Although traditionally robust and highly interpretable, these models often rely on restrictive assumptions such as linearity, normal error distribution and structural stability, which are frequently violated in real-world industrial environments. Sensitivity to outliers, influential observations, and structural breaks further compromises forecast reliability, particularly in settings characterised by data heterogeneity and operational volatility [10–14]. To address these limitations, scientific research has increasingly focused on applying artificial intelligence techniques. According to Lolla et al. [15], these methods have significantly advanced predictive analytics. Recent computational developments, particularly the use of graphical processing units (GPUs), have further stimulated interest in the applicability of machine learning and deep learning approaches [16,17]

Building upon this digital foundation, the field of Machine Learning (ML) has gained increasing importance in industrial analytics. ML, a subfield of Artificial Intelligence (AI), allows systems to identify patterns in data and improve performance without explicit programming [18]. In industrial contexts, ML has become essential for predictive maintenance, quality control, and process optimisation, enabling organisations to move from reactive to proactive decision-making. ML techniques are generally divided into three main categories:

(1) Supervised Learning—models are trained on labelled data to perform prediction tasks such as regression or classification. Algorithms such as Random Forest (RF) and Extreme Gradient Boosting (XGBoost) are especially effective for industrial applications, as they can handle non-linear relationships, multicollinearity, and noisy datasets [19]. Both have been widely used for defect prediction in the automotive industry due to their robustness and interpretability. RF, for instance, builds multiple decision trees and aggregates their results, reducing variance and overfitting, while XGBoost optimises gradient boosting through regularisation and parallel computation.

(2) Unsupervised Learning—focuses on identifying structures or clusters within unlabelled data. Techniques such as k-means clustering, hierarchical clustering, and Principal Component Analysis (PCA) are often used for dimensionality reduction and exploratory analysis, allowing engineers to identify underlying process behaviours, anomalies, or operational modes [20].

(3) Reinforcement Learning—employs trial-and-error interactions between an agent and an environment to optimise decision-making. It is widely used in robotics, autonomous systems, and process control, where systems can dynamically learn optimal policies from feedback [21].

In addition to these paradigms, Deep Learning (DL) represents a more complex subset of ML, characterised by artificial neural networks (ANNs) with multiple hidden layers. DL models, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have achieved outstanding results in image analysis, signal processing, and time-series forecasting, outperforming traditional algorithms when large volumes of data are available [22,23]. In the automotive manufacturing context, DL models have been used for visual inspection of paint defects, weld analysis, and surface finish evaluation, achieving accuracy rates above 80% [22]. Beyond visual inspection, ANN-based approaches have also been applied to manufacturing-related processes, such as the optimisation of ternary metal alloy electrodes for CH_4 gas detection [24]. Moreover, RNN-based architectures such as Long Short-Term Memory (LSTM) networks can learn long-term dependencies by selectively retaining relevant information, a property explored by Ramos et al. [25] in time-series forecasting applications.

Despite these successes, DL approaches remain highly data-dependent and computationally intensive, which restricts their applicability in laboratory-scale or exploratory industrial studies where datasets are often limited, heterogeneous, and predominantly structured. In such contexts, ensemble learning methods frequently offer a more favourable balance between predictive accuracy, interpretability, and computational efficiency [26]. Consequently, tree-based models such as Random Forest and XGBoost are often emphasised in the literature, as they can accommodate nonlinearities and interactions without explicit parametric assumptions, efficiently handle multiple predictors, and reduce manual modelling effort.

This perspective is reinforced by a growing body of empirical research. Peres et al. [19] demonstrated the effectiveness of Random Forest and XGBoost in predicting quality deviations across automotive production stages using structured process data, while Yao and Feng [27] showed that ensemble learning methods provide a more balanced trade-off between accuracy, stability, and explainability than deep learning models in manufacturing datasets dominated by physicochemical parameters. Although applied in different domains, Tang [28] similarly highlighted that emerging technologies and big data tools have expanded opportunities for predictive analytics in complex systems.

Further industrial evidence is provided by Fouad et al. [29], Ma et al. [30], and Hammam et al. [31], who reported superior robustness and forecasting performance of ML models in industrial decision-support and operational management scenarios characterised by noisy data and high variability.

Recent comparative studies further support the suitability of ensemble learning in industrial time-series contexts. Santoro et al. [32], for example, compared classical models, machine learning algorithms, and deep learning approaches in traffic forecasting and showed that XGBoost outperformed both traditional statistical methods and LSTM-based models in terms of Mean Absolute Error (MAE) and Mean Square Error (MSE). These findings suggest that less complex models may adapt more effectively to industrial datasets exhibiting stationarity and local disturbances.

Beyond predictive performance, causal inference techniques have gained increasing attention for their ability to identify true cause–effect relationships rather than mere statistical associations [33]. While traditional predictive models are effective at identifying correlations, they cannot determine whether changes in one variable genuinely produce changes in another. In industrial environments, causal inference addresses this limitation by modelling the underlying causal structure of a system, enabling practitioners to distinguish process variables that actively drive quality outcomes from those that are merely correlated. Recent advances in computational causal analysis, including Causal Forests, DoWhy frameworks, and Double Machine Learning (DML), integrate machine learning with statistical inference to provide interpretable and data-efficient causal estimates [34]. These methods also support counterfactual "what-if" analyses, allowing engineers to evaluate the potential impact of process adjustments prior to implementation—an especially valuable capability in electrodeposition processes, where identifying root causes of coating defects can guide preventive actions and reduce rework costs.

Given the constraints typically associated with industrial laboratory data, the modelling strategy adopted in this work is primarily based on tree-based ensemble methods, namely Random Forest and XGBoost. These algorithms were selected due to their robustness when applied to structured datasets with limited sample sizes and heterogeneous variable types. Nevertheless, a Neural Network model was also developed in order to explicitly assess the feasibility of applying Deep Learning techniques within the scenario under study, which justifies its inclusion in the modelling framework. The proposed methodology integrates these algorithms with Recursive Feature Elimination with Cross-Validation (RFECV) to support feature selection and enhance interpretability. All analytical steps were structured under the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework to ensure methodological rigour, transparency, and reproducibility.

The CRISP-DM is a widely adopted framework for projects related to Machine Learning and Deep Learning studies. It's comprised by six iterative phases—Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment—that guide the project systematically from conception to implementation [35,36]. By having a strong alignment with business objectives, flexibility and support for iterative refinement, the clarity and reproducibility are widely enhanced. However, CRISP-DM can be resource intensive, rely heavily on stakeholders engagement and often lacks guidance on project management, particularly, many studies mention that the Deployment phase is insufficiently addressed in practice. In real-world scenarios, this methodology has proven very effective, for instance, in manufacturing, CRISP-DM has been successfully used to predict assembly cycle time using real production data and in surface-technology projects [36,37]. Such cases demonstrate the applicability and effectiveness across sectors, while highlighting the need to augment it with better deployment strategies and project management practices.

The selection of ensemble learning techniques is supported by their documented ability to achieve a favourable balance between predictive accuracy, interpretability, and computational efficiency when applied to structured industrial datasets [26]. While deep learning approaches have demonstrated superior performance in domains dominated by unstructured data, such as image and text analysis, their effectiveness typically relies on large labelled datasets and high computational resources. In contrast, ensemble methods have been shown to provide more stable and interpretable solutions in manufacturing contexts characterized by limited sample sizes and high-dimensional physicochemical variables.

Empirical evidence from the literature reinforces this perspective. Peres et al. [19] reported that Random Forest and XGBoost delivered robust predictive performance and meaningful feature importance analysis in an automotive quality control setting based on structured process data. Conversely, Luckow et al. [22] highlighted the strengths of convolutional neural networks for image-based defect detection but also emphasized their limited applicability in small-scale or data-scarce environments. Similarly, Yao and Feng [27] demonstrated that ensemble learning methods offer a more balanced trade-off between accuracy, stability, and explainability than deep learning models in manufacturing datasets dominated by process parameters.

The dataset used in this study comprised a limited number of observations and a heterogeneous set of structured variables, including continuous physicochemical parameters, categorical identifiers, and defect counts.

Under these conditions, ensemble learning offers a robust analytical framework for exploring predictive and causal relationships while effectively mitigating the risks of overfitting. The integration of feature engineering and RFECV further enhances model interpretability by identifying the most influential process variables, thereby supporting data-driven root cause analysis and facilitating actionable insights for laboratory and production-oriented environments [38,39]. Ultimately, the use of ML in laboratory data analysis represents not merely a computational improvement but a methodological transformation, enabling laboratories to evolve from isolated measurement routines to integrated analytical systems. Such systems continuously capture, validate, and model data to support real-time decision-making, in alignment with the digital twin and smart manufacturing principles of Industry 4.0 [40].

Complementing predictive modelling, causal inference was applied in the present research as a subsequent analytical step to strengthen interpretability and support decision-making. Multiple causal approaches—Linear Regression, DoWhy, Causal Forest, and Double Machine Learning—were employed to ensure robustness and consistency of results. This hybrid analytical perspective validates causal findings across different modelling assumptions and aligns with recent literature advocating the combined use of predictive accuracy and causal understanding to support data-driven decision-making in industrial systems.

To evaluate the performance of the regression models developed in this study, three standard error-based metrics were used: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE). These metrics quantify the deviation between the predicted values ($\hat{y}_i$) and the observed values (y_i) across all n samples.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{1}$$

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{3}$$

- MAE measures the average magnitude of the errors without considering their direction, providing a clear interpretation in the same unit as the target variable.
- MAPE expresses the error as a percentage, allowing comparison of models across different datasets or scales.
- RMSE penalises larger deviations more strongly, making it sensitive to outliers and useful for assessing overall model stability.

Unlike accuracy, which is commonly used for classification problems where predictions belong to discrete categories, the models developed in this work address a regression problem, predicting continuous numerical values representing defect counts. Therefore, error-based metrics were deemed more appropriate to assess model performance in this context.

For the causal analysis, the coefficient of determination (R^2) was calculated to quantify the proportion of variance in the observed values explained by the model:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{4}$$

Finally, the application of ML in automotive manufacturing demonstrates the transformative potential of data-driven methods across production optimisation, predictive maintenance, and quality control. Yao and Feng [27] illustrated ML's versatility across domains including road marking recognition, sales forecasting, and defect detection, while Peres et al. [19], at Volkswagen AutoEuropa, reported the successful use of Random Forest, XGBoost, and Support Vector Machines to predict dimensional defects across 18,148 vehicles, achieving recall rates of up to 97.9%. Similarly, Luckow et al. [22] explored Deep Learning (DL) for visual inspection in automotive assembly, highlighting the potential of Convolutional Neural Networks (CNNs) to identify surface defects with high precision. However, their work also exposed critical limitations such as the need for extensive labelled datasets, computational intensity, and difficulty in model interpretability—challenges that remain common in industrial AI adoption. Beyond technical performance, these studies emphasise a broader trend: the integration of digitalisation and analytics as essential enablers for competitiveness. As noted by Brennen and Kreiss [5] and Verhoef et al. [6], the convergence of Industry 4.0, data governance, and machine learning is redefining how manufacturing organisations approach process control and decision-making. Despite this significant progress, the adoption of ML in industrial laboratories—especially in electrodeposition quality control—remains limited. Most initiatives focus on production line automation rather than laboratory data digitalisation, leaving a gap that this study addresses. By combining digital transformation with ML-based causal analysis, this research contributes to bridging that gap, extending the lessons from large-scale industrial deployments to the laboratory environment.

3. Research Methodology

This study employed a hybrid methodological approach that integrates digital transformation with exploratory data analysis using Machine Learning (ML) and causal inference techniques. Two complementary methodological frameworks were employed to ensure both scientific rigor and industrial applicability: (i) the Design Science Research Methodology (DSRM), used to guide the development, implementation, and validation of the digital artifact [23,24]; and (ii) the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework, applied to structure the data understanding, preparation, modelling, and evaluation phases [36,37]. The methodology is therefore organised into two tightly connected components: the digitalisation of data collection and processing in the laboratory, and the exploratory predictive–causal analysis of historical electrodeposition (ED) process data.

To facilitate the reader's understanding, Figure 1 presents a visual representation of the developed framework, which is divided according to the two methodologies used, as well as the associated processes. The arrows represent both the workflow and the information flow.

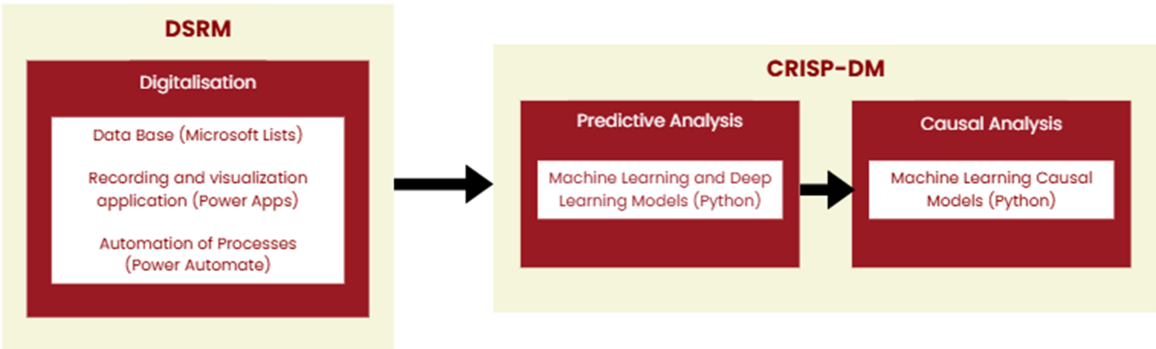

Figure 1. Framework developed.

3.1. Digitalisation of Data Collection and Processing

Prior to this work, measurement data from the laboratory of the automotive factory used to validate this study were manually recorded on paper sheets and subsequently transcribed into monthly Excel spreadsheets. This manual workflow was highly error-prone and time-consuming, introducing transcription inconsistencies and limiting data traceability and accessibility. The digitalisation stage aimed to (i) eliminate manual data transcription; (ii) centralise measurement records; and (iii) enable future integration with automated analytics and IoT-based monitoring systems.

This stage was structured according to the DSRM cycle [41], ensuring systematic artifact development and validation. The problem identification and motivation phase revealed that transcription errors affected approximately 40.01% of historical records, while data entry consumed on average 34% of an operator's working shift. Based on these findings, the definition of objectives focused on designing a digital system capable of reducing human error, improving data integrity, ensuring traceability, and supporting secure, structured storage. The design and development phase resulted in a low-code digital solution built using Microsoft Power Apps for the user interface, Microsoft Lists as a centralised relational database, and Power Automate to orchestrate automated workflows. The artifact was then demonstrated through pilot deployment in the laboratory, incorporating real-time feedback from operators and iterative refinements. The evaluation phase assessed efficiency gains, error reduction, and improvements in data accessibility, while the communication phase involved internal reporting and dissemination through academic publication.

The digital artifact included the following main components:

- New Database—A new database was created to compile the information from monthly Excel spreadsheets, using Microsoft Lists. By standardizing name conventions, measurement units, timestamps, and connecting data in the records, previously dispersed, became centralized.

- Data Capture and Storage—The application, developed by using Microsoft Power Apps, was designed to allow direct input of new measurements, storing them automatically in the database. Future integration with IoT sensors (e.g., pH, conductivity, and temperature sensors) was considered for fully automated data acquisition.
- Visualization and Monitoring of the Data—Using Microsoft Power Apps in conjunction with Microsoft Automate, the application can display the data, allowing the user to sort it for analysis. Visuals alerts, automated calculations and automated emails with important information also help this kind of analysis. The new database also allows the exportation of selected data to Microsoft Power BI, for dashboard creation and reporting.
- Error Handling and Traceability—The digital workflow incorporated data validation and traceability were enforced through built-in rules that restrict invalid entries (e.g., out-of-range physicochemical values, missing mandatory fields) and through automatic timestamping and user identification. These mechanisms ensure full traceability of each measurement, significantly reducing the risk of information loss or corruption.

3.2. Exploratory Analysis Using CRISP-DM

Following the digitalisation, the CRISP-DM methodology was applied to organise the predictive and causal analysis [42,43]. This framework was selected for its widespread adoption in industrial analytics and its flexibility when dealing with heterogeneous, real-world datasets, including those generated in laboratory environments [43].

3.2.1. Business Understanding

The analytical phase was designed to explore whether laboratory-collected process parameters could predict or causally explain electrodeposition coating defects.

To this end, the following research sub-questions were formulated:

- Which electrodeposition process parameters have the greatest influence on the occurrence of defects?
- Are the currently collected parameters sufficient and relevant to enable defect prediction?
- To what extent can ML models complement the classical analytical methodology adopted by the laboratory?
- What is the robustness and stability of the different ML models in identifying causal and predictive relationships?
- What limitations or gaps in the current data become evident through the application of ML?
- Could this type of approach serve as a starting point for the future implementation of more advanced decision support tools?

These questions ensured that the analysis went beyond pure predictive accuracy, aiming instead to extract actionable insights for process optimisation and future decision-support systems.

To quantify the impact of manual transcription prior to digitalisation, the time required to transfer laboratory measurements into Excel spreadsheets was measured across different processes. The results are presented in Table 1, showing that total transcription time ranged from approximately 40 min to more than 2.5 h per shift, depending on production variability. For the Minimum Total the values are for the measure of only two cars and the Maximum Total is for fourteen cars, (the maximum that the company produced during the study). The processes listed in Table 1 correspond to measurements performed by the company. For the pre-treatment and electrodeposition processes, these relate to physico-chemical data collected in both processes. The thickness values refer to the layer added during electrodeposition, as well as the final paint layer (enamel) and its gloss. Additionally, the company used to validate the new framework produces a document referred to as L10, which contains information related to production, including the number of defects and cars produced, as well as the number of materials consumed during the pre-treatment and electrodeposition processes.

Table 1. Time used for transcription.

Process	Unit Transcription Time (min)	Minimum Total Transcription Time (min)	Maximum Total Transcription Time (min)
Pre-Treatment (PT)	2.16	4.32	4.32
Electrodeposition (ED)	2.45	2.41	2.41
Production and Consumption (L10)	1.25	1.47	1.47
Chassis Thicknesses	3.25	3.25	3.25
ED Thicknesses	4.74	9.47	66.32
Enamel Thicknesses	4.74	9.47	66.31
Enamel Gloss	4.84	9.67	19.35
Total	23.43	40.07	163.42

3.2.2. Data Understanding and Preparation

The data used in this study, for the predictive and causal models, was collected from the electrodeposition (ED) (see Figure 2) coating process by the laboratory of the automotive factory used to validate this study. Electrodeposition is a critical pre-painting process, responsible for ensuring corrosion resistance and the uniform adhesion of subsequent paint layers. It comprises four main stages. First, the vehicle undergoes a manual wash with demineralized water, followed by immersion in the paint tank. At this stage, the application of an electric current results in the formation of a thin paint uniform film on the metallic surfaces of the vehicle by electrochemical deposition. This process is followed by rinsing and curing phases to stabilise the coating. After curing, the number of surface defects is assessed. It is important to note that unused paint is treated through ultrafiltration to remove impurities and is then reintroduced into the tank, enabling its reuse and minimizing waste.

Given that some variables share similar characteristics, such as pH and conductivity, a standardized nomenclature was established for parameter identification, indicating first the source of the variable and then the measured characteristic (e.g., "paint_ed_ph").

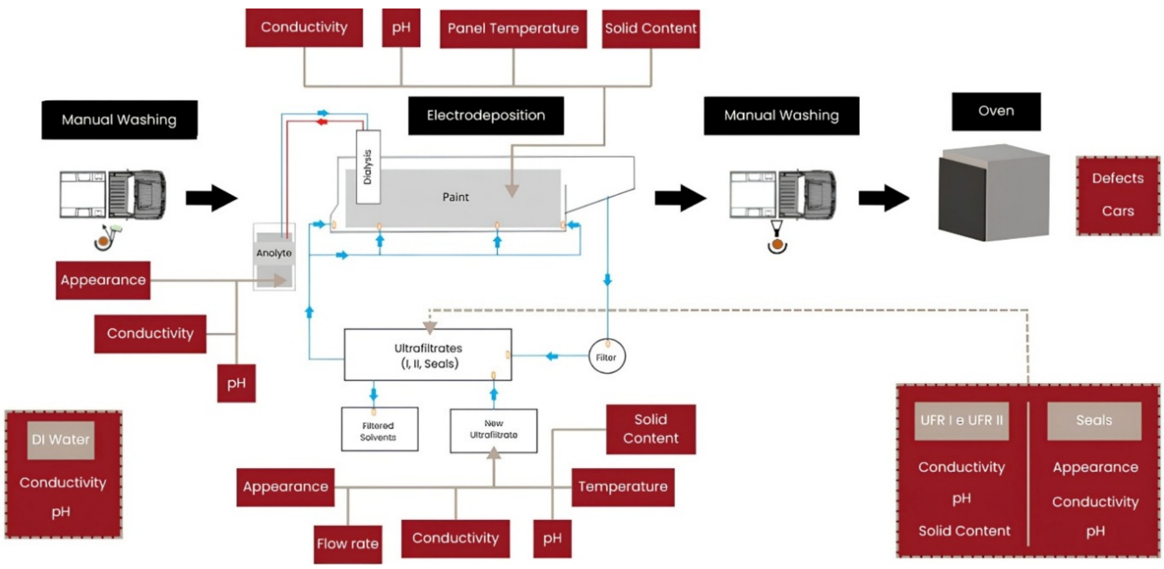

Figure 2. Electrodeposition Process.

Within the laboratory, several physicochemical parameters are routinely monitored to ensure process stability. These parameters include:

- pH and conductivity of the ED bath, of the ultrafiltrate I, of the ultrafiltrate II, and of the anolyte;
- Solid content (%) of ultrafiltrates I and II;
- Bath temperature and appearance assessments (qualitative variables);
- Production and consumption indicators related to paint usage and replenishment cycles.

The target variable for this study—Defects by Car—was created by dividing the total number of ED coating defects recorded during laboratory inspections by the total number of cars that were painted daily. All other variables represent independent process parameters used as predictors to model the defect behavior.

A Python-based data integration pipeline was developed, using Visual Studio Code, to consolidate historical laboratory records from multiple Excel sources into a unified structured dataset (.csv). This pipeline standardised variable naming conventions according to parameter type and measurement origin (e.g., pH_ED_Bath, Cond_UF1, Solid_UF2, Temp_ED, Defects). Data preprocessing followed CRISP-DM best practices and included data cleaning (removal of duplicated, missing, or invalid records), outlier treatment based on physicochemical specifications, and imputation of missing values using forward-fill, backward-fill, or linear interpolation, selected according to each variable's coefficient of variation (CV).

After preprocessing, the resulting dataset (df_corrected) contained 946 valid records—corresponding to 31.94% of all available measurements—with complete defect information (as seen on Table 2) The reduction from the original dataset reflects the elimination of incomplete and inconsistent records rather than random sampling, a factor later considered when interpreting model robustness. To reduce multicollinearity and improve model robustness, a secondary dataset—df_No_Correlations—was generated by removing highly correlated variables (Pearson correlation coefficient $|r| > 0.6$) [44]. This step ensured that the models could better capture independent effects among process parameters. Feature engineering was subsequently applied to derive composite variables, and Recursive Feature Elimination with Cross-Validation (RFECV) was used to identify the most informative feature subset, producing the df_feature_engineering_rfecv dataset. All datasets were split into training (70%) and testing (30%) subsets.

Table 2. Electrodeposition process records.

	Number of Records	Number of Parameters	Incorrect Records	Empty Records	Total of Errors	Information on Defects
Before Treatment	2962	28	11.82%	28.19%	40.01%	31.94%
After Treatment	946	26	0%	0%	0%	100%

3.2.3. Modelling

The predictive modelling phase focused on evaluating whether the laboratory-monitored parameters could be used to predict electrodeposition (ED) coating defects. Three supervised Machine Learning algorithms were tested: Random Forest, XGBoost, and Artificial Neural Networks. Model performance was evaluated using Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE), which are appropriate for continuous regression problems and enable direct interpretation of prediction deviations in both absolute and relative terms.

To establish a reference for performance comparison, a baseline model was defined in which the predicted value was always equal to the median of the target variable. Consequently, the ML models were required to outperform this baseline, corresponding to MAE = 1.25, MAPE = 19.08% and RMSE = 1.70, in order to demonstrate added predictive value beyond trivial estimation.

The target variable used in all predictive models corresponds to the daily defect ratio, calculated as the number of ED coating defects divided by the number of vehicles painted per day. Model predictions were compared against the observed values of this target variable to compute MAE, MAPE, and RMSE.

During model development and hyperparameter tuning, a 5-fold cross-validation strategy [45,46]. was applied to the training data to reduce overfitting and ensure robustness. For each model, performance metrics were calculated on the validation folds and averaged to obtain stable estimates. The final performance values reported in this study correspond to the mean results obtained on the held-out test sets, ensuring that the metrics reflect the generalisation capability of the models rather than their fit to the training data.

In parallel with predictive modelling, causal inference analysis was conducted to identify process variables with statistically significant causal effects on defect formation. Four causal models were applied: Simple Linear Regression, DoWhy, Causal Forest, and Double Machine Learning (DML). These models were trained using the same preprocessed datasets to ensure methodological consistency. The combination of predictive and causal approaches allowed not only the evaluation of forecasting performance but also the identification and interpretation of key process drivers influencing defect occurrence.

All modelling and statistical analyses were implemented in Python, by means of Visual Studio Code, using established libraries for machine learning and causal inference. Residual analysis was performed to assess model stability and detect potential systematic biases. If the residual distributions are approximately normal and centered around zero, this is an indicative of no systematic bias or overfitting [47].

4. Results and Discussions

4.1. Impact of Digitalisation

The transition from analogue data recording to a digitalised workflow produced substantial operational and informational gains. Before the implementation of the new system, operators spent between 40 min and 2 h per shift (≈34% of total working time) transcribing measurements into Excel spreadsheets. This process was completely eliminated after the deployment of the digital platform, freeing time for higher-value analytical activities.

As illustrated in Figure 3, this transformation led to a substantial improvement in operational efficiency by completely eliminating the need for data transcription in the implemented solution. The minimum values represent the smallest number of vehicles measured by the operator on any given day during the study period, whereas the maximum values correspond to days with the highest measurement volume, driven by laboratory requirements. This variability occurred across different days rather than within a single shift. As the number of vehicles measured increased, the time required for manual data transcription grew proportionally, significantly impacting operator workload. The digitalized solution demonstrates the effect of eliminating manual transcription—a task whose duration scaled directly with measurement volume. By removing this activity, the operator’s workload becomes less sensitive to daily fluctuations, thereby freeing time for analytical and value-added tasks. A comparative analysis before and after the implementation, using key performance indicators defined in collaboration with the operators and supervisors of the laboratory—efficiency, error reduction, analysis time, flexibility, and data security—demonstrated an improvement from scores of 1–2 (pre-digitalisation) to 4–5 (post-digitalisation) (see Figure 4). This KPIs were measured by retrieving data at the start of the project and at the end, being translated to the present scale by mathematical calculus, (see Figure 4).

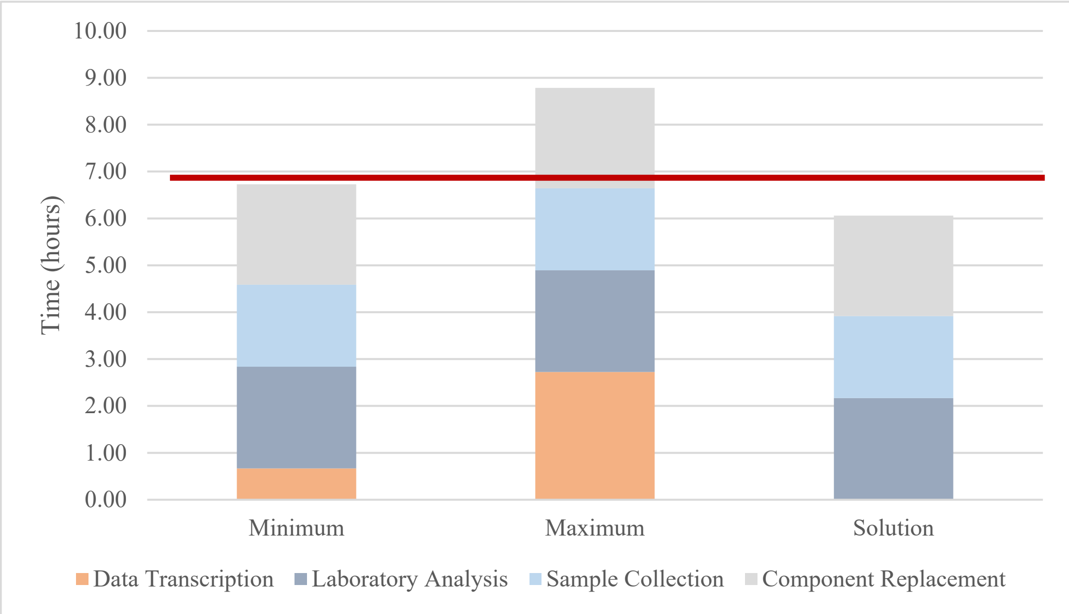

Figure 3. Comparison of activity times.

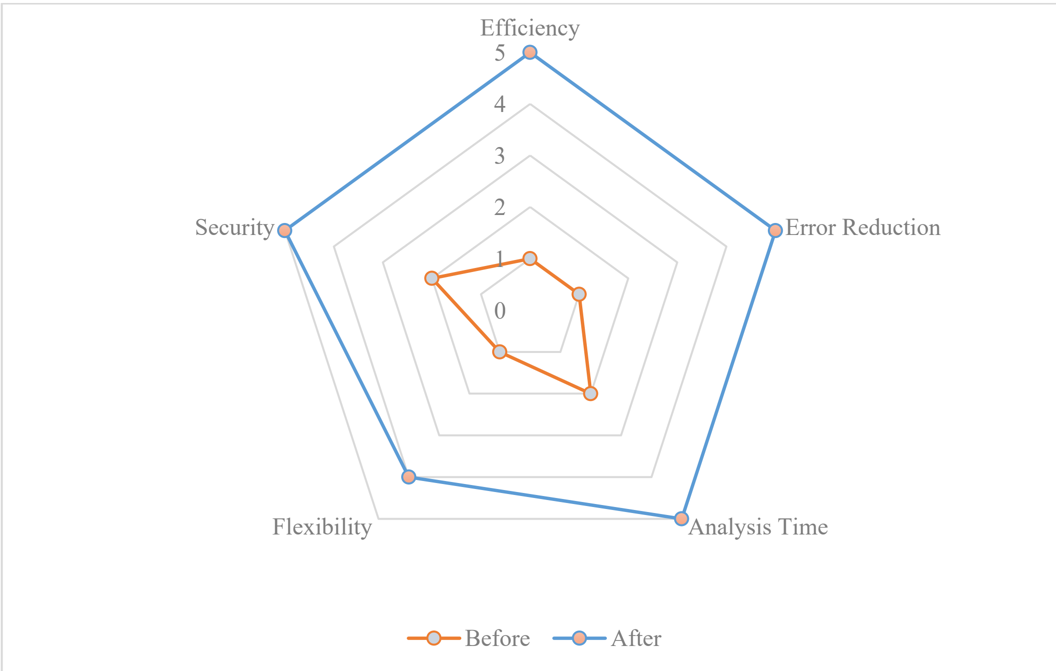

Figure 4. Comparison of process indicators: before vs. after digitalization.

Error rates also improved markedly. Historical data indicated that approximately 40.01% of all records contained transcription errors due to manual entry and repeated file manipulation. Through automated data validation and real-time database synchronisation, the digital workflow reduced these errors to virtually zero. These findings are consistent with the literature, where Pop [2] and Verhoef et al. [6] reported similar improvements in accuracy and traceability after adopting digital tools in industrial environments.

Furthermore, process flexibility was enhanced. In scenarios such as sudden changes in paint type, previously requiring new paper templates or manual updates, the digital platform enabled immediate adaptation by simply updating input parameters in the database. This adaptability is characteristic of low-code platforms, which according to Bock and Frank [7], significantly reduce development and change management time.

Finally, the data security and governance aspects were also strengthened. The new system introduced user authentication, permission levels, and audit trails, ensuring traceability of all operations and compliance with data integrity standards. Previously, Excel-based workflows lacked version control, increasing the risk of accidental or malicious data loss. The implemented digital approach aligns with best practices reported by Haddara and Zach [9] in ERP-based environments, confirming that data governance is a key benefit of industrial digitalisation.

Overall, these results justify the scientific relevance of digitalisation not only as an operational improvement but as a data quality enabler, without which the subsequent machine learning analyses would not be statistically reliable.

4.2. Predictive Modelling Results

4.2.1. Initial Dataset (df_corrected)

Using the df_corrected dataset, the XGBoost model achieved the best results among the tested algorithms, with MAE = 1.002, MAPE = 17.02%, and RMSE = 1.343. Although these values indicate a reasonable fit, they remain above the industry threshold for high precision (MAPE < 10%) [48].

Residual analysis (see Figure 5) showed an approximately normal distribution, with mild skewness (−0.22) and kurtosis (2.13), suggesting that the model captured the main structure of the data but slightly underestimated extreme defect counts. This behavior is expected given the limited dataset size (946 observations) and the non-linear relationships inherent to the electrodeposition process. The Residue Distribution also let us see that the models don't overfit the data, since the curve is approximately normal [47].

Physicochemical variations in parameters such as pH, conductivity, and ultrafiltrate solids can produce non-proportional effects on coating defects, behavior that tree-based models can partially capture, but only with sufficient training diversity.

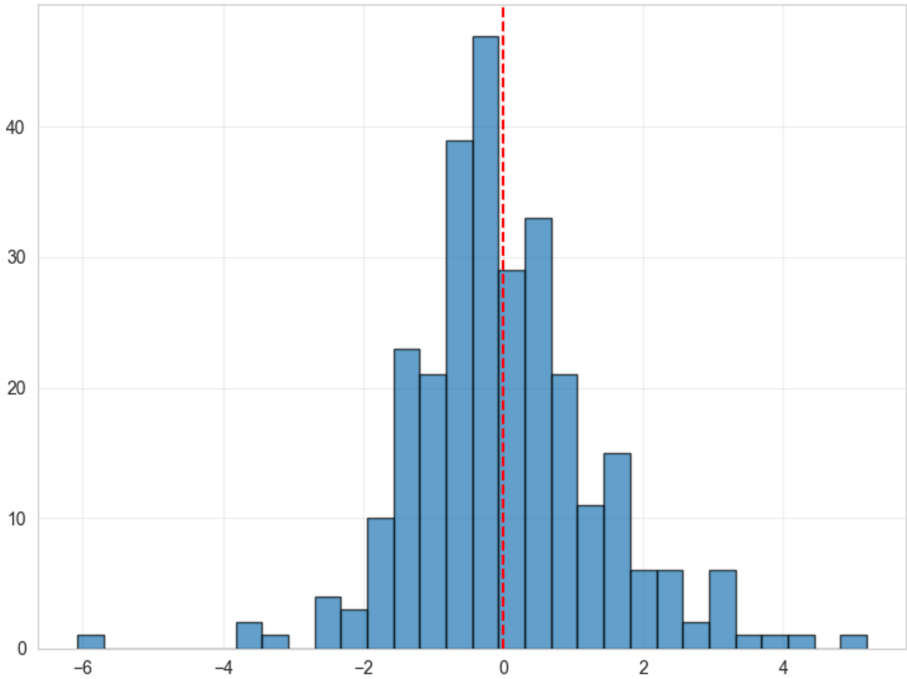

Figure 5. Residue distribution (residues/frequency).

4.2.2. Reduced Correlations Dataset (df_No_Correlations)

When highly correlated features ($|r| > 0.6$) were removed to produce the df_No_Correlations dataset, Random Forest outperformed all other algorithms, achieving MAE = 0.972, MAPE = 16.45%, and RMSE = 1.307.

The improvement indicates that multicollinearity among physicochemical variables (for instance, between pH and conductivity in ultrafiltrate solutions) was reducing model interpretability and increasing noise.

After removing redundant predictors, model generalisation improved, as seen by the tighter residual dispersion around zero.

This result supports the findings emphasised that reducing inter-variable dependency enhances predictive robustness and interpretability in multivariate industrial datasets.

4.2.3. Feature Engineering (df_feature_engineering) and RFECV Datasets (df_feature_engineering_rfecv)

The df_feature_engineering dataset, which introduced derived attributes such as ratios between conductivity and solids, initially produced slightly inferior results (MAE = 1.025; MAPE = 16.84%; RMSE = 1.417). This outcome suggests that not all engineered features provided meaningful new information, possibly introducing collinearity or noise.

Conversely, applying Recursive Feature Elimination with Cross-Validation (RFECV) generated a more parsimonious feature subset (df_feature_engineering_rfecv), achieving MAE = 0.977; MAPE = 16.70%; RMSE = 1.323. When averaging all models across datasets, this configuration delivered the best overall performance (MAE = 1.051; MAPE = 17.63%; RMSE = 1.403).

This demonstrates that structured feature selection contributes more effectively to predictive accuracy than indiscriminate feature expansion, a conclusion aligned with Peres et al. [19], who observed similar results in automotive quality control using tree-based ensembles.

4.2.4. Interpretation

Overall, Random Forest consistently provided the most stable and interpretable results across datasets, confirming its robustness for heterogeneous, small-to-medium industrial datasets.

Nevertheless, none of the models achieved high-precision thresholds (MAPE < 10%), primarily due to the limited number of defect-labelled records and the narrow variability of process parameters, which constrained model learning.

From a process standpoint, this implies that the parameters currently collected by the laboratory are only partially explanatory of defect formation. Important factors, such as bath ageing, circulation flow rates, or surface pre-treatment residues, are not digitally captured, limiting the predictive scope.

Hence, while ML models demonstrated potential, their performance justifies the need for additional data acquisition and sensor integration (e.g., real-time pH and temperature logging) to expand model accuracy and causal interpretability.

4.3. Causal Modelling Results

Causal inference analysis complemented the predictive models by identifying which process variables exhibited statistically significant causal relationships with defect counts. The models created and tested in Python—Linear Regression, DoWhy, Causal Forest, and Double Machine Learning (DML)—were trained using both the original and correlation-reduced datasets, using the same data handling approach described in the Data Preparation phase of the CRISP-DM methodology.

Across all models, the solid content of ultrafiltrate I consistently emerged as the dominant causal driver of coating defects. This finding aligns with the process knowledge of electrodeposition, as variations in solid concentration alter bath conductivity and paint particle availability, directly influencing deposition uniformity.

Other variables with moderate causal influence included the Ph of the ED bath and the Ph of the ultrafiltration solutions, both of which affect electrochemical balance and particle migration rates during deposition.

Model performance metrics (calculation of the R^2 of the models) for the df_corrected dataset are summarised as follows (Figure 6):

- Linear Regression: 0.954
- DoWhy: 0.880
- Causal Forest: 0.747
- DML: 0.710

These results indicate that simpler linear models provided sufficiently robust causal estimates given the dataset size, corroborating findings by Luckow et al. [22], who noted that more complex models do not always outperform simpler approaches when data volume is limited.

The consistency of results across both datasets (original and correlation-reduced) reinforces the validity of the identified causal relationships. Collectively, these findings justify the conclusion that the ultrafiltrate I solid content is a key control parameter for defect mitigation, and that causal analysis can serve as a diagnostic complement to traditional quality monitoring, guiding future process optimisation.

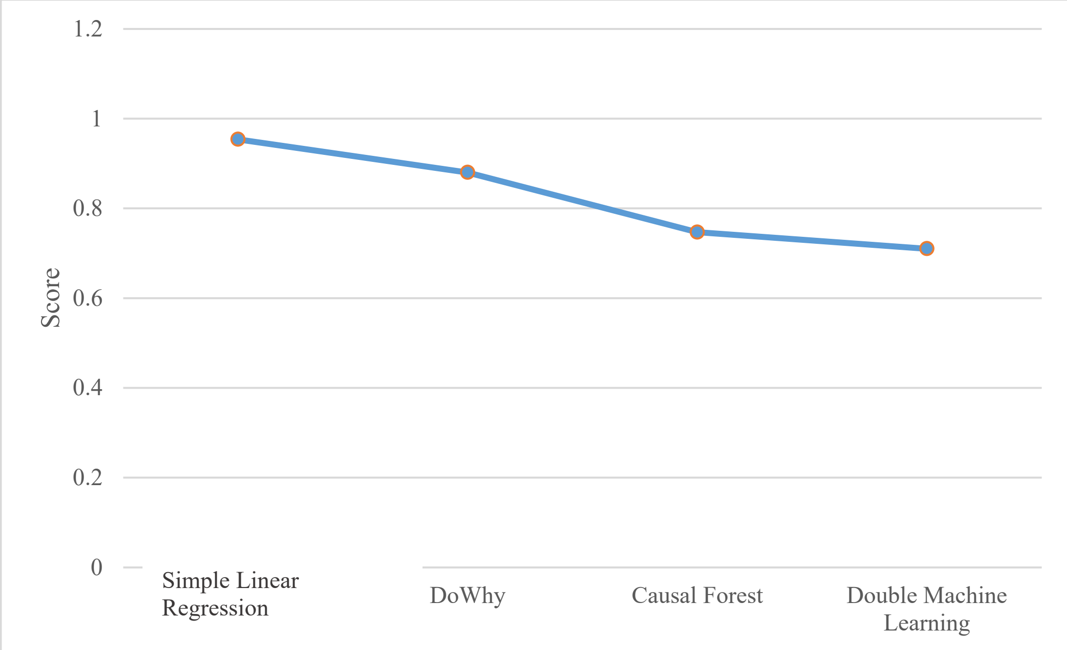

Figure 6. Causal models scores.

5. Critical Analysis

The results of this study clearly demonstrate that digitalisation represented the most transformative intervention, delivering immediate operational and informational benefits. The elimination of manual transcription, which previously consumed up to 34% of the operators' working time, aligns with findings by Pop [2] and Mehta and Rastogi [3], who reported efficiency gains of 25–35% in similar industrial digitalisation projects. Moreover, the near-total removal of transcription errors (previously 40.01%) confirms that low-code digital tools can achieve reliability levels comparable to integrated ERP or MES systems, as also highlighted by Verhoef et al. [6] and Haddara and Zach [9]. These outcomes reinforce that digitalisation is not only an operational enabler but a scientific prerequisite for reliable analytics, since data integrity directly determines model performance.

The predictive modelling results further emphasise this dependency. While the Random Forest and XGBoost algorithms achieved Mean Absolute Percentage Errors (MAPE) between 16% and 17%, these values remain above the high-precision industrial benchmark (MAPE < 10%) defined by Hyndman and Koehler [49]. Similar results were reported by Peres et al. [19] in their study at Volkswagen AutoEuropa, where tree-based ensembles achieved strong but not perfect prediction of dimensional defects, due to unmeasured environmental variables. The present study corroborates these challenges: the limited dataset (946 labelled records) and the absence of key variables such as bath ageing or circulation rate constrained model learning. This finding is consistent with Yao and Feng [27] and Luckow et al. [22], who showed that data completeness and sensor integration are critical for achieving robust industrial ML models.

The causal inference analysis provided valuable complementary insights by identifying ultrafiltrate I solid content as the most influential causal factor in defect formation, a result that is coherent with electrochemical process theory and prior research by Fu et al. [50] on electrodeposition bath dynamics. The fact that simpler models (e.g., linear regression) produced causal estimates comparable to more complex methods, loke Causal Forest and Double Machine Learning, confirms the literature's observation that model complexity does not necessarily guarantee better causal interpretability in small-sample industrial datasets [51,52].

Overall, this study contributes by empirically validating that digitalisation combined with ML and causal analysis can yield actionable insights even under resource and data constrains, a condition frequently encountered in laboratory-scale operations. The observed efficiency gains and the identification of meaningful causal variables align with the broader shift toward smart manufacturing and data-driven quality control noted by Brennen and Kreiss [5] and Verhoef et al. [6]. However, the work also reinforces a key limitation repeatedly stressed in the literature: data quality, consistency, and representativeness remain major obstacles to fully leveraging AI in manufacturing. Future work should therefore prioritise real-time sensor integration, expanded data collection, and model retraining, enabling predictive systems capable of near-real-time defect prevention and continuous process optimization.

6. Conclusions

This research addressed the question:

> "*How can the efficiency of data collection and processing in an industrial painting laboratory be increased through digitalisation mechanisms and data-driven analysis?*"

The findings demonstrate that the integration of low-code digitalisation tools with exploratory Machine Learning (ML) and causal inference models provides a novel and effective framework for enhancing both operational efficiency and analytical capability in laboratory environments. The developed Microsoft Power Apps platform, supported by centralised databases and automated workflows, eliminated manual transcription tasks, previously consuming up to 34% of operational time, and reduced transcription errors (≈40%) to virtually zero. This confirms the transformative potential of digitalisation not only as an operational improvement but as an enabler of trustworthy, analytics-ready data, a prerequisite often overlooked in industrial research.

From an analytical standpoint, this study offers a novel integration of causal inference and predictive ML in the context of electrodeposition quality analysis, using real laboratory data. While the Random Forest model achieved moderate predictive accuracy (MAE ≈ 1.0; MAPE ≈ 16–17%), the causal analysis consistently identified the ultrafiltrate I solid content as the primary factor influencing coating defects. This dual methodological approach offers both predictive and explanatory insights—bridging a gap in existing literature where causal understanding is rarely integrated into industrial ML studies.

The novelty of this work lies in the proposed framework that integrates digitalization tools, machine learning techniques, and causal analysis, and applies them to a manufacturing support laboratory, something that was not found in other literature, since most existing works focus on the processes occurring in production lines rather than on companies' support areas, such as laboratories.

This study demonstrates that combining digitalization, machine learning, and causal analysis can generate actionable insights even under resource and data constraints—a scenario frequently encountered in laboratory-scale operations within the automotive manufacturing sector.

The main limitations of this study include the relatively small dataset, the lack of additional sensors that could enhance data collection, and the resulting constraints on model accuracy. Beyond the automotive sector, the proposed framework is generalizable to other industries facing similar challenges of fragmented data collection and limited process digitalisation, such as chemical processing, pharmaceuticals, food manufacturing, and energy systems. By demonstrating that substantial analytical and operational value can be extracted even from constrained datasets, this work provides a replicable model for laboratories and production units seeking to transition toward smart, data-driven quality control.

6.1. Limitations

- The main limitations of this study were related to data collection, which was carried out manually. This approach constrained the historical volume of data obtained with a likelihood of transcription errors during transfer to the digital system. Due to this the dataset had to reduce from 2962 into 946 data examples (Table 2).
- The number of variables collected by the laboratory used to validate the study did not include certain variables from the electrodeposition phase, such as manual rinsing steps or the temperature values of the drying oven—variables that may contain relevant information for the models.
- The laboratory-only scope, which does not currently incorporate upstream or downstream production data.

6.2. Future Work

- Future research should address these limitations by expanding the dataset, increasing operational variability, and integrating data from full-scale production processes to enhance the generalizability and robustness of the proposed framework. This can be achieved dynamically and in a fully automated manner through:
- Automating data acquisition via IoT-connected sensors to ensure continuous, high-frequency data capture;
- Developing interactive dashboards for real-time monitoring of process indicators;
- Extending the model scope to incorporate production line variables and multivariate time-series models, enabling predictive maintenance and supporting digital twin applications.

In summary, this work presents a validated proof-of-concept that uniquely integrates low-code digitalisation with machine learning and causal analysis to enhance quality control in industrial laboratories. It contributes practically, by demonstrating measurable efficiency and data-quality improvements, and scientifically, by establishing a novel, transferable framework for data-driven process optimisation that can be extended beyond the automotive sector to other manufacturing domains within the paradigm of Industry 4.0.

Author Contributions

M.T.R.P., M.J.G.P., A.M.G. and M.G.T.: conceptualisation; M.T.R.P., M.G.P and M.G.T.: methodology; M.T.R.P., M.J.G.P., A.M.G., H.V. and M.G.T.: validation; M.T.R.P., M.J.G.P. and M.G.T.: formal analysis; M.G.T.: investigation; M.T.R.P., M.J.G.P. and M.G.T.: data curation; M.G.T.: writing original draft preparation; M.T.R.P., M.J.G.P., A.M.G., H.V. and M.G.T.: writing review and editing; M.T.R.P., M.J.G.P., A.M.G. and M.G.T.: visualisation; M.T.R.P.,M.J.G.P., A.M.G., H.V. and: supervision; M.G.T.: project administration. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding

Institutional Review Board Statement

Not applicable

Informed Consent Statement

Not applicable.

Data Availability Statement

Data available on demand.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

No AI tools were utilized for this paper.

References

1. Pinto, B.; Silva, F.J.G.; Costa, T.; et al. A strategic model to take the first step towards industry 4.0 in SMEs. *Procedia Manuf.* **2019**, *18*, 637–645. https://doi.org/10.1016/j.promfg.2020.01.082.
2. Pop, L.D. Digitalization of the System of Data Analysis and Collection in an Automotive Company. *Procedia Manuf.* **2020**, *46*, 238–243. https://doi.org/10.1016/j.promfg.2020.03.035.
3. Mehta, S.; Rastogi, A.K. Impact of Digitalization on Automotive Industry: Challenges & Opportunities. *BISM* **2017**, *1*, 81–92. Available: http://bism2017.globalconf.org.
4. Baiyere, A.; Salmela, H.; Tapanainen, T. Digital transformation and the new logics of business process management. *Eur. J. Inf. Syst.* **2020**, *29*, 238–259. https://doi.org/10.1080/0960085X.2020.1718007.
5. Brennen, J.S.; Kreiss, D. Digitalization. In *The International Encyclopedia of Communication Theory and Philosophy*; Wiley: Hoboken, NJ, USA, 2016; pp. 1–11. https://doi.org/10.1002/9781118766804.wbiect111.
6. Verhoef, P.C.; Broekhuizen, T.; Bart, Y.; et al. Digital transformation: A multidisciplinary reflection and research agenda. *J. Bus. Res.* **2021**, *122*, 889–901. https://doi.org/10.1016/j.jbusres.2019.09.022.
7. Bock, A.C.; Frank, U. Low-Code Platform. *Bus. Inf. Syst. Eng.* **2021**, *63*, 733–740. https://doi.org/10.1007/s12599-021-00726-8.
8. Waszkowski, R. Low-code platform for automating business processes in manufacturing. *IFAC PapersOnLine* **2019**, *52*, 376–381. https://doi.org/10.1016/j.ifacol.2019.10.060.
9. Haddara, M.; Zach, O. ERP Systems in SMEs: A Literature Review. In Proceedings of the 44th Hawaii International Conference on System Sciences, Kauai, HI, USA, 4–7 January 2011. https://doi.org/10.1109/HICSS.2011.191.
10. Akhtar, S.; Shahzad, S.; Zaheer, A.; et al. Short-Term Load Forecasting Models: A Review of Challenges, Progress, and the Road Ahead. *Energies* **2023**, *16*, 4060. https://doi.org/10.3390/en16104060.
11. Benti, N.E.; Chaka, M.D.; Semie, A.G. Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects. *Sustainability* **2023**, *15*, 7087. https://doi.org/10.3390/su15097087.
12. Chatfield, C.; Xing, H. *The Analysis of Time Series*, 7th ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2019. https://doi.org/10.1201/9781351259446.

13. Pesaran, M.H.; Timmermann, A. How costly is it to ignore breaks when forecasting the direction of a time series? *Int. J. Forecast.* **2004**, *20*, 411–425. https://doi.org/10.1016/S0169-2070(03)00068-2.
14. Yaseen, Z.M. A New Benchmark on Machine Learning Methodologies for Hydrological Processes Modelling: A Comprehensive Review for Limitations and Future Research Directions. *Knowl. Based Eng. Sci.* **2023**, *4*, 65–103. https://doi.org/10.51526/kbes.2023.4.3.65-103.
15. Lolla, R.; Harper, M.; Lunn, J.; et al. Machine Learning Techniques for Predicting Risks of Late Delivery. In *Data Science and Emerging Technologies*; Springer: Singapore, 2023; pp. 343–356. https://doi.org/10.1007/978-981-99-0741-0_25.
16. Liu, X.; Wang, W. Deep Time Series Forecasting Models: A Comprehensive Survey. *Mathematics* **2024**, *12*, 1504. https://doi.org/10.3390/math12101504.
17. Lopes, D.R.; Ramos, F.R.; Costa, A.; et al. Forecasting Models for Time-Series: A Comparative Study Between Classical Methodologies and Deep Learning. In Proceedings of the SPE 2021—XXV Congresso Da Sociedade Portuguesa de Estatística, Online, 13–16 October 2021. https://doi.org/10.13140/RG.2.2.12559.92328.
18. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. https://doi.org/10.1007/s12525-021-00475-2.
19. Peres, R.S.; Barata, J.; Leitao, P.; Garcia, G. Multistage Quality Control Using Machine Learning in the Automotive Industry. *IEEE Access* **2019**, *7*, 79908–79916. https://doi.org/10.1109/ACCESS.2019.2923405.
20. Crowley, J.L. Pattern Recognition and Machine Learning Generative Networks: EigenSpace Coding, Auto-Encoders, Variational Autoencoders and Generative Adversarial Networks. 2020.
21. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2018.
22. Luckow, A.; Cook, M.; Ashcraft, N.; et al. Deep Learning in the Automotive Industry: Applications and Tools. In Proceedings of the 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, USA, 5–8 December 2016; pp. 3759–3768. https://doi.org/10.1109/BigData.2016.7841045.
23. Samek, W.; Wiegand, T.; Müller, K.-R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *arXiv* **2017**, arXiv:1708.08296.
24. Ghosal, S.G.; Dey, S.; Chattopadhyay, P.P.; et al. Designing optimized ternary catalytic alloy electrode for efficiency improvement of semiconductor gas sensors using a machine learning approach. *Decis. Mak. Appl. Manag. Eng.* **2021**, *4*, 126–139. https://doi.org/10.31181/dmame210402126g.
25. Ramos, F.R.; Pereira, M.T.; Oliveira, M.; et al. The memory concept behind deep neural network models: An application in time series forecasting in the e-Commerce sector. *Decis. Mak. Appl. Manag. Eng.* **2023**, *6*, 668–690. https://doi.org/10.31181/dmame622023695.
26. Marchese Robinson, R.L.; Palczewska, A.; Palczewski, J.; et al. Comparison of the Predictive Performance and Interpretability of Random Forest and Linear Models on Benchmark Data Sets. *J. Chem. Inf. Model.* **2017**, *57*, 1773–1792. https://doi.org/10.1021/acs.jcim.6b00753.
27. Yao, B.; Feng, T. Machine learning in automotive industry. *Adv. Mech. Eng.* **2018**, *10*. https://doi.org/10.1177/1687814018805787.
28. Tang, T. Analysis and Demand Forecasting Based On e-Commerce Data. In Proceedings of the 2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 26–29 May 2023; pp. 64–68. https://doi.org/10.1109/ICAIBD57115.2023.10206072.
29. Fouad, A.M.; Wefki, H.; Kouta, H. AI-Driven Raw Material Demand Forecasting: Towards Project Management Practices. *J. Supply Chain Manag. Sci.* **2025**, *6*. https://doi.org/10.59490/jscms.2025.8136.
30. Ma, B.J.; Jackson, I.; Huang, M.; et al. A data-driven and context-aware approach for demand forecasting in the beverage industry. *Int. J. Logist. Res. Appl.* **2025**, *6*. https://doi.org/10.59490/jscms.2025.8136.
31. Hammam, I.M.; El-Kharbotly, A.K.; Sadek, Y.M. Adaptive demand forecasting framework with weighted ensemble of regression and machine learning models along life cycle variability. *Sci. Rep.* **2025**, *15*, 38482. https://doi.org/10.1038/s41598-025-23352-w.
32. Santoro, D.; Ciano, T.; Ferrara, M. A comparison between machine and deep learning models on high stationarity data. *Sci. Rep.* **2024**, *14*, 19409. https://doi.org/10.1038/s41598-024-70341-6.
33. Yao, L.; Chu, Z.; Li, S.; et al. A Survey on Causal Inference. *ACM Comput. Surv.* **2021**, *54*, 1–46. https://doi.org/10.1145/3444944.
34. Kaddour, J.; Lynch, A.; Liu, Q.; et al. Causal Machine Learning: A Survey and Open Problems. *arXiv* **2022**, arXiv:2206.15475.
35. Saltz, J.S.; Hotz, N. Identifying the Most Common Frameworks Data Science Teams Use to Structure and Coordinate Their Projects. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 2038–2042. https://doi.org/10.1109/BigData50022.2020.9377813.
36. Schröer, C.; Kruse, F.; Gómez, J.M. A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Comput. Sci.* **2021**, *181*, 526–534. https://doi.org/10.1016/j.procs.2021.01.199.

37. Shafique, U.; Qaiser, H. A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA). *Int. J. Innov. Sci. Res.* **2014**, *12*, 217–222.
38. Huang, H.N.; Chen, H.M.; Lin, W.W.; et al. Employing feature engineering strategies to improve the performance of machine learning algorithms on echocardiogram dataset. *Digit. Health* **2023**, *9*. https://doi.org/10.1177/20552076231207589.
39. Ara Mowri, R.; Siddula, M.; Roy, K. A Comparative Performance Analysis of Explainable Machine Learning Models with and without RFECV Feature Selection Technique Towards Ransomware Classification. *IEEE Access* **2022**, *10*, 130180–130195. https://doi.org/10.1109/ACCESS.2022.3228041.
40. Paolanti, M.; Romeo, L.; Felicetti, A.; et al. Machine Learning Approach for Predictive Maintenance in Industry 4.0. In Proceedings of the 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), Oulu, Finland, 2–4 July 2018. https://doi.org/10.1109/MESA.2018.8449150.
41. Peffers, K.; Tuunanen, T.; Rothenberger, M.A.; Chatterjee, S. A Design Science Research Methodology for Information Systems Research. *J. Manag. Inf. Syst.* **2007**, *24*, 45–77. https://doi.org/10.2753/MIS0742-1222240302.
42. Eiras, E.; Silva, F.J.G.; Campilho, R.D.S.G.; et al. A Novel Fully Automatic Concept to Produce First Subset of Bowden Cables, Improving Productivity, Flexibility, and Safety. *Machines* **2023**, *11*, 992. https://doi.org/10.3390/machines11110992.
43. Tojal, M.C.; Silva, F.J.G.; Campilho, R.D.S.G.; et al. Case-Based Product Development of a High-Pressure Die Casting Injection Subset Using Design Science Research. *FME Trans.* **2022**, *50*, 32–44. https://doi.org/10.5937/fme2201032T.
44. Schober, P.; Boer, C.; Schwarte, L.A. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth. Analg.* **2018**, *126*, 1763–1768. https://doi.org/10.1213/ANE.0000000000002864.
45. Szeghalmy, S.; Fazekas, A. A Comparative Study of the Use of Stratified Cross-Validation and Distribution-Balanced Stratified Cross-Validation in Imbalanced Learning. *Sensors* **2023**, *23*, 2333. https://doi.org/10.3390/s23042333.
46. Ramezan, C.A.; Warner, T.A.; Maxwell, A.E. Evaluation of Sampling and Cross-Validation Tuning Strategies for Regional-Scale Machine Learning Classification. *Remote Sens.* **2019**, *11*, 185. https://doi.org/10.3390/rs11020185.
47. Li, W.; Cook, D.; Tanaka, E.; et al. A Plot is Worth a Thousand Tests: Assessing Residual Diagnostics with the Lineup Protocol. *J. Comput. Graph. Stat.* **2024**, *33*, 1497–1511. https://doi.org/10.1080/10618600.2024.2344612.
48. Lewis, C.D. Industrial and business forecasting methods. *J. Forecast.* **1983**, *2*, 194–196.
49. Hyndman, R.J.; Koehler, A.B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688.
50. Fu, X.; Li, J.; Zhang, H.; et al. Simulation and experimental research on nickel-based coating prepared by jet electrodeposition at different scanning speeds. *Int. J. Adv. Manuf. Technol.* **2022**, *121*, 1851–1866. https://doi.org/10.1007/s00170-022-09308-8.
51. Pearl, J. The seven tools of causal inference, with reflections on machine learning. *Commun. ACM* **2019**, *62*, 54–60. https://doi.org/10.1145/3241036.
52. Schölkopf, B.; Locatello, F.; Bauer, S.; et al. Toward Causal Representation Learning. *Proc. IEEE* **2021**, *109*, 612–634. https://doi.org/10.1109/JPROC.2021.3058954.