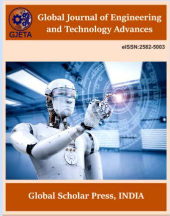
Global Journal of Engineering and Technology Advances

eISSN: 2582-5003

Cross Ref DOI: 10.30574/gjeta

Journal homepage: https://gjeta.com/



(RESEARCH ARTICLE)

Check for updates

Comparative study of sensor-generated operational data, equational calculation, and temporal data acquisition for optimal predictive maintenance decisions

Ahiamadu Jonathan Okirie [1,*], Mack Barnabas [2] and Nnorom Obinichi [1]

[1] *Department of Mechanical Engineering, University of Port Harcourt, Rivers State, Nigeria.*
[2] *Department of Industrial/Technical Education, University of Nigeria, Nsukka, Nigeria.*

Global Journal of Engineering and Technology Advances, 2024, 20(02), 061–073

Publication history: Received on 23 June 2024; revised on 02 August 2024; accepted on 05 August 2024

Article DOI: https://doi.org/10.30574/gjeta.2024.20.2.0143

Abstract

Predictive maintenance greatly enhances equipment health monitoring and performance, this strategy predicts failures before they occur, allowing for focused and timely interventions; it does this by combining real-time sensor data trending, machine learning, and sophisticated data analytics. This study compares equational calculations with sensor-generated data, it further investigates the stability, accuracy, and reliability of electric motor RPM data collected at different intervals to find the most proficient data acquisition techniques for effective predictive maintenance decisions. Statistical metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Absolute Error (AE) were utilized to quantify variances, the assessment approach consisted of two distinct phases: an initial accuracy assessment to measure discrepancies between sensor-generated data and calculated value; this was followed by evaluating the stability, accuracy, and reliability of RPM data collected at short intervals and those collected at longer intervals. Key findings indicate that sensor-generated RPM readings at short intervals provide detailed insights into electric motor transient behaviours despite greater variability and error margins. This high-frequency data provides a detailed insight into motor function by capturing extensive deviation patterns that long-term data trending could not. Also, sensor-generated data shows a substantial ability to give precise insights into transitory behaviors compared to equationally computed value, making it an appropriate option for predictive maintenance. The study advances the subject of predictive maintenance by providing useful recommendations for improving maintenance procedures, increasing equipment reliability, and reducing downtime and costs.

Keywords: Comparative analysis; Predictive maintenance; Reliability; Electric motor; Revolution per minute (RPM); Equational calculations; Sensor-generated data; Variability

1. Introduction

With the advent of the Internet of Things (IoT) and predictive technologies, there has been tremendous growth in the amount of data generated [1]. The advent of big data and attendant machine learning renaissance offers opportunities for and challenges to data quality research [2]. In the realm of predictive maintenance, accurate and reliable data is crucial for predicting equipment failures and optimizing performance. Machine-generated data, derived from advanced sensors monitoring and recording systems offers real-time insights into equipment conditions. However, the accuracy and reliability of this data can be influenced by various factors such as sensor calibration, data processing algorithms, and environmental conditions. Relational databases with a high degree of quality may be the gateway for predictive modelling and enhanced business analytics [3], this data is invaluable for its granularity and immediate relevance, potentially enabling more precise predictions about equipment health. In contrast, equational calculations rely on mathematical models and theoretical frameworks to predict equipment behaviour, these models are based on established principles and historical data, providing a different approach to predictive maintenance. While these

* Corresponding author: Ahiamadu Jonathan Okirie.

calculations can offer a stable and theoretically sound basis for predictions, their effectiveness can be limited by the assumptions and simplifications inherent in the models. The study aims to evaluate the comparative performance of these two approaches by analyzing data accuracy, quality, and reliability across different temporal intervals. Short intervals, characterized by high-frequency data collection capture detailed and transient behaviors of equipment. Long intervals, on the other hand, provide a more seemingly stable and aggregated view of performance trends over time. By assessing both machine-generated data and equational calculations within the framework of short and long temporal intervals, this research seeks to identify the strengths and limitations of each method. The findings will contribute to the understanding of how different data acquisition strategies affect predictive maintenance, offering valuable insights for optimizing maintenance practices and decisions, and enhancing equipment reliability. This investigation not only addresses the technical aspects of data accuracy and reliability but also aims to provide practical recommendations for implementing effective predictive maintenance strategies. Through rigorous comparative analysis, the study aspires to advance the field of predictive maintenance and improve the overall efficiency and lifespan of industrial equipment.

This study examines how operational data from sensors has been integrated into industrial settings by analyzing earlier research on the subject. In an attempt to close a research gap, it examines the advantages and disadvantages of utilizing sensor data to improve predictive maintenance techniques in industrial settings. Utilizing data can significantly boost manufacturing efficiency [4][5], companies are increasingly taking advantage of data analytics to make informed decisions that enhance production and profitability [6] [7] [8] [9], data science facilitates this by enabling large-scale, data-driven decision-making and automation, supported by technologies for big data storage and processing [10]. However, maintaining data security and integrity is crucial, requiring robust frameworks to avoid risks such as financial losses and legal issues [11], according to [12], continuous condition monitoring, which involves frequent sensor measurements is advantageous for early fault detection in critical equipment and is further improved by advanced data processing techniques. [13], further adduced that the advent of Industry 4.0 introduced machine learning-based solutions for predictive maintenance, utilizing historical data to predict and prevent failures. However, extracting valuable insights from sensor data remains challenging and is often used reactively rather than proactively [14], because, enhancing data quality is essential for effective data analytics and decision-making [15] [16]. In its study, [17] introduced a six-step approach for improving data quality, which is crucial for precise analysis and effective monitoring in performance systems. The accuracy of online performance monitoring systems heavily depends on the quality of the input data, which is often compromised due to the low precision of installed plant sensors, automated predictive maintenance is closely linked with advancements in production automation and intelligent sensor technologies[18], necessitating efficient data analysis to support complex decision-making and system management.

2. Materials and Methods

The study's methodology involves defining and establishing key parameters and variables and a comprehensive review of sensor-generated operational data and equational calculations to evaluate data accuracy, reliability, and quality. RPM data from an electric motor was collected using condition monitoring, a human-machine interface, and a recorder over 60 seconds, with records taken at 10-second and 1-second intervals. Statistical error metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Absolute Error (AE) were used to assess the differences between sensor-generated and equationally computed data. The data analysis process included two stages: initially, an accuracy assessment using these metrics, quantified the discrepancies between the sensor-generated and equationally calculated data. Following this, a reliability and qualitative assessment evaluated the consistency of sensor data over time (10-second and 1-second intervals) and under the same operational conditions, ensuring its robustness and predictability to determine the best methods of data collection for predictive maintenance. The results of the comparative analysis of these scenarios were visualized using graphs and error analysis was conducted to identify discrepancies, leading to a detailed investigation into potential causes of deviations.

2.1. Sensor-generated operational data

Operational data generated by sensors is the data produced by machinery and equipment during operation. Examples of this type of data include vibration, temperature, pressure, speed, RPM etc. The gathering and examination of this data provide many advantages. Continuous data collection allows for equipment performance to be monitored in real-time, which enables the prompt identification of irregularities and possible malfunctions. By examining past data and recognizing trends, companies can adopt predictive maintenance strategies to predict equipment breakdowns in advance, ultimately decreasing downtime and maintenance expenses. In-depth understanding of how equipment functions helps improve operations by making them more efficient and productive. Moreover, precise and thorough information aids in making well-informed decisions at every stage, from day-to-day activities to long-term strategy. Data generated by machines also guarantees accurate and timely reporting for regulatory compliance and internal audits. In general, utilizing this information improves the reliability, productivity, and security of industrial processes.

2.2. Equational calculations

Equational computations use mathematical equations and formulas to determine precise values and outcomes using given parameters and variables. These computations play a fundamental role in multiple scientific, engineering, and technical fields by providing representations, assessments, and forecasts of behaviors and results. Equational calculations in predictive maintenance can aid in predicting equipment's remaining useful life, failure rates, and optimizing maintenance schedules. The precision of equational computations relies on the correctness of the data provided and the suitability of the selected models and equations. Despite being accurate and based on theory, equational calculations may not always accurately represent intricate real-world dynamics like data-driven methods do. However, they continue to be a vital resource for engineers and analysts, providing a systematic method for problem-solving and decision-making across different industries.

2.3. Data accuracy, reliability, and quality

In the field of data management and analytics, data quality, accuracy, and reliability are key elements. The term "data accuracy" describes the correctness of data, making sure that it accurately depicts real-world values or situations. Precise analysis and well-informed decision-making depend on accurate data. On the other hand, data credibility and uniformity throughout time are related to data reliability. It is essential to observe trends and make prediction because consistent data yields the same outcomes under consistent conditions. Accuracy and reliability are only two aspects of the larger idea of data quality, which also includes timeliness, completeness, and relevance. Good data is complete, relevant to the situation, and readily available. When combined, these qualities guarantee that data is reliable and useful for a range of analytical objectives, eventually resulting in improved operational efficiency and decision-making.

2.4. Mean absolute error (MAE)

Mean Absolute Error is commonly used in statistical analysis and machine learning to assess the precision of predictive models. MAE calculates the mean error size between predicted and actual values, regardless of their direction. MAE is calculated by taking the average of the absolute differences between sensor-generated data (y_i) and equational calculations (y), where n represents the total number of observations. The equation for MAE is:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - y) \ldots\ldots\ldots\ldots (1)$$

In contrast to other error measurements, MAE calculates errors without squaring them, resulting in a linear score that gives equal importance to all errors. This feature of MAE makes it easy to understand, as it is in the same units as the target and shows the average error size clearly. MAE stands out because it is more robust when dealing with outliers compared to metrics such as MSE and RMSE which are more sensitive to outliers. Nevertheless, it does not reprove significant errors as harshly as MSE or RMSE which could be a disadvantage in situations where large errors are unwanted. MAE is commonly used in regression analysis and model evaluation to measure predictive precision in real-world applications. A model is considered more accurate when it has a lower MAE, as this indicates less average errors in the predictions. MAE is a useful tool for professionals in different fields who want to assess and enhance their predictive models' performance due to its straightforwardness and interpretability.

2.5. Mean squared error (MSE)

Mean Squared Error is a basic measurement utilized in statistical analysis and machine learning to evaluate the mean of the errors' squares, precisely the mean squared discrepancy between the predicted values and the genuine values. It acts as an essential signal of the precision of a forecasting model. The MSE formula is as follows:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - y)^2 \ldots\ldots\ldots\ldots (2)$$

A smaller MSE implies a stronger match of the model to the data, showing that the estimated values are similar to the real values, whereas a larger MSE indicates weak model performance, with notable differences between the estimated and real values. MSE is commonly utilized in regression analysis to assess the effectiveness of regression models, aiding in understanding the extent to which the model accounts for variability in the data. In the field of machine learning, MSE is frequently utilized as a loss function for optimizing model parameters to minimize MSE in order to enhance prediction accuracy. MSE's sensitivity to large errors is a key benefit because it reproves larger errors more than smaller ones, making it useful in scenarios where significant deviations are unwanted. Yet, this high sensitivity to significant errors could also pose a drawback since squaring the errors may overly reprove extreme values, possibly resulting in an inflated estimation of the model's error when outliers are included in the dataset. However, despite this, MSE remains a versatile and widely used tool that offers a precise measurement of predictive model accuracy, crucial for model assessment comparison and enhancement, allowing practitioners to effectively evaluate and enhance the performance of their models.

2.6. Root mean squared error (RMSE)

Root mean squared error is a frequently utilized metric in statistical analysis and machine learning to assess the precision of a predictive model. Insufficient understanding of the theoretical characteristics of the RMSE and MAE estimators has restricted their optimal utilization and sparked discussions on determining the superior error metric [19]. In model evaluation research, both the RMSE and MAE are frequently employed [20]. By computing the square root of the average squared discrepancies between the predicted and actual values, the RMSE metric quantifies prediction errors. RMSE is a metric that quantifies the prediction errors by calculating the square root of the average squared discrepancies between predicted and actual values. The equation for RMSE is:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y\right)^2} \ \ldots\ldots\ldots (3)$$

This metric is beneficial as it is measured in the same units as the target variable, which simplifies its interpretation and comparison with the actual data. A smaller RMSE suggests increased precision and improved model effectiveness since the projected values are nearer to the real values. The metric is very helpful in situations where significant errors can cause serious issues, as it reproves bigger errors more severely than smaller ones. Just like MSE, RMSE can be impacted by outliers, potentially magnifying the model's error in the presence of outliers in the data. In general, RMSE offers a reliable and easy-to-understand way to evaluate prediction accuracy, helping with the analysis, comparison, and enhancement of predictive models. It is a crucial instrument for professionals looking to comprehend and improve the effectiveness of their models.

2.7. Absolute error (AE)

Absolute error is a basic principle in statistics and data analysis that is utilized to assess how close a predicted value is to the true value. It is described as the total discrepancy between the value that was observed and the value that was predicted. AE is equal to the difference between y_i and y. The equation used to determine absolute error is:

$$\text{AE} = y_i - y \ \ldots\ldots\ (4)$$

Absolute error is a simple way to quantify the size of errors in a dataset regardless of their direction, assuming it is always a positive value. This simplicity makes it a straightforward measure to understand the average difference between predictions and actual observations. The key benefit of AE is its resistance to outliers since it does not magnify larger errors as MSE does. This quality of absolute error is especially valuable in cases where outliers could distort the outcomes. Nevertheless, as it does not harshly reprove significant mistakes. AE is a useful method to assess how well predictive models perform, offering a straightforward and understandable way to determine prediction precision. By emphasizing the size of mistakes, analysts and practitioners can gain insight into how accurate and precise their models are by understanding the average difference between predicted and actual values.

2.8. Case study equipment: electric motor overview

Electric motors are crucial in contemporary technology, converting electrical energy into mechanical energy and being utilized in a variety of applications for their effectiveness, reliability, and flexibility. The classification consists of three main types: AC motors, DC motors, and special-purpose motors. AC motors, such as induction and synchronous motors run on alternating current and are utilized in industrial machinery and applications that need accurate speed regulation. DC motors operated by direct current consist of brushed and brushless varieties, recognized for their precise speed regulation and effectiveness, utilized in electric cars, household gadgets, and technology. Types of motors specifically designed for certain functions such as stepper and servo motors are utilized in precise tasks like operating 3D printers, robots, and CNC machines. Electric motors operate by the combination of magnetic fields and electricity, generating torque to rotate the rotor. They play a critical role in different industries, providing energy to industrial machines, HVAC systems, for businesses, home appliances, electric cars, and automation systems used in robotics and manufacturing. Advantages they provide include high efficiency, minimal maintenance, precise control, and environmental friendliness because of zero emissions while operating. Nevertheless, it I challenged by reliable power supplies, upfront expenses, advanced control systems, and environmental issues related to disposing and recycling motor parts. The electric motor's frequency is determined by how many cycles that occurs in one second of the alternating current (AC) power that powers it. Frequency is quantified in units of hertz (Hz). According to [21], the speed at which a motor operates can be influenced by the frequency of the AC power supply. Variable frequency drives for electric motors have the ability to adapt pump performance to different operating conditions by decreasing motor and pump RPM. Typically, the standard frequencies in many countries are 50 Hz and 60 Hz. The frequency of 50 Hz is widespread in various regions across the globe, such as Europe, Asia, and Africa. On the other hand, the 60 Hz frequency is commonly found in North America and other areas. The AC motor's synchronous speed, the rate at which the magnetic field spins, can be determined using the motor's frequency and pole count. The equation used to calculate this velocity is:

$$\text{RPM} = \frac{120\times f}{N} \text{.........} \quad (5)$$

Where N is the number of poles, f is the frequency, and RPM is the synchronization speed. It's essential to remember that slip causes the motor's real speed to be somewhat lower than the synchronous speed. Slip is the difference between the rotor's real speed and the synchronous speed, and it requires torque from the motor. The number of full rotations or cycles that a rotating equipment completes in a minute is measured in RPM. This metric is crucial for evaluating the rotational speed of engines, motors, turbines, and other machinery in a variety of mechanical and engineering contexts. Comprehending RPM is crucial to guaranteeing the appropriate functioning, effectiveness, and security of these systems. The basic formula to calculate RPM is:

$$\text{RPM} = \frac{\text{Number of revolutions}}{\text{Time taken}} \text{........} \ (6)$$

Revolutions per minute is an important metric for monitoring the speed at which engines and equipment rotate in different sectors. Tachometers detect and count rotations to measure RPM by tracking movement over a period. In the automotive field, RPM is used to track engine performance, improve fuel efficiency, power output, and maintain safe engine operations by providing drivers with real-time data. In industrial environments, RPM monitoring helps ensure machines function properly, boosting efficiency and avoiding malfunctions. The aerospace sector depends on RPM readings for jet engines and propellers to uphold ideal thrust and fuel efficiency. Household appliances, like washing machines and blenders indicate RPM to let users know their operating speeds. Various factors impact RPM, such as load - usually lowering RPM - and the frequency of the electrical power source in AC motors. RPM is also influenced by mechanical efficiency, wear, lubrication, and maintenance. Keeping an eye on RPM is essential to avoid overheating, mechanical issues, and accidents, guaranteeing engines and motors operate effectively and have a longer lifespan. Accurate RPM maintenance is crucial in performance-driven activities such as racing and high-precision manufacturing. To sum up, RPM is vital for evaluating how well rotating machinery performs, its efficiency, and safety in different sectors. Precise RPM tracking aids in the consistency of ideal operational levels, enhancing both equipment lifespan and efficiency.

The number of poles in an electric motor plays a crucial role in determining its speed and functionality. The arrangement of the poles, which are the magnetic poles produced by the motor's windings affects the torque and synchronous speed of the machine. Because of their inverse relationship, more poles result in a lower synchronous speed. A 6-pole motor operates at 1200 RPM at 60 Hz, whereas a 2-pole motor operates at 3600 RPM. Reduced pole count motors generate less torque but run faster, which makes them suitable for high-velocity applications like fans and compressors. Conversely, motors with more poles run more slowly but produce more torque, which makes them ideal for applications like controlling cranes and hoists. In addition, motors with more poles tend to operate more smoothly and quietly, making them perfect for environments where noise pollution is a concern. Conversely, motors with fewer poles are often more efficient at higher speeds but may also make more noise. It is necessary to take into account the application's particular needs as well as speed, torque, and efficiency while selecting the right motor. To maximize motor performance and make sure it satisfies operating criteria efficiently and reliably, engineers must carefully evaluate the number of poles in the motor.

Table 1 Case study electric motor parameters and variables.

Motor type	**Frequency, f (Hz)**	**Number of poles (N)**	**RPM $[\frac{120\times f}{N}]$**	**Data transmission method**	**Data collection interval (seconds)**
Induction type	50	4	1500	Sensors	1
					10

3. Results and Discussion

3.1. One-second interval RPM recorded values versus equational RPM

Table 2 One-second interval RPM recorded values

Time(s) i	Sensor-generated RPM (y_i)	Equational RPM (y)	Absolute Error [$y_i - y$]
1	1498	1500	2
2	1502	1500	2
3	1505	1500	5
4	1500	1500	0
5	1497	1500	3
6	1501	1500	1
7	1499	1500	1
8	1500	1500	0
9	1499	1500	1
10	1500	1500	0
11	1486	1500	14
12	1498	1500	2
13	1497	1500	3
14	1525	1500	25
15	1520	1500	20
16	1520	1500	20
17	1497	1500	3
18	1502	1500	2
19	1499	1500	1
20	1505	1500	1
21	1499	1500	1
22	1502	1500	2
23	1498	1500	2
24	1497	1500	3
25	1500	1500	0
26	1499	1500	1
27	1500	1500	0
28	1499	1500	1
29	1505	1500	5
30	1497	1500	3
31	1502	1500	2
32	1504	1500	4
33	1470	1500	30

34	1480	1500	20
35	1481	1500	19
36	1499	1500	1
37	1504	1500	4
38	1497	1500	3
39	1504	1500	4
40	1499	1500	1
41	1505	1500	5
42	1498	1500	2
43	1496	1500	4
44	1504	1500	4
45	1497	1500	3
46	1502	1500	2
47	1496	1500	4
48	1498	1500	2
49	1505	1500	5
50	1504	1500	4
51	1505	1500	5
52	1496	1500	4
53	1498	1500	2
54	1500	1500	0
55	1496	1500	4
56	1504	1500	4
57	1505	1500	5
58	1499	1500	1
59	1500	1500	0
60	1498	1500	2

Table 3 MAE, MSE, and RMSE of one-second interval RPM recorded values

Metric	**Value**
MAE [$\frac{1}{n}\sum_{i=1}^{n}(y_i - y)$]	4.57
MSE [$\frac{1}{n}\sum_{i=1}^{n}(y_i - y)^2$]	91,40
RMSE [$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y)^2}$]	9.56

In Table 2, the sensor-generated RPM values largely cluster around the equational value of 1500 RPM, though notable fluctuations occur. Absolute errors range from 0 to 30 RPM with the most significant deviations observed at specific times such as 14, 15, 16, 33, and 34 seconds, where errors surpass 10 RPM. These deviations are both positive and negative, indicating variability in the direction of deviation from the equational value. The sensor-generated RPM readings are reliable in offering detailed insights into transient behaviors and short-term trends because of their high sampling frequency.

Three metrics, MAE, MSE, and Root RMSE, are computed in Table 3 to evaluate the accuracy and reliability of the sensor-generated RPM values. The MAE is 4.57 RPM, showing that the average deviation of the sensor-generated RPM values from the equational value is about 4.57 RPM, indicating that, on average, the sensor-generated RPM values deviate from the equational value by approximately 4.57 RPM. The MSE, calculated at 91.40, reflects the average squared magnitude of errors, indicating that while some errors are small, others are significantly larger. The RMSE, a more interpretable measure, stands at 9.56 RPM, suggesting that the typical deviation of sensor-generated RPM from the equational value is around 9.56 RPM. These metrics collectively reveal the accuracy and reliability of the RPM readings from the sensor in comparison to equational expectations. While the average deviation (MAE) is relatively modest, the larger errors contribute to higher MSE and RMSE values, highlighting variability in the machine's performance. Further analysis could focus on identifying the causes of these larger deviations and exploring ways to minimize them to achieve more precise RPM monitoring and control.

3.2. Ten-second interval RPM recorded values versus equational RPM

Table 4 Ten-second interval RPM recorded values

Time(s) (i)	Sensor-generated RPM (y_i)	Equational RPM (y)	Absolute Error [$y_i - y$]
10	1500	1500	0
20	1505	1500	5
30	1497	1500	3
40	1499	1500	1
50	1504	1500	4
60	1498	1500	2

Table 5 MAE, MSE, and RMSE of ten-second interval RPM recorded values

Metric	Value
MAE [$\frac{1}{n}\sum_{i=1}^{n}(y_i - y)$]	2.50
MSE [$\frac{1}{n}\sum_{i=1}^{n}(y_i - y)^2$]	9.17
RMSE [$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y)^2}$]	3.03

Table 4 displays recorded RPM values at 10-second intervals, including values at 10, 20, 30, 40, 50, and 60 seconds. The sensor-generated RPM values exhibit slight variations from the calculated 1500 RPM. The range of absolute errors is from 0 to 5 RPM, showing fairly minor differences. Table 5 presents the MAE, MSE, and RMSE for ten-second interval RPM data, providing essential metrics for evaluating the accuracy and reliability of the recorded values. The sensors's RPM values have an average deviation of 2.50 RPM from the equational value, as indicated by the MAE of 2.50 RPM. The MSE is 9.17, showing the average squared size of the errors, suggesting that the differences are usually minor. The RMSE is 3.03 RPM, offering a clear measure of average deviation in the original units. These metrics seemingly show that the RPM values generated by the sensor are fairly accurate over ten-second intervals, with minimal deviations from the expected value. The MAE, MSE, and RMSE values seems relatively small, indicating consistent performance with minimal errors.

3.3. Comparative analysis

3.3.1. Sensor-generated data versus equational calculations comparison

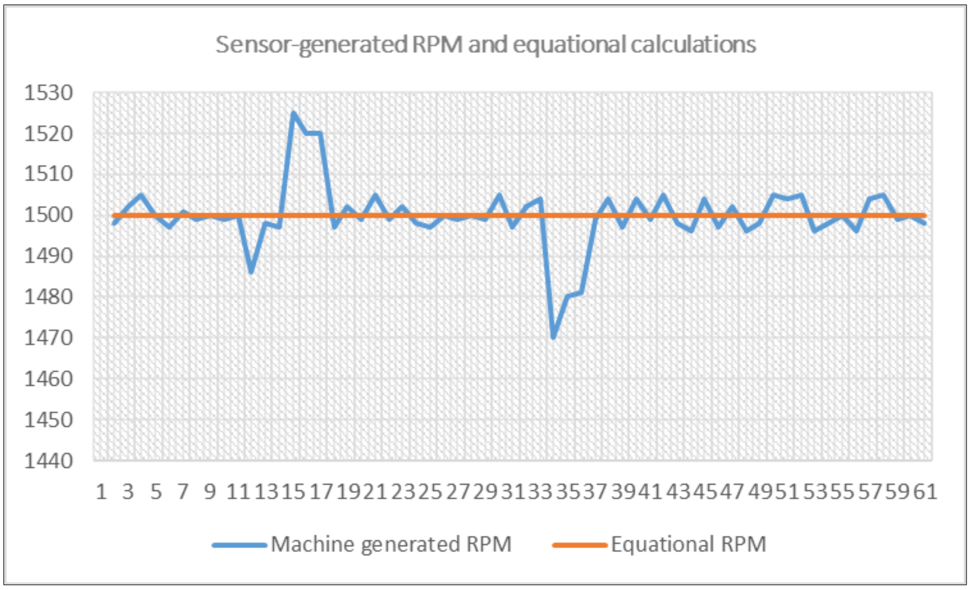


Figure 1 Comparison of machine-generated RPM and equational RPM

Figure 1 illustrates sensor-generated RPM data recorded at one-second intervals, it shows a range of absolute errors from 0 to 30 RPM. While the equational RPM value remains consistently at 1500. The sensor-generated values fluctuate, indicating variability and deviations. Despite this variability, the data provides detailed insights into the motor's performance, capturing transient behaviors and short-term trends that equational calculations is not able to capture. Significant deviations at specific points (e.g., 14, 15, 16, 33, and 34 seconds) suggest that sensor-generated data exhibits greater level of accuracy and reliability compared to equational value. The equational calculations, being consistent and devoid of fluctuations seems to offer more stability. Nonetheless, the sensor-generated data's higher granularity and ability to reflect real-time operational dynamics provide a comprehensive understanding that equational calculations alone cannot achieve. Therefore, while equational calculations seems to guarantee immediate data stability, sensor-generated data offers valuable detailed insights crucial for thorough performance analysis and real-time monitoring.

3.3.2. Data points and volume of generated data

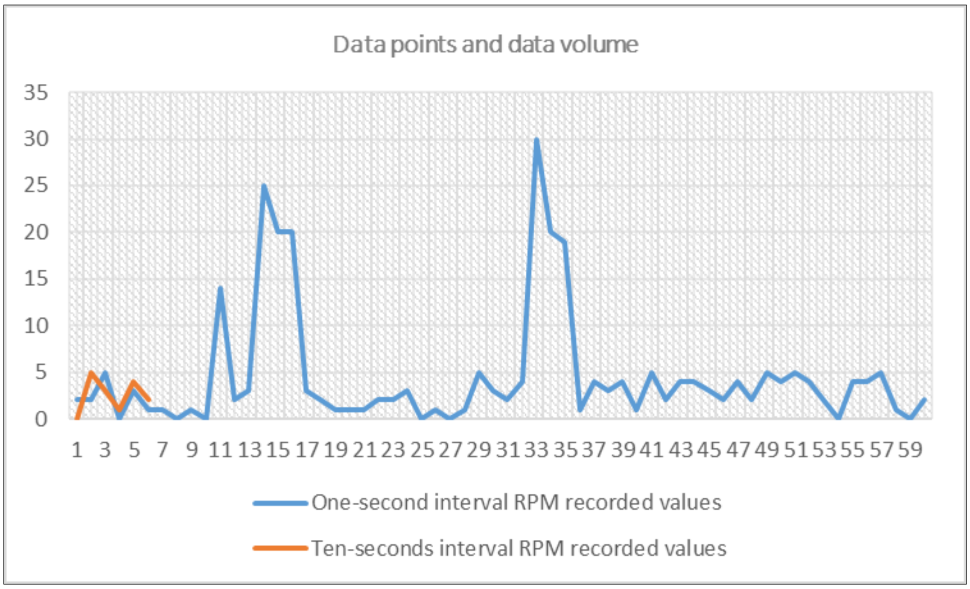


Figure 2 Comparison of data points and data volume of one-second and ten-second interval RPM recorded values.

From a data storage perspective, choosing between one-second and ten-second intervals for RPM data collection has significant implications. In Figure 2, the one-second interval generates a larger dataset with 60 data points, demanding more storage space and increased management overhead. This high volume necessitates robust storage solutions and greater processing power, potentially straining infrastructure and impacting performance if not managed efficiently. In contrast, the ten-second interval produces a smaller dataset of only 6 data points, reducing storage costs and simplifying data handling and analysis. While the one-second interval provides detailed insights into RPM variations, crucial for precise performance analysis but with higher data noise, the ten-second interval offers a seemingly cleaner overview suitable for broader trend analysis with lower data variability.

Comparatively, the one-second interval captures transient behaviors and short-term trends with a significant level of granularity despite larger errors and variability ranging from 0 to 30 RPM. This detailed data is advantageous for applications requiring precise operational insights but requires careful consideration of storage and processing resources. On the other hand, the ten-second interval, with its lower variability (errors ranging from 0 to 5 RPM), provides a seemingly more stable dataset, suitable for general performance assessments and trend analysis. The choice between these intervals hinges on balancing the need for detailed data against practical considerations of storage, processing capacity, and the specific analytical requirements of the application.

3.3.2. Mean absolute error, mean squared error and root mean squared error

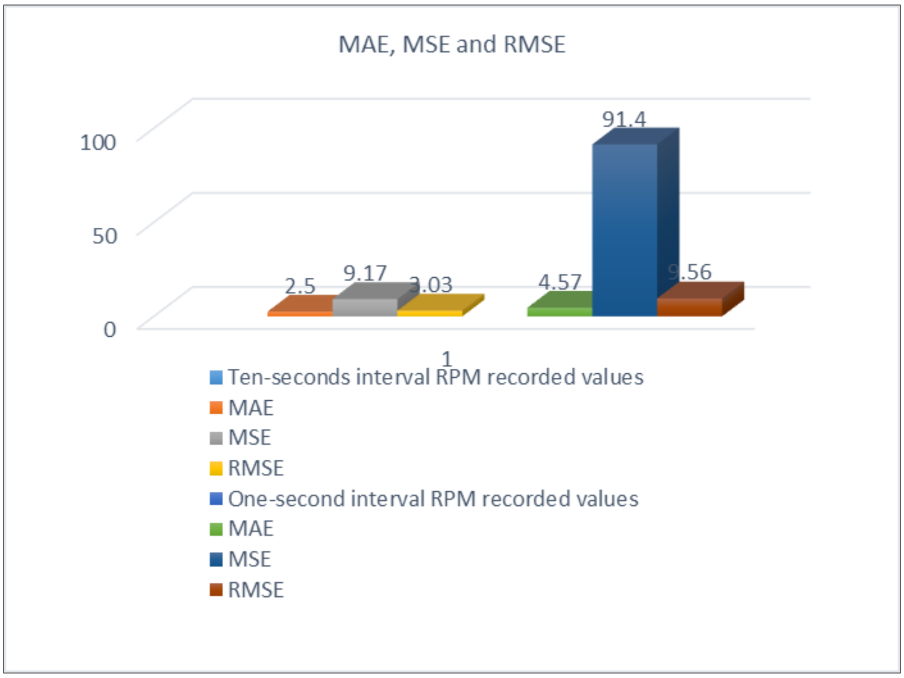


Figure 3 Comparison of MAE, MSE, and RMSE of one-second and ten-second interval RPM recorded values.

The comparative analysis of RPM recorded values between one-second and ten-second intervals reveals significant differences in the accuracy of metrics. In Figure 4, for the one-second interval data, the MAE is higher at 4.57 compared to 2.50 for the ten-second interval, indicating that shorter sampling intervals result in greater deviations from the true RPM values. Similarly, the MSE for the one-second interval is substantially larger at 91.40 compared to 9.17 for the ten-second interval, suggesting increased variability and error accumulation over shorter timeframes. This trend is further emphasized by the RMSE, which is 9.56 for the one-second interval and notably lower at 3.03 for the ten-second interval. The lower RMSE for the ten-second interval indicates better overall stability over longer sampling periods, albeit with potentially less detailed resolution of transient RPM changes compared to the one-second interval data.

The one-second interval data with an MAE of 4.57, MSE of 91.40, and RMSE of 9.56, exhibits higher variability and error. However, this shorter interval captures detailed trends and transient behaviors in the RPM values that might be crucial for identifying quick deviations and patterns necessary for real-time monitoring and responsive decision-making. This granularity allows for the detection of short-lived anomalies or rapid changes that could indicate emerging issues or performance shifts, making it valuable for applications that require high-resolution data for dynamic analysis.

Conversely, the ten-second interval data with significantly lower error metrics (MAE of 2.50, MSE of 9.17, and RMSE of 3.03), seems to offers better overall stability. However, it might smooth out or miss short-term fluctuations and critical deviations that occur within the shorter intervals. This can be a limitation in scenarios where immediate detection of rapid changes is essential for maintaining operational efficiency and responding promptly to potential problems. Therefore, while the ten-second interval provides higher data stability, it may fall short in capturing critical deviations necessary for efficient data analytics in dynamic environments. The choice between these methods depends on the specific requirements of the analysis: whether it prioritizes immediate accuracy and stability or a detailed and comprehensive understanding of behavioral trends.

3.4. Variability in electric motor RPM

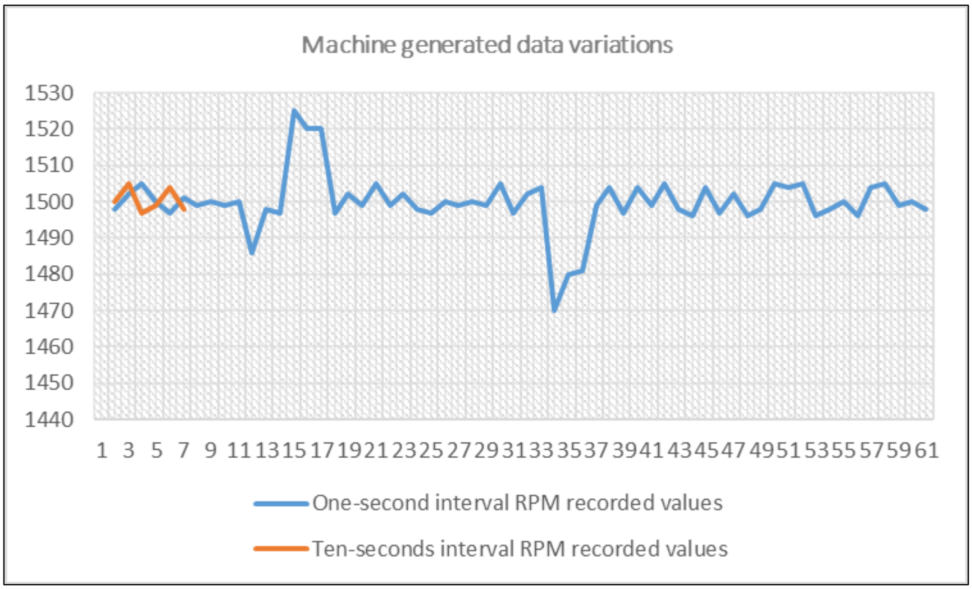


Figure 4 Comparison of machine-generated data of one-second and ten-second interval recorded values.

The operational performance of electric motors as highlighted in Table 2 reveals significant variations in RPM, Figure 3 also illustrates these trends with the aid of a graph. The figure shows a deviation in the one-second interval data acquisition and the ten-second data acquisition from the equational calculated value of 1500 RPM. These trends are influenced by diverse mechanical, electrical, and environmental factors.

Motors running below their rated RPM can stem from load-related issues such as overloading or variable load demands which strain the motor beyond its capacity and cause it to slow down. Voltage drops, phase imbalances in three-phase systems, and fluctuations in supply frequency also contribute to decreased motor speed. Mechanical problems such as bearing wear, shaft misalignment, and physical obstructions within the motor further hinder RPM performance. Electrical faults like winding issues and rotor defects can similarly impair motor efficiency, exacerbated by environmental factors such as high temperatures and humidity which degrade insulation and increase resistance.

Conversely, motors running above their rated RPM often result from power supply anomalies like overvoltage or frequency deviations where the motor receives excess energy beyond its designed limits. Malfunctions in control systems, including faulty VFDs or misconfigured speed controllers can also lead to unintended speed increases. Mechanical issues such as sudden load reductions or coupling problems between the motor and driven equipment can cause RPM spikes. Feedback loop errors, arising from faulty sensors or calibration issues in the control mechanisms further contribute to erratic motor speeds. To maintain consistent operation at rated RPM, regular maintenance practices, effective load management, and ensuring the integrity of power supply and control systems are essential, alongside monitoring and prompt resolution of mechanical wear and environmental conditions impacting motor performance.

4. Conclusion

Predictive maintenance, which integrates real-time sensor data monitoring, machine learning, and advanced data analytics, significantly enhances equipment health monitoring by predicting failures before they occur, thereby enabling targeted and prompt interventions. This study compared equational calculations and sensor-generated data, focusing on the stability, accuracy, and reliability of RPM data from an electric motor, recorded at varying time intervals to determine the optimal data collection methods for predictive maintenance.

Key findings indicate that sensor-generated data, when compared to equationally calculated values, provides detailed insights into transient behaviors, making it a highly reliable option for predictive maintenance. This scenario highlights the robustness of sensor-generated data in predictive maintenance applications. Furthermore, data collected at shorter intervals, such as one-second intervals, captures comprehensive trends and transient behaviors, offering a more nuanced and accurate understanding of electric motor performance. Despite the associated greater variability and larger errors with high-frequency data, it proves highly suitable for analyses requiring detailed operational insights.

This study underscores the benefits of short-interval data acquisition in capturing detailed operational insights, it also highlights that sensor-generated data can provide more detailed and timely insights into transient behaviors compared to equationally calculated value, making it a valuable tool for predictive maintenance. While both methods have their merits, sensor data offers the advantage of real-time monitoring and detailed trend analysis, enhancing the ability to predict and prevent equipment failures, aiding industrial stakeholders in making informed decisions regarding predictive maintenance and overall station performance optimization, Although statistical assessment metrics such as MAE, MSE, RMSE, and AE are traditionally used to evaluate model accuracy, they have been employed here to illustrate the extent of deviation and accuracy between sensor-generated data and theoretically computed values.

Compliance with ethical standards

Acknowledgments

This study was not supported by any grants from funding bodies in the public, private, or not-for-profit sectors.

Disclosure of conflict of interest

No conflict of interest to be discussed.

References

[1] Bhadani AK, and Jothimani D. Big data: challenges, opportunities, and realities. Effective big data management and opportunities for implementation. (2016); 1-24.

[2] Gudivada V, Apon A, and Ding J. Data quality considerations for big data and machine learning: going beyond data cleaning and transformations. International Journal on Advances in Software. (2017); 10(1): 1-20.

[3] Juriga D. Improving manufacturing data quality with data fusion and advanced algorithms for improved total data quality management. (2019).

[4] Tao F, Qi Q, Liu A, and Kusiak A. Data-driven smart manufacturing. Journal of Manufacturing Systems. (2018); 48: 157-169.

[5] Goosen AE. A system to quantify industrial data quality (Doctoral dissertation, North-West University (South Africa). (2019).

[6] Jabir B, and Falih N. Big data analytics for strategic and operational decisions. In Artificial Intelligence and Industry 4.0. Academic Press. (2022); 195-214.

[7] Mandinach EB, Honey M, and Light D. A theoretical framework for data-driven decision making. In annual meeting of the American Educational Research Association, San Francisco, CA. (2006); 39-52.

[8] Janssen M, Van Der Voort H, and Wahyudi A. Factors influencing big data decision-making quality. Journal of business research. (2017); 70: 338-345.

[9] Jeble S, Kumari S, and Patil Y. Role of big data in decision making. operations and supply chain management: An International Journal. (2017); 11(1): 36-44.

[10] Provost F, and Fawcett T. Data science and its relationship to big data and data-driven decision making. Big Data. (2013); 1(1): 51-59.

[11] Duggineni S. Data integrity and risk. Open Journal of Optimization. (2023); 12(2): 25-33.

[12] Lahdelma S, and Juuso E. Trend analysis in condition monitoring of process equipment. In Proceedings of the 8th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM. (2011): 904-913.

[13] Arena S, Manca G, Murru S, Orrù PF, Perna R, and Reforgiato Recupero D. Data science application for failure data management and failure prediction in the oil and gas Industry: a case study. Applied Sciences. (2022); 12(20): 1-29.

[14] Garcia ACB, Bentes C, de Melo RHC, Zadrozny B, and Penna, TJ. Sensor data analysis for equipment monitoring. Knowledge and Information Systems. (2011); 28: 333-364.

[15] Guo A, Liu X, and Sun T. Research on key problems of data quality in large industrial data environment. In Proceedings of the 3rd International Conference on Robotics, Control and Automation. 2018; 245-248.

[16] Naydenova I, Kovacheva Z, and Kaloyanova K. Important data quality accents for data analytics and decision making. automation, robotics and communications for industry 4.0. (2021); 90.

[17] Makinen MV. Data quality in smart manufacturing. (2020).

[18] Pech M, Vrchota J, and Bednar J. Predictive maintenance and intelligent sensors in smart factory. Sensors. (2021); 21(4): 1470.

[19] Karunasingha DSK. Root mean square error or mean absolute error? Use their ratio as well. Information Sciences. (2022); 585: 609-629.

[20] Chai T, and Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE). Geoscientific model development discussions. (2014); 7(1): 1525-1534.

[21] Hanson B, Weigand C, and Orloff S. Performance of electric irrigation pumping plants using variable frequency drives. Journal of Irrigation and Drainage Engineering. (1996); 122(3): 179-182.