A data-driven regression framework for predictive maintenance using process parameters in manufacturing operations

Fredrick Nnaemeka Okeagu[1]*, Raphael Olumese Edokpia[2], Emmanuel Ikpoza[2]

[1]Department of Industrial/Production Engineering, Nnamdi Azikiwe University, Awka, Nigeria
[2]Department of Production Engineering, University of Benin, Benin City, Nigeria
*Corresponding author, email: fn.okeagu@unizik.edu.ng

Article history
Submitted: 10 November 2025
Revised: 15 December 2025
Accepted: 27 December 2025
Published: 30 December 2025

Keywords
Downtime
Plastic manufacturing
Predictive maintenance
Process parameters
Regression framework

Abstract
Unplanned machine downtime remains a persistent challenge in manufacturing industries, resulting in productivity loss, increased operational costs, and reduced equipment reliability. This study develops a data-driven regression framework for predicting machine downtime using measurable process parameters from plastic manufacturing operations. Real production data were obtained from Cup and Plate production lines, with uptime, product weight, cycle time, and production quantity adopted as the independent predictor variables, while downtime served as the dependent response variable. Descriptive data analysis was conducted to understand the statistical behavior and variability of the datasets. Multiple linear regression modeling was then employed to establish quantitative relationships between the process parameters and downtime. The developed models demonstrated strong goodness-of-fit statistics and revealed that downtime can be significantly explained as a function of key operational parameters. The study provides empirical evidence that integrating process parameters into predictive modeling offers an effective foundation for predictive maintenance decision-making. This work establishes a robust modeling framework upon which further validation, sensitivity analysis, and economic assessment are subsequently built. The outcomes support proactive maintenance planning, enhance understanding of downtime behavior, and contribute to the advancement of data-driven maintenance intelligence in manufacturing environments.

1. Introduction

Efficient maintenance management is a critical determinant of productivity, profitability, and competitiveness in modern manufacturing systems. Among the various challenges faced in production environments, machine downtime remains one of the most disruptive, as it leads to production interruptions, throughput reduction, poor delivery performance, and increased operational expenditure. Traditional maintenance approaches such as corrective or time-based preventive maintenance have proven insufficient, as they either respond after failure or rely on fixed schedules that do not consider the actual operating condition of machines. Consequently, industries are increasingly adopting predictive maintenance strategies driven by operational data and analytical models capable of forecasting potential failure conditions before they escalate into downtime events [1].

Predictive maintenance relies fundamentally on the ability to extract meaningful information from process data, establish relationships between operating conditions and machine behaviour, and use such knowledge to predict future performance. In manufacturing environments, measurable process parameters such as uptime, product weight, cycle time, and production quantity inherently reflect the state of machine operation [2]. Their variations often provide early indicators of system instability or stress conditions that may eventually manifest as downtime. However, many existing studies in predictive maintenance literature have predominantly focused on heavy industries, rotating machinery, or purely sensor-based approaches, with limited emphasis on production-driven parameters in plastic manufacturing operations.

The total production downtime for each production line per shift and developed an algorithm for an uptime maximization [3]. A case study was conducted on one of the leading Nigerian plastic manufacturing firms experiencing a decline in its production efficiency. Robust interview sessions supplemented with well-structured questionnaire were used for data gathering. The data on machine maintenance effectiveness and process pit-stop were evaluated. Also, a factory layout of the company was obtained to highlight some issues involving material movement to the plant floor. A multiple regression analysis was also used to compare downtime and other variables such as cycle time, capacity, weight, overall equipment effectiveness (OEE). From the optimization model, uptime was maximized by 332 min per shift operation thereby reducing the downtime.

© 2025 The Author(s)
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 3110-1577

A state-of-the-art sensor utilization and applications in the injection molding industry, ranging from additional sensors utilized in sensor-integrated injection molding machines to advanced sensor applications incorporating artificial intelligence in the injection molding industry in the future [4]. The selection and utilization of ML techniques, including Linear Regression (LR), Artificial Neural Network (ANN), Decision Tree (DT), K-nearest neighbors (KNN), and Cluster Analysis (CA) [5]. Implemented at the Innovative Plastic Manufacturing Company in Egypt, this research enhances the consistency of plastic bottle production by addressing issues such as surface marks, flashes, bubbles, and variations in liter capacity. Ashhar (2025) proposed a blind calibration method with occasional physical calibration of one pilot sensor. The physical calibration of the pilot sensor is used to disseminate the traceability and avoid systematic drift while doing data-driven calibration of five temperature sensors in an injection molding machine [6]. Plastic washing machine from locally available materials suitable for washing polythene materials [7]. The primary motivation of the work is to save time and labor and wash efficiently, the wash time and mass of polythene sachets washed are 0.02 kg in 3 minutes, 0.04 kg in 4 minutes and 0.06 kg in 5 minutes respectively. The washing machine was found to be effective.

1.1. Machine learning in predictive maintenance

Machine learning algorithms can be categorized into several types can be seen in Figure 1 and Figure 2.

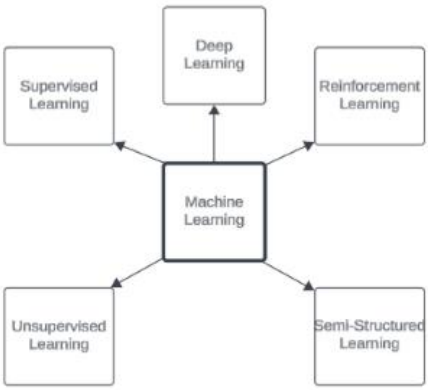


Figure 1. Types of Machine Learning [8]

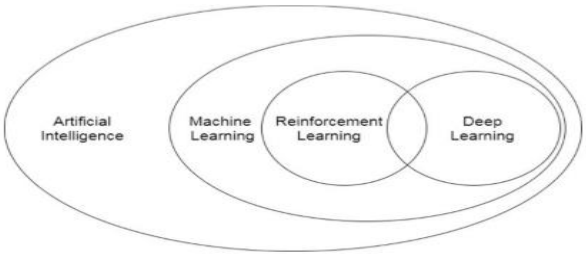


Figure 2. Sub-Classes of Artificial Intelligence [8]

Machine learning is a subset of artificial intelligence that focuses on developing algorithms and models that can learn from data and make predictions or take actions without being explicitly programmed [9]. It involves the use of mathematical and statistical techniques to create models that can generalize patterns and make accurate predictions or decisions on new, unseen data.

1.1.1. Supervised Learning

That in supervised learning, the algorithm learns from labeled training data that consists of input variables and corresponding output labels, Figure 3 [10]. The goal is to build a model that can make accurate predictions or classifications on new, unseen data. When the target variable that we're trying to predict is continuous, such as in housing price prediction, we call the learning problem a regression problem. Examples of supervised learning algorithms include decision trees, support vector machines, and neural networks. On the other hand, when the target variable can take on only a small number of discrete values (such as if, given the living area, we wanted to predict if a dwelling is a house or an apartment), we call it a classification problem.



Figure 3. Different stages of the supervised learning [11]

1.1.2. Unsupervised Learning

Unsupervised learning involves learning from unlabeled data, where the algorithm uncovers patterns or structures in the data without any specific target output. It is used to discover hidden relationships or groupings within the data [12]. Clustering and non-clustering algorithms (Figure 4) is an example of unsupervised learning. Example of clustering problem can be seen in market segmentation which is the process of dividing a larger market into distinct, homogeneous groups of consumers who share similar characteristics, needs, or behaviors. Also, an example of non-clustering algorithm is the "Cocktail Party Algorithm". It allows you to find structure in a chaotic environment. For instance, identifying individual voices and music from a mesh of sounds at a cocktail party).

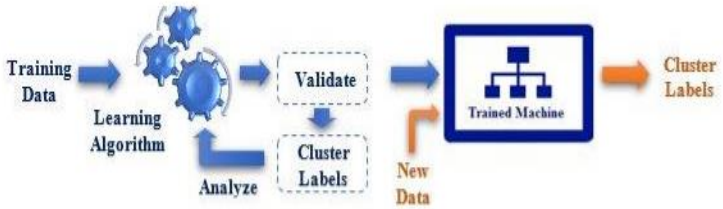


Figure 4. The different stages of the unsupervised learning [11]

1.1.3. Semi-Supervised Learning

Semi-supervised learning is a combination of supervised and unsupervised learning [13]. It leverages a small amount of labeled data along with a larger amount of unlabeled data to improve the learning process as in Figure 5.

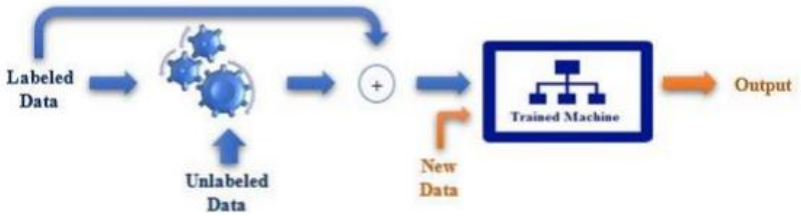


Figure 5. Different stages of the semi-supervised learning [11]

1.1.4. Reinforcement Learning

Reinforcement learning involves training an agent to interact with an environment and learn optimal actions through trial and error. The agent receives feedback in the form of rewards or penalties based on its actions, and it learns to maximize the cumulative reward over time (Figure 6). Reinforcement learning has been successfully applied in areas like robotics and game playing.

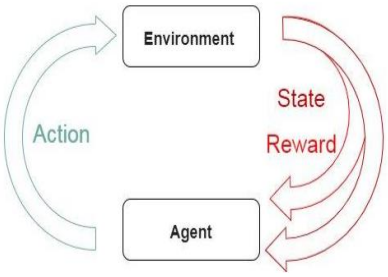


Figure 6. Reinforcement Learning [8]

1.1.5. Deep Learning

Deep learning is a subfield of machine learning that focuses on training deep neural networks with multiple layers to learn complex features and representations from data [14]. Deep learning has achieved remarkable success in various domains, including image and speech recognition, natural language processing, and autonomous driving. Each type of machine learning algorithm possesses unique strengths and weaknesses, making the selection of the appropriate algorithm crucially dependent on the specific problem and the characteristics of the available data. Supervised learning algorithms excel in tasks were labeled training data is available. Unsupervised learning algorithms, on the other hand, are useful in uncovering hidden patterns and structures within unlabeled data. Semi-supervised learning strikes a balance between the two, using a combination of labeled and unlabeled data. Reinforcement learning is ideal for training agents to make optimal decisions in dynamic environments, while Deep learning algorithms, with their ability to learn complex representations, have exhibited exceptional performance in tasks involving large-scale data and intricate patterns (Figure 7).



Figure 7. Deep Learning [15]

1.2. Application of machine learning in predictive maintenance

In the context of maintenance, machine learning offers significant potential for improving efficiency, reducing costs, and enhancing equipment reliability. By analyzing historical data, machine learning algorithms can identify patterns, correlations, and anomalies that may indicate impending failures or maintenance needs as shown in Figure 8.

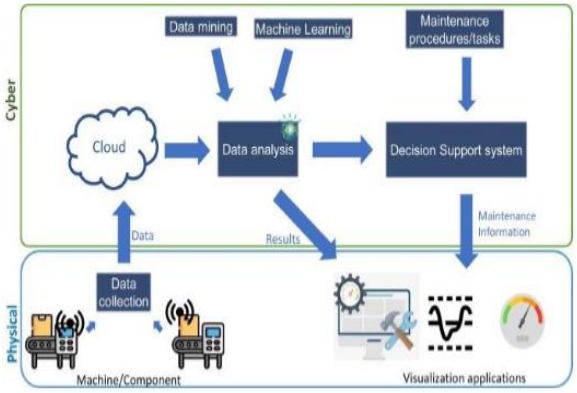


Figure 8. Machine learning in predictive maintenance [16]

This allows maintenance teams to move from traditional, calendar-based maintenance schedules to more proactive and targeted approaches. Machine learning can be applied to various maintenance tasks, such as predictive maintenance, anomaly detection, fault diagnosis, and optimization of maintenance plans and resource allocation. The ability to leverage machine learning technologies enables organizations to harness the power of data to make informed decisions, optimize maintenance strategies, and maximize equipment uptime. [17] states that machine learning plays a crucial role in predictive maintenance across various industries. It is a subfield of artificial intelligence that focuses on developing algorithms and models capable of learning from data and making predictions or taking actions without explicit programming. By analyzing machine data and detecting anomalies, machine learning techniques enable the forecasting of maintenance requirements, optimizing operations, and ensuring safety [9], [18]. In the maintenance sector, artificial intelligence algorithms are extensively applied to enhance enhance maintenance systems, utilizing techniques such as Multilayer Perceptron, Random Forest, and Support Vector Regression. Moreover, the integration of Industrial Internet of Things (IIoT) with machine learning tools provides valuable insights for predictive maintenance, as seen in Industry 4.0 initiatives[19] [20]. Machine learning methods, such as random forest regression and linear regression, have shown high prediction accuracy in anticipating component failures.

However, many studies adopt complex machine learning approaches without first establishing statistically grounded analytical models that clearly explain the contribution of individual process parameters to downtime.

For effective industrial application, there remains a need for simple, interpretable, and statistically reliable models that can be easily implemented by plant engineers as decision-support tools. Therefore, this study develops a multiple linear regression-based predictive framework using real production datasets from Cup and Plate manufacturing lines to model the relationship between operational parameters and machine downtime.

The study first conducts descriptive statistical analysis to understand the inherent behavior, variability, and distribution characteristics of the collected production data. Thereafter, regression models are formulated to quantify the influence of uptime, product weight, cycle time, and production quantity on downtime. By establishing a clear mathematical representation of downtime behavior, the model provides insight into how operational adjustments may influence downtime occurrence and duration. This forms the foundation for broader maintenance intelligence, supporting proactive decision-making, improved machine reliability, and enhanced production efficiency.

This work contributes to predictive maintenance literature by demonstrating the relevance of process-driven parameters as reliable predictors of downtime in plastic manufacturing operations, and by developing a structured regression framework that provides clarity, interpretability, and industrial applicability. The framework developed herein forms the basis upon which subsequent studies on model validation, sensitivity analysis, and economic assessment are further extended.

2. Method

Based on the nature of this work, there is exchange of numeric data (Quantitative) between the hardware and the software components of the developed system. Narrative data (Qualitative) are also gathered from the machine operators and technicians on the maintenance history of the equipment under study, both the numeric and the narrative data are utilized in developing the intelligent system for the equipment maintenance prediction. Hence, the research approach adopted for this study is the Mixed Approach [21], [22].

This project presents a data-centric approach, which is all-encompassing to ensure modelling of downtime in plastic manufacturing process, balancing computational effort and accuracy. It is also based on three phases to reduce modelling effort and deal with the high complexity and amount of system components without resulting in extremely detailed models. The first phase is to identify and analyze key process parameters that influence machine performance and downtime in plastic manufacturing, design and develop a system architecture that integrates IoT (Internet of Things) devices, cloud computing, user interface (UI), and other hardware for data collection, the second is to carry out a Descriptive Data Analysis of the collected data, plus data Fitting for model selection using Excel Data Analysis Toolpak. The third is to develop a predictive model using machine learning technique (Jamovi), and deploy it for predicting machine downtime.

2.1. Data Capturing

After an extensive study of the facility maintenance history and geometry, the IoT network which includes the proximity sensor, microprocessor, and other hardware such as the stopwatch and weighing scale were strategically installed in their respective positions for optimal operation (data capturing).

2.2. Descriptive Data Analysis (DDA) of the Collected Data, and Data Fitting for Model Selection Using Excel Data Analysis Toolpak

The study was conducted using real-time production data collected from Nnamdi Azikiwe Production Cluster plastic manufacturing facility operating in Awka, Nigeria. The equipment under study included injection moulding, Pelleting, Crushing and Washing machines, used to manufacture two products: plastic cups and plates. The data was obtained from proximity sensor, weighing scale, stopwatch, programmable logic controller (PLC)-integrated devices and recorded over 100 production runs for each product.

2.3. Variables Considered

Five key production parameters were included in the dataset:

a. Downtime (minutes): Dependent variable, representing machine stoppage time per production cycle.

b. Uptime (minutes): Total operating time excluding stoppages.

c. Product Weight (g): Mass of each finished unit.

d. Cycle Time (seconds): Time required to complete one production cycle.

e. Quantity: Number of units produced per shift.

Descriptive statistics were used to summarize and understand the structure of the dataset. Key measures included; Central tendency (mean, median), Dispersion (standard deviation, range, minimum, and maximum), Distribution shape (skewness and standard error of skewness). This analysis served to identify potential outliers or anomalies, evaluate the spread and symmetry of each variable, verify statistical assumptions (e.g., normality) prior to regression modelling. The descriptive statistics were generated using Jamovi version 2.4, a user-friendly open-source statistical software based on R.

2.4. Data Fitting

The type of regression model that best fit the data to ensure accurate analysis is determined by the use of Excel Data Analysis Toolpak. The input variables are plotted individually against the output variable (for both cup and plate). Different types of regression models are used, and the model with the best R^2 value is selected for the modelling. Specifically, the model depicts the relationship between the process parameters and the output parameter under study. The results are summarized in a table, and the best model selected. The regression models considered are linear, exponential, logarithmic, polynomial and power.

2.5. Development of a Predictive Model Using Machine Learning Technique (Jamovi), and Machine Downtime Prediction Using Python.

Computer Aided Engineering (CAE) tools have become very important in modern engineering practice, they are used to replicate real-world system behavior for detailed description of processes. Jamovi, Pythton, ThinkSpeak and the Arduino IDE are part of the CAE tool.

JAMOVI data analysis software was selected among other data analysis tool because of its prowess in different variety of industrial applications. It provides users with tools to perform statistical analysis, especially regression analysis, after the Excel Data Analysis ToolPak must have been used for model selection.

Python has very strong mathematical calculation ability, it is particularly good at signal processing, image processing. Cloud IoT is a fully managed service that allows you to easily and securely connect, manage, and ingest data from millions of globally dispersed devices. ThingSpeak as an IoT analytics platform, allows to aggregate, visualize and analyses live data streams in the cloud. These features make ThingSpeak fit to serve as the data base for the system. It's ability to receive and export numerous data at same time informed the decision to select it for this purpose in this study. Its function basically will be to receive data from the hardware (sensors), and export's same to PYTHON for downtime prediction – all thanks to internet.

C++ Arduino IDE, when compared to other existing microcontrollers has minimum power consumption and an easily programable interface. It is available in a cheap amount and provides simple interfacing to analog circuits, and enables users to build their kit. Arduino IDE is for the training of the hardware to capture and transmit data in real-time.

The machine learning tools are deployed at different stages in the system. JAMOVI is for the generation/testing of the analysis model. PYTHON performs the final data analysis and downtime prediction. These tools are integrated with the hardware to develop the intelligent system.

Four major process parameters as input variables for downtime prediction in plastic manufacturing [3]. These variables will produce a robust generic monitoring and control model for adaptive control of process parameters. The equipment uptime, product weight, product cycle time and the quantity of products per shift are grouped under these process parameters. Finding an empirical relationship between the parameters can be made possible using Computer Aided Engineering (CAE) such as Jamovi, Python etc.

In this study, the input variables are the equipment uptime, product weight, product cycle time and the quantity of products produced per shift. The model is to predict downtime in real-time using regression analysis. The data are monitored and collected for about a period of six (6) months, but recorded daily, that's 8hrs per day/shift (480mins), Monday to Friday. The comprehensive data table contains readings at this rate, but was later filtered to generate/test an accurate analysis model using Jamovi. However, an algorithm to recognize process parameters, analyze the data and make downtime prediction is developed using Python Language. The analysis model is imported into the environment for prediction. To model and predict machine downtime, multiple linear regression (MLR) was employed. Model fit was evaluated using:

a. R^2 (coefficient of determination): proportion of variance in downtime explained by predictors.

b. p-values and t-statistics: to assess statistical significance of each variable.

c. Standard error of coefficients: to estimate uncertainty in predictions.

The regression analysis was also performed in JAMOVI, utilizing the "Regression" module with default ordinary least squares (OLS) estimation.

3. Results and Discussion

The data are monitored and collected at the Nnamdi Azikiwe Production Cluster, for about a 6months period, covering daily working shift of 8 hours, Monday-Friday. At the firm, three major operations were carried out during plastic manufacturing namely: Washing, Crushing, Mixing, Pouring and Injection molding in which the outputs are Cup and Plate. Machine uptime, product weight, number of defects, quantity per shift, and cycle time, were considered for downtime prediction. Regular periods of inspection and repair of the machines were performed every 24 hours for 30 minutes and it normally affects the shift of 8h. Cup and Plate (8 h) Operating time for the shift is 8 h × 60 min = 480 min. The following activities constitute planned downtime:

a. Movement of raw materials from where it is stored to where mixing take place.

b. Movement of mixed materials from mixing machine to the hopper

c. Regular maintenance (pit-stop) = 30 min.

d. Mixing of raw materials

e. Pouring of mixed materials into the hopper

f. Break time = 60 min.

The data is filtered to generate an accurate predictive model that perfectly represents the real-world behavior of the firm's manufacturing process.

3.1. Descriptive Data Analysis (DDA) of Data, and Model Selection Using Excel Data Analysis Toolpak.

3.1.1. Descriptive data analysis for cup

The DDA was done using Jamovi, the results are presented in Table 1 and charts below, followed by the analysis.

Table 1. Result of DDA for Cup

Descriptives					
	Downtime (m)	Uptime (m)	Weight (g)	Cycle Time (s)	Quantity/Shift (pcs)
N	100	100	100	100	100
Missing	0	0	0	0	0
Mean	120	363	70.0	20.6	1095
Median	122	362	70.0	21.0	1092
Standard deviation	8.43	9.74	0.674	0.607	29.5
Range	26	30	3	3	92
Minimum	104	350	68	19	1055
Maximum	130	380	71	22	1147
Skewness	-0.429	0.264	-0.190	-0.274	0.257
Std. error skewness	0.241	0.241	0.241	0.241	0.241

As shown in Figure 9, the histogram plots provide a visual representation of the data distribution for downtime, uptime, weight, cycle time, and quantity per shift, enabling a clearer understanding of data spread, skewness, and overall production performance. Next, Figure 10 shows the box plot for the DDA cup.

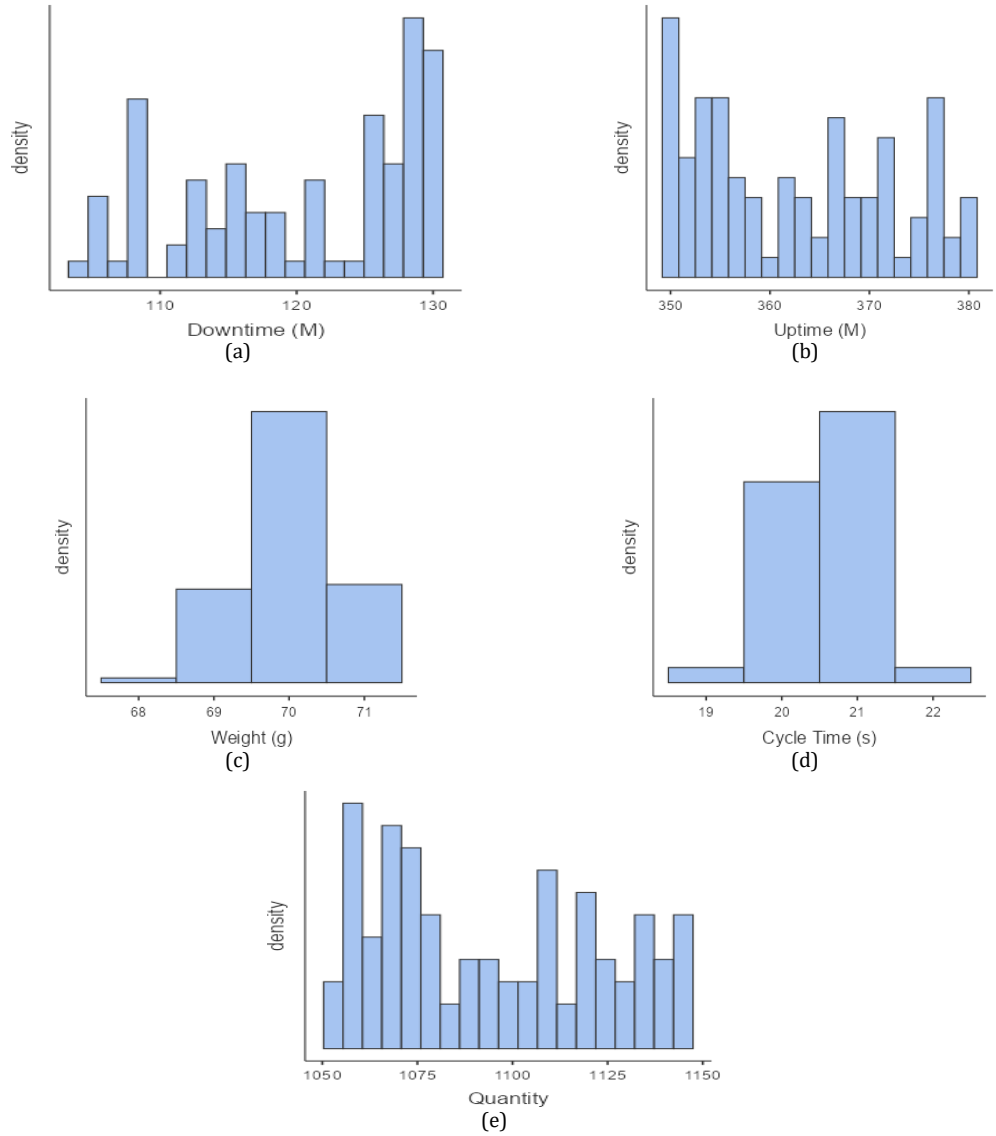
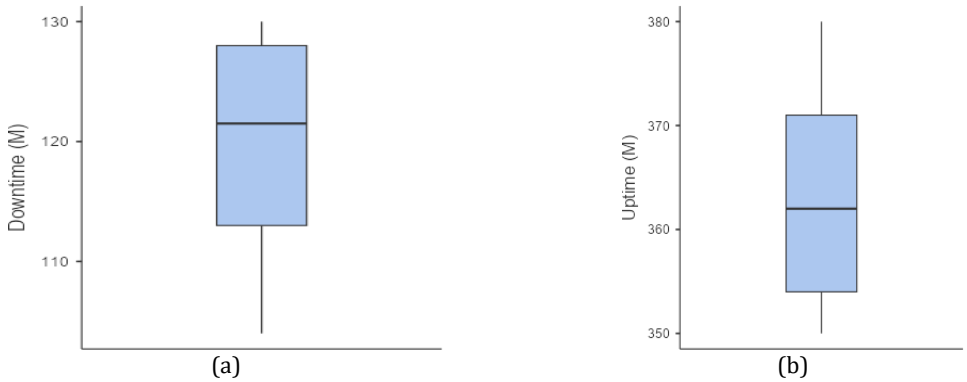


Figure 9. Histogram Plots for Cup DDA



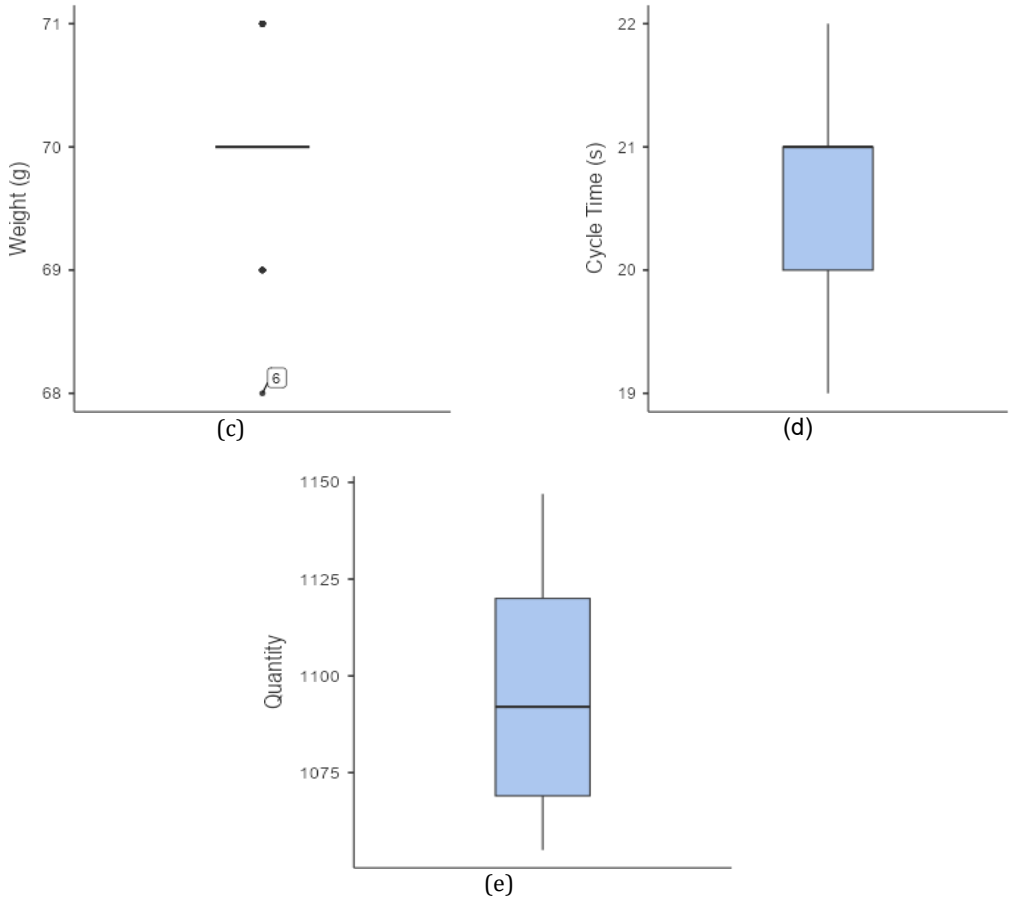


Figure 10. Box Plots for Cup DDA

3.1.2. Interpretation and Discussions

Dataset Overview of Appendices A and B, shows that the Sample size (N) is 100 for each variable, while Missing values = 0, indicating a complete dataset with no missing entries, which strengthens the reliability of subsequent statistical inferences. Central tendency and dispersion for cup can be seen in Table 2.

Table 2. Central Tendency and Dispersion for Cup

Variable	Mean	Median	Std. Dev.	Min	Max	Range
Downtime (m)	120.0	122	8.43	104	130	26
Uptime (m)	363.0	362	9.74	350	380	30
Weight (g)	70.0	70.0	0.674	68	71	3
Cycle Time (s)	20.6	21.0	0.607	19	22	3
Quantity	1095	1092	29.5	1055	1147	92
Uptime (M)	363.0	362	9.74	350	380	30

From Table 2, the average downtime is 120 minutes with a standard deviation of 8.43, suggesting low to moderate variability. The median (122) is slightly above the mean, indicating a left-skewed distribution, confirmed by the skewness value (-0.429). For uptime, Mean and median are close (363 and 362), with a relatively low standard deviation, suggesting a symmetric and consistent uptime process. Weight and Cycle Time have very low standard deviations (0.674 and 0.607 respectively) suggest that weight and cycle time are tightly controlled process parameters, with negligible variation which is critical for manufacturing consistency. Quantity Produced Shows moderate variability with a range of 92 and standard deviation of 29.5, due to variation in shift length, production efficiency, or minor operational disruptions. Shape of the distributions for cup can be seen in Table 3.

Table 3. Shape of the Distributions for Cup (Skewness)

Variable	Skewness	Std. Error of Skewness	Interpretation
Downtime (M)	-0.429	0.241	Slight negative skew
Uptime (M)	0.264	0.241	Slight positive skew
Weight (g)	-0.190	0.241	Approximately symmetric
Cycle Time (s)	-0.274	0.241	Slight negative skew
Quantity	0.257	0.241	Slight positive skew

In Table 3, all variables fall within ±0.5 skewness, indicating approximately normal distributions, a positive sign for applying parametric tests. Slight skewness in downtime and cycle time suggests the need to watch for potential process delays or early terminations that might pull averages down. If the document had accompanying boxplots or histograms, here is what would likely be observed based on the statistics:

From the Boxplots (Figure 10), Downtime has Mild left tail (due to negative skewness), suggesting occasional short downtimes. uptime and quantity have Mild right tails (positive skew), indicating occasional longer uptime or higher output days. The also indicates that weight and cycle time have Very tight interquartile range, likely indicating high process control and machine stability. The histogram plots (Figure 9), shows symmetrical bell-shaped curves for weight and cycle time, and light asymmetry in downtime and quantity with minor outliers at the tails.

Analyzing the operation significance on plastic manufacturing operations and predictive maintenance perspective, the narrow variability in product weight and cycle time suggests well-calibrated machinery and stable production processes. downtime variability, although moderate, indicates room for optimization. Combined with the slight negative skew, this could point to proactive maintenance strategies already being in place. The tight range in uptime implies predictable machine performance, which is an ideal condition for implementing downtime predictive models. For the statistical readiness, DDA indicates that variables are approximately normally distributed, there is no missing data, and the variation is sufficiently present in the independent variables to support model training.

The descriptive analysis of the Cup data reveals a well-behaved dataset in statistical terms with minimal skewness, tight standard deviations in key process parameters, and no missing data. These characteristics are indicative of a controlled production environment, and they support the validity and reliability of the predictive modeling efforts that follow. Such a robust data foundation enhances confidence in the use of machine learning or multiple linear regression models for forecasting downtime or identifying hidden inefficiencies in the plastic manufacturing process.

Similarly, the descriptive statistics for the Plate dataset demonstrate a stable and well-controlled production environment, similar to the Cup line. However, slightly higher variability in downtime and quantity suggests different production dynamics or operator behavior. These findings serve as the basis for regression modelling.

3.2. Model Selection

The type of regression model that best fit the data, to ensure accurate analysis is determined through the use of Excel Data Analysis Toolpak. From the collected data (cup), the input variables are plotted individually against the output variable. Different types of regression models are used, and the model with the best R^2 value is selected for the modelling.

Specifically, the model depicts the relationship between the process parameters (machine uptime, product weight, product cycle time, and product quantity) and the output variable (downtime) under study. The results are summarized in Table 4, and the best model selected. The regression models considered are linear, exponential, logarithmic, polynomial and power. Cup data is used for the model selection.

Table 4. Model Selection Summary Table

S/N	Independent Variables	Dependent Variable	Linear (R^2)	Exponential (R^2)	Logarithmic (R^2)	Polynomial (R^2)	Power (R^2)
1	Uptime	Downtime	1	0.8999	0.7986	0.9121	0.8975
2	Weight	Downtime	0.2521	0.0520	0.0576	0.1541	0.0575
3	Cycle Time	Downtime	0.2283	0.0281	0.3330	0.1678	0.0331
4	Quantity	Downtime	1	0.8999	0.7986	0.8342	0.7975

The model with R^2 value of 1 or closest is the best fit, thus linear model was selected for the modelling.

3.3. Development of a Predictive Model Using Machine Learning Technique (Jamovi), and Machine Downtime Using Python.

The model is created using Jamovi. The parameter data in Appendix 1 is used to develop a multiple linear regression model for analysis, downtime (Y) is the output variable, while equipment uptime (X_1), product weight (X_2), product cycle (X_3), and product quantity per shift (X_4), are the input variables. The JAMOVI results are shown below.

Regression Analysis: Y versus X_1, X_2, X_3, X_4

The regression Equations 1 and Equation 2;

$$Y_{Cup} = 401.200 - 0.417X_1 + 0.292X_2 + 0.428X_3 - 0.145X_4 \quad (1)$$

$$Y_{Plate} = 498.5410 - 0.9751X_1 - 0.1016X_2 - 0.0411X_3 - 0.0140X_4 \quad (2)$$

Where Y is Downtime (m), X_1 = Uptime (m), X_2 = Weight (g), X_3 = Cycle time (s), X_4 = Quantity/shift

Table 5 shows the fit measure model, Table 6 shows the cup model coefficients - downtime (M), Table 7 shows the plate model coefficients - downtime (M), and Figure 11 shows the residual plots for cup, and Figure 12 shows the residual plots for plate

Table 5. Model Fit Measures

Model	R	R^2
1 (Cup)	0.988	0.976
2 (Plate)	0.993	0.987

Note. Models estimated using sample size of N=100

Table 6. Cup Model Coefficients - Downtime (M)

Predictor	Estimate	Standard Error	T-Value	P-Value
Intercept	401.200	15.5602	25.78	<.001
Uptime (M)	-0.417	0.2453	-1.70	0.092
Weight (g)	0.292	0.1991	1.46	0.146
Cycle Time (s)	0.428	0.2192	1.95	0.054
Quantity	-0.145	0.0811	-1.79	0.077

Table 7. Plate Model Coefficients - Downtime (M)

Predictor	Estimate	Standard Error	T-Value	P-Value
Intercept	498.5410	15.4308	32.308	<.001
Uptime (M)	-0.9751	0.1735	-5.619	<.001
Weight (g)	-0.1016	0.1752	-0.580	0.563
Cycle Time (s)	-0.0411	0.2005	-0.205	0.838
Quantity	-0.0140	0.0694	-0.201	0.841

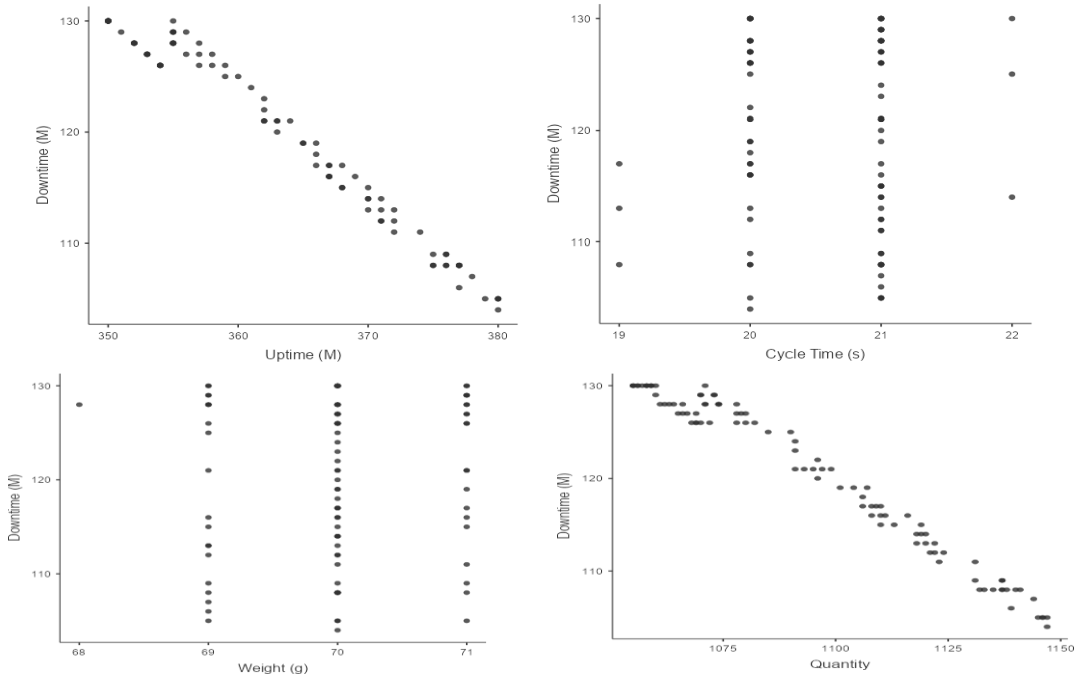


Figure 11. Residual Plots for Cup

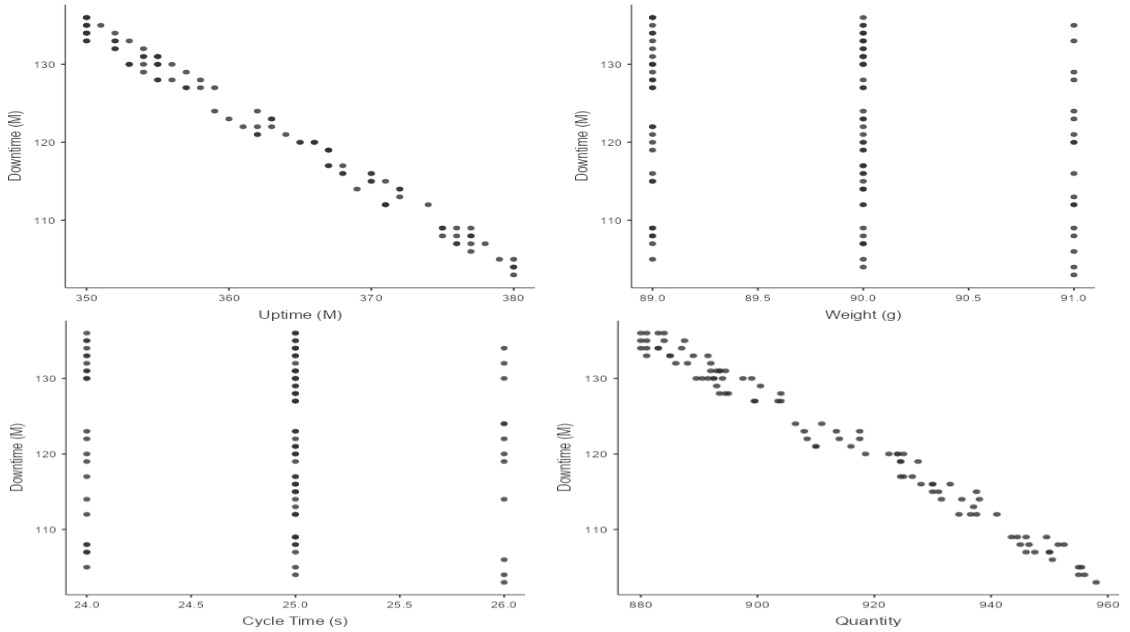


Figure 12. Residual Plots for Plate

3.3.1. Cup Linear regression interpretation and analysis

R^2 = 0.976 (Table 6); this means that 97.6% of the variance in downtime is explained by the four predictor variables (Uptime, Weight, Cycle Time, Quantity). This is a very high explanatory power, suggesting that the model fits the data exceptionally well. R = 0.988; Strong positive correlation between observed and predicted values of downtime. With N = 100, the sample size is adequate to yield statistically stable and reliable results. This high R^2 suggests that the model can be confidently used in predictive maintenance and process optimization applications.

As uptime increases, downtime is expected to decrease (Figure 11), holding other variables constant, so for every 1-minute increase in uptime, there is a reduction of 0.417 minutes in downtime (Table 8). Marginal (p = 0.092), indicating that the relationship, while logically sound, may require further data or control of confounding factors to be statistically robust.

Higher product weight correlates with increased downtime; this could imply that heavier products (perhaps due to material inconsistencies or machine strain) require longer changeover or maintenance time. The model suggests that longer cycle times tend to be associated with greater downtime. Every 1-second increase in cycle time contributes to an approximate 0.43-minute increase in downtime. Therefore, machines taking longer to complete a cycle may be closer to fault thresholds or require recalibration. Higher output (Quantity), is associated with reduced downtime, likely because efficient runs produce more units without stopping the line.

Overall Model Implications for cup, reveals that downtime is inversely related to production uptime and output quantity, while being positively linked to heavier weights and longer cycle times. This suggests a potential operational cycle as machines work longer and produce more, downtime tends to reduce, but when weights increase or cycles lengthen, it may indicate wear and tear, impending failure, or inefficiency thus increasing downtime risk. This regression model (Equation 1), is both statistically and operationally meaningful. It provides a predictive mechanism for estimating downtime based on measurable production parameters. While not all predictors achieved traditional statistical significance, their directionality aligns with logical manufacturing principles. It offers a robust base for developing real-time monitoring systems, threshold alert mechanisms, and downtime minimization strategies in plastic manufacturing.

3.3.2. Plate Linear regression interpretation and analysis

R^2 = 0.987 (Table 6); this indicates that 98.7% of the variability in downtime is explained by the four predictors, demonstrating an excellent model fit. R = 0.993; a very strong linear correlation between predicted and observed downtime values. The high R^2 also suggests that the model is robust and highly predictive within the sample used. Similarly, for every 1-minute increase in uptime, downtime decreases by approximately 0.98 minutes (Table 8), this further suggests that consistent machine operation without interruption drastically reduces total downtime. Quantity also has effect size but less statistically significant.

The regression model for Plate production (Equation 2), reveals that Uptime is the only statistically significant factor influencing downtime (Figure 12). The model exhibits a very high R^2 value, suggesting strong predictive capability. However, Weight, Cycle Time, and Quantity, though theoretically relevant, do not significantly explain downtime variations for Plate under current operational conditions. These findings underscore the dominant influence of machine runtime stability in reducing production l1/27/2026osses.

3.4. Machine Downtime Using Python

Figure 8 presents the results of machine downtime predictions for cup and plate components in the range of 1 to 22 days using Python-based modeling.

Table 8. Predicted Downtime for Cup and Plate

Number of Days (n)	Cup Predicted Downtime (min)	Plate Predicted Downtime (min)
1.0	119.39	123.09
2.0	117.06	129.35
3.0	110.5	116.16
4.0	117.94	110.28
5.0	123.71	125.58
6.0	118.3	116.52
7.0	118.68	120.81
8.0	121.49	117.31
9.0	120.77	131.8
10.0	121.63	125.48
11.0	117.16	130.22
12.0	121.18	126.58
13.0	114.54	117.96
14.0	107.87	106.18
15.0	125.16	131.88
16.0	122.05	123.0
17.0	116.83	125.73
18.0	128.59	132.29
19.0	122.09	133.7
20.0	119.71	112.21
21.0	119.71	109.44
22.0	119.2	125.03

(n = 22 observations; values in minutes)

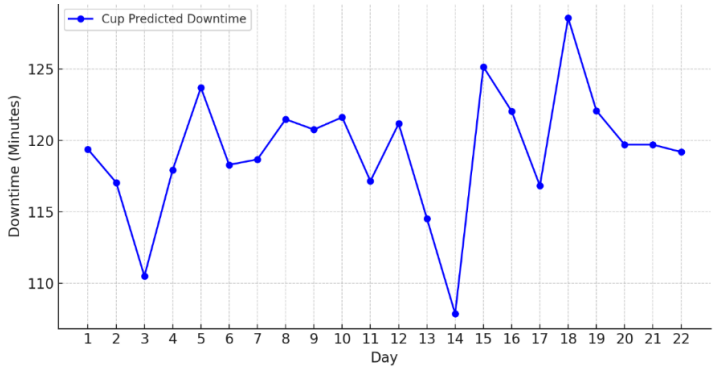


Figure 13. Cup Monthly Downtime Prediction

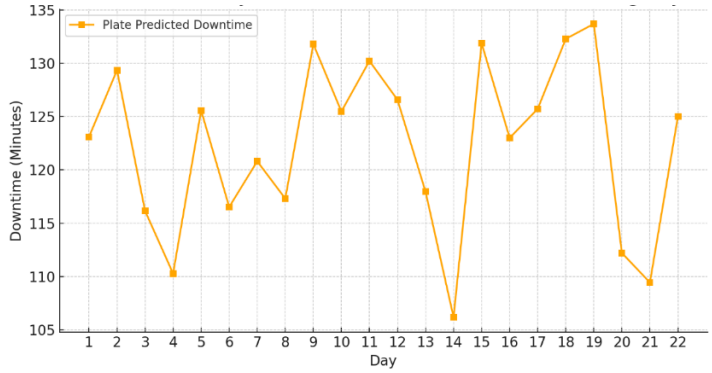


Figure 14. Plate Monthly Downtime Prediction

In Figure 13 and Figure 14 plotted from Table 8, the simulated monthly downtime predictions for Cup and Plate production lines over a 22-working-day period were presented, based solely on the regression models developed in equation 1 and equation 2. The simulation reflects the pure predicted values derived directly from the regression models.

These predictions serve as a baseline for proactive maintenance planning and performance evaluation. The consistent fluctuation in values across the 22 days reflects the inherent dynamics of manufacturing operations, where factors such as cycle time, machine uptime, and production weight contribute to downtime variability. The insights derived from these simulated values can be further integrated into predictive maintenance systems, offering manufacturers a data-driven foundation for decision-making.

4. Conclusion

This study presented a data-driven regression framework for predicting machine downtime using process parameters obtained from real plastic manufacturing operations. Production datasets from Cup and Plate lines were utilized, with uptime, product weight, cycle time, and production quantity integrated as predictor variables. Descriptive statistical analysis provided insight into the stability and behavior of the datasets, while multiple linear regression modeling successfully established the quantitative relationship between these operational parameters and machine downtime.

The developed models demonstrated strong explanatory capability, confirming that downtime can be effectively predicted from measurable production variables. The study thereby reinforces the role of data-driven techniques in transitioning from reactive to predictive maintenance strategies. By providing a transparent and interpretable modeling structure, the framework supports engineers and decision makers in understanding downtime behavior and anticipating machine disruptions before they occur.

Beyond establishing this predictive foundation, the study sets the stage for further research extensions, including advanced model validation, one-at-a-time sensitivity assessment of model parameters, incorporation of real-time monitoring, and integration of economic cost analysis to evaluate maintenance decisions. Overall, the proposed regression framework contributes to the advancement of predictive maintenance intelligence, promotes improved machine reliability, and supports efficient production system management in manufacturing environments.

Author Contributions

All authors have equal contributions to the paper. All the authors have read and approved the final manuscript.

Funding

No funding support was received.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/ or publication of this article.

Data Availability

The datasets generated during and/ or analysed during the current study are available from the corresponding author on reasonable request.

Declaration on AI Use

The authors declare that no artificial intelligence (AI) or AI-assisted tools were used in the preparation of this manuscript. AI were used only to improve readability and language under strict human oversight; no content, ideas, analyses, or conclusions were generated by AI.

References

[1] M. Bevilacqua, F. E. Ciarapica, and G. Mazzuto, “A Fuzzy Cognitive Maps Tool for Developing a RBI&M Model,” *Qual. Reliab. Eng. Int.*, vol. 32, no. 2, 2014, doi: 10.1002/qre.1756.

[2] D. O. Onuoha, C. E. Mgbemena, H. C. Godwin, and F. N. Okeagu, “APPLICATION OF INDUSTRY 4 . 0 TECHNOLOGIES FOR EFFECTIVE REMOTE MONITORING OF CATHODIC PROTECTION SYSTEM OF OIL AND GAS PIPELINES -A SYSTEMATIC REVIEW To cite this version : MONITORING OF CATHODIC PROTECTION SYSTEM OF OIL AND GAS PIPELINES -A HAL Id : hal,” *Int. J. Ind. Prod. Eng.*, vol. 1, no. 2, pp. 29–50, 2022, doi: https://journals.unizik.edu.ng/index.php/ijipe/.

[3] S. C. Nwanya, J. I. Udofia, and O. O. Ajayi, “Optimization of machine downtime in the plastic manufacturing

Optimization of machine downtime in the plastic manufacturing," *Cogent Eng.*, vol. 20, no. 1, 2017, doi: 10.1080/23311916.2017.1335444.

[4] S. Shin, K. Baek, J. Oh, Y. Bin, and M. Dok, "Sensors and Actuators : A . Physical Recent developments of in-situ process and in-line quality monitoring in injection molding using intelligent sensors," *Sensors Actuators A. Phys.*, vol. 383, no. November 2024, p. 116248, 2025, doi: 10.1016/j.sna.2025.116248.

[5] Z. A. Elnaby, A. Zaher, Ragab K. Abdel-Magied, and H. I. Elkhouly, "Improving plastic manufacturing processes with the integration of Six Sigma and machine learning techniques : a case study," *J. Ind. Prod. Eng.*, vol. 41, no. 1, pp. 1–18, 2024, doi: https://doi.org/10.1080/21681015.2023.2260384.

[6] K. Ashhar and Y. F. C. Shan, "Data-driven calibration of temperature sensors for monitoring and control of injection molding machine," *Meas. Sensors*, vol. 38, 2025, doi: 10.1016/j.measen.2024.101631.

[7] E. Eyere, A. I. Olabisi, and O. P. Aruoture, "DESIGN AND DEVELOPMENT OF A LOW-COST WASHING MACHINE SUITABLE FOR POLYTHENE MATERIALS," *Asian J. Curr. Res.*, vol. 2, no. 1, pp. 22–32, 2017.

[8] F. Fourati and M. Alouini, "Artificial intelligence for satellite communication : A review," *Intell. Converg. Networks*, vol. 2, no. 3, p. 213–244, 2021, doi: 10.23919/ICN.2021.0015.

[9] E. Quatrini, F. Costantino, G. Di Gravio, and R. Patriarca, "Condition-Based Maintenance — An Extensive Literature Review," *Machines*, vol. 8, no. 31, pp. 1–28, 2020.

[10] V. Nasteski, "An overview of the supervised machine learning methods," *Horizons.*, pp. 1–11, 2017, doi: 10.20544/HORIZONS.B.04.1.17.P05.

[11] A. B. NASSIF, I. SHAHIN, I. ATTILI, M. AZZEH, and K. SHAALAN, "Speech Recognition Using Deep Neural Networks : A Systematic Review," *IEEE access,* vol. 7, pp. 19143–19165, 2019, doi: 10.1109/ACCESS.2019.2896880.

[12] V. Patel, "A framework for secure and decentralized sharing of medical imaging data via blockchain consensus," *Health Informatics J.*, vol. 25, no. 4, pp. 1398 –1411, 2019, doi: 10.1177/1460458218769699.

[13] Y. C. A. P. Reddy, P. Viswanath, and B. E. Reddy, "Semi - supervised learning : a brief review," *Int. J. Eng. Technol.*, vol. 7, pp. 81–85, 2018.

[14] N. K. Chauhan and K. Singh, "A Review on Conventional Machine Learning vs Deep Learning," *Int. Conf. Comput. Power Commun. Technol.*, pp. 347–352, 2018.

[15] F. He, P. K. Olia, R. J. Oskouei, M. Hosseini, Z. Peng, and T. Banirostam, "Applications of Deep Learning Techniques for Pedestrian Detection in Smart Environments : A Comprehensive Study," *Hindawi J. Adv. Transp.*, vol. 2021, no. 1, pp. 1–11, 2021, doi: 10.1155/2021/5549111.

[16] J. Dalzochio *et al.*, "Computers in Industry Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges," *Comput. Ind.*, vol. 123, p. 103298, 2020, doi: 10.1016/j.compind.2020.103298.

[17] Z. M. Çınar, A. A. Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, "Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0," *sustainability*, vol. 12, no. 19, 2020, doi: https://doi.org/10.3390/su12198211.

[18] F. N. Okeagu and C. E. Mgbemena, "A SYSTEMATIC REVIEW OF DIGITAL TWIN SYSTEMS FOR IMPROVED PREDICTIVE MAINTENANCE OF EQUIPMENT IN SMART FACTORIES," *Int. J. Ind. Prod. Eng.*, vol. 1, no. 1, pp. 1–20, 2022, doi: https://journals.unizik.edu.ng/index.php/ijipe/.

[19] A. Angelopoulos *et al.*, "Tackling Faults in the Industry 4.0 Era-A Survey of Machine-Learning Solutions and Key Aspects," *Sensors*, vol. 20, pp. 1–34, 2020, doi: 10.3390/s20010109.

[20] C. E. Mgbemena and F. N. Okeagu, "Development of an IoT-based real-time remote monitoring device for the maintenance of injection moulding machines in plastic industries," *UNIZIK J. Eng. Appl. Sci.*, vol. 2, no. 1, pp. 260–278, 2023, doi: https://journals.unizik.edu.ng/index.php/ujeas.

[21] K. C. Onyeneho, M. C. Uzochukwukanma, and F. N. OKEAGU, "Optimization of Brewing Extract for Optimum Yield of Beer from Industrial Beet Molasses," *ICONIC Res. Eng. JOURNALS*, vol. 9, no. 2, pp. 846–861, 2025.

[22] O. C. Chukwunedum, F. N. Okeagu, C. S. Anyaora, and D. C. Anizoba, "Increasing method performance of long-dimensional hydraulic cylinders for road and construction machinery," *Greenation Int. J. Eng. Sci.*, vol. 3, no. 1, 2025, doi: https://doi.org/10.38035/gijes.v3i1.