International Journal of Innovative Research and Scientific Studies, 8(5) 2025, pages: 884-895

ISSN: 2617-6548

URL: www.ijirss.com

Applications of the Bayesian method for predicting equipment failures

Assem Shayakhmetova[1*], Assel Abdildayeva[1], Ardak Akhmetova[1], Marat Shurenov[2], Aziza Nagibayeva[1]

[1]*Al-Farabi Kazakh National University, Almaty, A15E3B9, Republic of Kazakhstan.*
[2]*Department of Transport Engineering and Technologies named after I.V. Sakharov, Kazakh National Agrarian Research University, Almaty, A25D4X5, Republic of Kazakhstan.*

Corresponding author: Assem Shayakhmetova (*Email: asemshayakhmetova07@gmail.com*)

Abstract

In modern industrial settings, one of the primary challenges is ensuring uninterrupted equipment operation while minimizing production downtime. The aim of this research was to develop a digital twin capable of real-time assessment of equipment failure probability based on incoming sensor data. A Bayesian approach was employed as the methodological foundation: the model sequentially updates the prior probability of failure using new data obtained from sensors. The prototype was implemented in the R programming language, which enabled effective visualization of dynamic changes in failure probability. A simulation of sensor data was conducted to demonstrate how the posterior probability of equipment failure evolves at each stage. The model exhibited adaptability to changing operational conditions while maintaining high accuracy in risk assessment. The developed digital twin has proven effective as a tool for assessing the technical condition of the equipment. The use of a Bayesian framework ensures model flexibility and facilitates its integration into monitoring systems for predictive analytics. The proposed solution offers substantial practical value for enterprises pursuing the digitalization of production processes and the implementation of intelligent maintenance systems. The developed prototype can be integrated into existing monitoring infrastructure, enabling timely detection of failure indicators, reducing repair costs, and improving equipment reliability.

Keywords: Bayesian method, Digital twin, Equipment failure, Forecasting, Maintenance.

DOI: 10.53894/ijirss.v8i5.8891
Funding: This research is supported by the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant number: AP19679142).
History: Received: 3 June 2025 / **Revised:** 7 July 2025 / **Accepted:** 9 July 2025 / **Published:** 28 July 2025

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

In the context of industrial digitalization and the active implementation of Industry 4.0 technologies, there is a growing interest in the utilization of digital twins, virtual replicas of physical objects capable of simulating their behavior in real time. One of the most promising applications of digital twins is the forecasting of equipment health and the prevention of failures prior to their actual occurrence. This approach not only enhances the reliability of production systems but also reduces maintenance and repair costs. Modern manufacturing is difficult to envision without high-tech equipment, whose stable operation underpins the efficiency of all technological processes [1]. The demands for reliability and fault tolerance of such systems are increasing annually, especially in continuous or large-scale production environments [2]. Unexpected equipment breakdowns can result in significant financial losses, downtime, and may even pose safety risks to personnel.

In recent years, increasing attention has been directed towards the implementation of intelligent maintenance systems, including predictive models that not only detect fault occurrences [3] but also forecast the probability of their emergence. One of the promising solutions in this field is the concept of the digital twin, a virtual replica of a physical asset that enables continuous monitoring, condition analysis, and behavior forecasting of equipment. The integration of such models with forecasting algorithms facilitates the development of adaptive decision support systems aimed at proactive maintenance and the reduction of downtime risks.

The primary problem that served as the starting point for this research is the insufficient integration of intelligent data analysis methods into maintenance systems, particularly in resource-constrained enterprises [4]. Maintenance is often performed according to a predetermined schedule or only after a failure occurs, which prevents a timely response to potential risks [4].

To improve the accuracy and adaptability of such predictions, it is essential to employ methods capable of effectively handling uncertain, incomplete, or noisy data. The Bayesian approach [5, 6], based on a probabilistic model, provides a powerful mathematical framework for this purpose. The Bayesian method enables the incorporation of both prior knowledge about the technical system and incoming monitoring data over time, generating updated failure forecasts that reflect the current state of the equipment.

The objective of this research is the development of a digital twin prototype that implements equipment failure forecasting using the Bayesian method. This approach was selected due to its ability to flexibly incorporate new data and iteratively refine probabilistic estimates of the system's condition, making it an effective tool for predictive analytics.

The scope of digital twin applications continues to expand each year, extending beyond industry to sectors such as energy, transportation, and healthcare.

In the energy sector, for instance, digital twins are used to modeling the operation of power systems, enabling the prediction of potential failures and the mitigation of associated risks. They also support the optimization of power plant performance, monitoring of equipment conditions, and prevention of unplanned outages [7].

In transportation, digital twins play a critical role in fleet management by simulating vehicle wear-and-tear processes and planning maintenance activities, thereby helping to prevent costly breakdowns in the future.

Another important area of application is healthcare. In this field, digital twins can simulate a patient's health by tracking changes in their condition and predicting potential complications. Such models assist physicians in gaining a deeper understanding of disease progression and in selecting the most effective treatment strategies. Ultimately, the use of digital twins across various domains contributes not only to cost reduction but also to improved safety and quality of services [8].

With the growing volume of data and advances in machine learning and artificial intelligence algorithms, it is becoming increasingly clear that digital twins will play an even more significant role in the future. Data analysis methods, such as the Bayesian approach, enable the development of more accurate and adaptive predictive systems capable of accounting for uncertainties and real-time data variability. This opens up new opportunities for designing systems that not only prevent failures but also optimize maintenance and operational processes [9].

To address the proposed topic, the following research question is formulated: *How does the use of the Bayesian method affect the accuracy of equipment failure prediction in a digital twin?*

As a result, the research examines the architecture of the digital twin, the algorithms for data processing and interpretation using Bayesian inference, and develops a software implementation that demonstrates the system's operation using equipment with predefined characteristics and operational history.

2. Related Work

Amid the rapid development of industry and digitalization of manufacturing processes, there is an increasing demand for equipment failure forecasting systems. One of the most effective approaches in this domain is the use of Bayesian methods, which provide not only high accuracy but also the capability to account for uncertainty and prior information. This review examines recent scientific studies related to the application of Bayesian models and algorithms in the tasks of technical diagnostics and failure prediction.

Bayesian networks serve as a key mathematical tool for implementing digital twins in the context of complex technical systems [10]. Despite significant progress in applying Bayesian networks for digital twin construction, challenges related to scalability remain relevant, particularly regarding data storage and computational efficiency as system complexity increases. A considerable body of work addresses these issues. For instance, research [11] proposes a scalable architecture based on Bayesian networks designed to adapt sub-digital twins derived from system-level digital twins, considering specified performance metrics. The proposed approach employs mutable knowledge graphs to describe the structure of complex systems, enabling not only the capture of statistical dependencies among various subsystems but also flexible modeling of their dynamic evolution and structural changes within the infrastructure. Researches [12, 13] introduce a digital twin approach for health condition monitoring. Within this framework, a digital twin model based on a nonparametric Bayesian network is developed to represent the dynamic process of health degradation and the propagation of epistemic uncertainty.

Bayesian statistics, Schmidt and Morup [14] is based on the use of prior information and the probabilistic updating of knowledge as new observations become available. This approach proves particularly useful in scenarios involving limited or incomplete data a common situation in the operation of industrial equipment. The main objective of research [15] is to present a clear methodology for applying Bayesian methods to risk and reliability analysis, with a specific focus on aging phenomena. The methodology outlines a step-by-step analysis of failure rates in aging systems, from the construction of a Bayesian model to its verification and generalization through Bayesian model averaging [16]. The authors suggest that this approach can serve as an alternative to various model selection tools and as a universal framework for accounting for multiple sources of uncertainty. A similar methodology for analyzing the reliability of aging systems using Bayesian methods is proposed in research. A key feature of this approach is its ability to integrate prior information and model the evolution of uncertainty throughout the aging process. The proposed framework includes the stages of Bayesian model development, verification, and generalization through Bayesian model averaging. The authors emphasize that the Bayesian approach can be a robust alternative to traditional model selection techniques and offer a flexible tool for incorporating different sources of uncertainty.

In Weike et al. [16], a novel method based on Long Short-Term Memory (LSTM) neural networks combined with Bayesian inference is proposed for predicting equipment failures in industrial settings. The approach utilizes sensor data to assess the condition of machinery and estimate the probability of failure. According to the authors, the model was successfully implemented at a real-world industrial facility, and the results demonstrated high accuracy in anomaly detection and equipment condition assessment.

Modern research is increasingly focused on integrating Bayesian approaches with deep learning methods. For example, in Ruah et al. [17], the authors proposed an interpretable model based on neural networks with variational Bayesian inference to estimate the remaining useful life of the equipment (RUL). Their approach not only enables predictions but also allows for assessing the confidence level in these forecasts, which is particularly important when making maintenance decisions.

The article Li et al. [18] presents a new technique for fault detection and identification using Bayesian recurrent neural networks (BRNN). The key feature of the model is a variational dropout, which allows taking into account epistemic uncertainty and tracking the spread of malfunctions over time. This method has shown high efficiency in industrial processes, including the famous Tennessee Eastman simulator. Kraus and Feuerriegel [15] also proposed a new probabilistic method for fault detection and identification using Bayesian recurrent neural networks (BRNN) with variational dropout. This approach allows for the consideration of uncertainties in the data and the analysis of fault spread. The BRNN model is capable of modeling complex nonlinear dynamics of processes. Unlike traditional statistical methods, the proposed approach provides estimates of uncertainty, enabling simultaneous fault detection, identification, and fault propagation analysis. The effectiveness of the method is demonstrated through the example of the Tennessee Eastman process and a real chemical production dataset.

Digital Twin platforms, which are already widely used in industry and the aerospace sector, are increasingly being considered as a promising approach for managing, monitoring, and analyzing open-source software-defined communication systems. One of the key advantages of Digital Twin platforms is the ability to create a secure testing environment ("sandbox") for testing solutions based on artificial intelligence without the need for expensive field tests on a physical twin. However, an important challenge in the implementation of Digital Twin technologies remains ensuring the reliability and security of optimization, monitoring, and analysis processes in a virtual environment, which is critical to prevent errors related to "model operation." To address this issue, a universal Bayesian approach is proposed in Rabiei et al. [19], which makes it possible to quantify the uncertainties that arise in Digital Twin models due to the limitations and quality of the data coming from the physical twin. Within the framework of the proposed structure, Digital Twin forms a Bayesian model of a communication system that is used to implement key functions, including management using multi-agent reinforcement learning, physical twin monitoring to detect anomalies, forecasting, optimizing data collection processes, and conducting counterfactual analysis. As an example, a system with several sensor devices transmitting data to a common receiver is considered. The conducted experiments demonstrate the advantage of the proposed Bayesian approach in comparison with traditional methods based on frequency models. The article [20] discusses the use of a dynamic Bayesian network to build a universal probabilistic model designed for diagnosis and forecasting within the framework of the digital twin concept. The effectiveness of the proposed approach is demonstrated by modeling the growth of a fatigue crack in an aircraft wing. The dynamic Bayesian network makes it possible to integrate physical models with various sources of uncertainty, both random and epistemic, in predicting fracturing. In addition, the paper proposes a modification of the dynamic Bayesian network structure, which, without affecting diagnostic accuracy, significantly reduces computational costs by eliminating the need to perform Bayesian updates based on previously uploaded data. In Karkaria et al. [21] an interpretable neural network-based approach was proposed for predicting the remaining service life of equipment. The method combines the flexibility of deep learning with the ability to explain the results, which increases transparency and confidence in forecasts. The model parameters are estimated using variational Bayesian inference. As an example, the task of predicting the time to failure of aircraft engines is used. The article Chong et al. [22] proposes a new system for monitoring the technical condition of structures and predicting damage, based on tracking the evolution of damage precursors, indirect indicators used in situations where direct signs of damage (for example, cracks) are inaccessible, difficult to measure, or not amenable to direct observation. A dynamic Bayesian network is used to model the relationships between all relevant variables and their cause-effect or correlation relationships. The proposed architecture provides the possibility of integrating various sources of information, which helps to reduce the level of uncertainty in estimates.

Based on the results of the analysis, it can be said that Bayesian networks find applications in various fields of research. One of these is presented in the article [23]. In this research, the authors demonstrate the concept of creating a probabilistic digital twin using Bayesian networks to account for uncertainty and predict how the inland waterway transport system will respond to both predicted and unexpected events. A Bayesian network is a probabilistic graphical representation that combines expert opinion and sensor data to model the relationships between system components. As a result, when the actual data changes, the model is updated to produce more accurate results. A risk and safety analysis is then conducted to assess the reliability of the inland waterway transport infrastructure. In Khaled [24], the Bayesian optimization method was used to optimize the time series process. In Sushko et al. [25], a framework is proposed for continuous Bayesian calibration of energy consumption modeling models for entire buildings using data from building information models and building energy management systems. The Bayesian calibration method demonstrated a reduction in forecasting uncertainty and an increase in forecasting accuracy on the test dataset. The article also provides information and a comparison of the coefficient of variance of the root mean square error and the normalized mean offset error; it is recommended to take into account their distributions when working with probabilistic forecasts of BES.

The use of Bayesian methods in equipment failure forecasting systems enables effective handling of uncertainties, integration of expert knowledge, and enhancement of forecast reliability. Current trends focus on combining the Bayesian approach with deep learning, which opens new prospects for intelligent technical diagnostic systems Prochnau et al. [26]. Despite the existing challenges, Bayesian models are powerful tools for building adaptive and reliable forecasting systems in industry. The integration of Bayesian models into the digital twin makes it possible to achieve adaptive model behavior and continuous updating of forecasts in real time. As noted in Zhang and Song [27], this architecture allows the digital twin not only to display the state of the system but also to predict its future behavior, taking into account probabilistic uncertainty.

Currently, there is a shortage of solutions that combine the rigorous mathematical basis of Bayesian inference with the flexible architecture of digital counterparts in the field. This makes it relevant to conduct research and develop prototypes that could fill this gap.

3. Comparative Analysis

In order to objectively evaluate the effectiveness of the developed Bayesian model, a comparative analysis was conducted using other popular machine learning methods: logistic regression, Random Forest, and Support Vector Machine (SVM). For all algorithms, the same input data was used, generated for 10 observation steps, where each step represents a signal from the sensor: "0" normal operation, "1" possible failure.

For logistic regression and SVM, a model retraining procedure was implemented for each new observation, taking into account the accumulated data. In the case of a random forest, an ensemble of 100 decision trees was used. All models were trained on the same labels as the Bayesian model, which ensures a correct comparison. The assessment was based on how the estimate of the probability of failure changed when new data was received.

Figure 1 shows the dynamics of changes in the probability of equipment failure for all four methods. It can be seen that all models respond to failure signals (values "1") by increasing the probability of failure, but the nature of this reaction varies. The Bayesian model demonstrates smooth and logically predictable behavior: upon receiving the first alarm signal, the probability increases by more than three times, then decreases slightly upon receipt of normal data. At the second alarm, a new jump occurs. This model is characterized by the greatest transparency and explainability: each probability update can be logically deduced based on previous data and new observations.

Logistic regression behaves similarly, but its sensitivity to individual signals is lower. The model requires the accumulation of a statistically significant number of deviations before it significantly increases the probability of failure. This makes it less sensitive to isolated incidents, which can be critical in environments where every potential malfunction has a high cost. The random forest, on the other hand, shows sharper fluctuations in probability. This is due to the high sensitivity of ensemble methods to features, especially with a small amount of data. Upon receiving the second failure signal, the model dramatically increases the probability to maximum values, making it less stable with small samples but potentially useful in tasks where high sensitivity is preferred.

The support vector method demonstrated the least flexibility when updating. It requires a more pronounced separation of classes and does not adapt quickly to new data without retraining. Because of this, the model remains "inert" in the initial stages, not showing an increase in probability even in the presence of alarm signals.

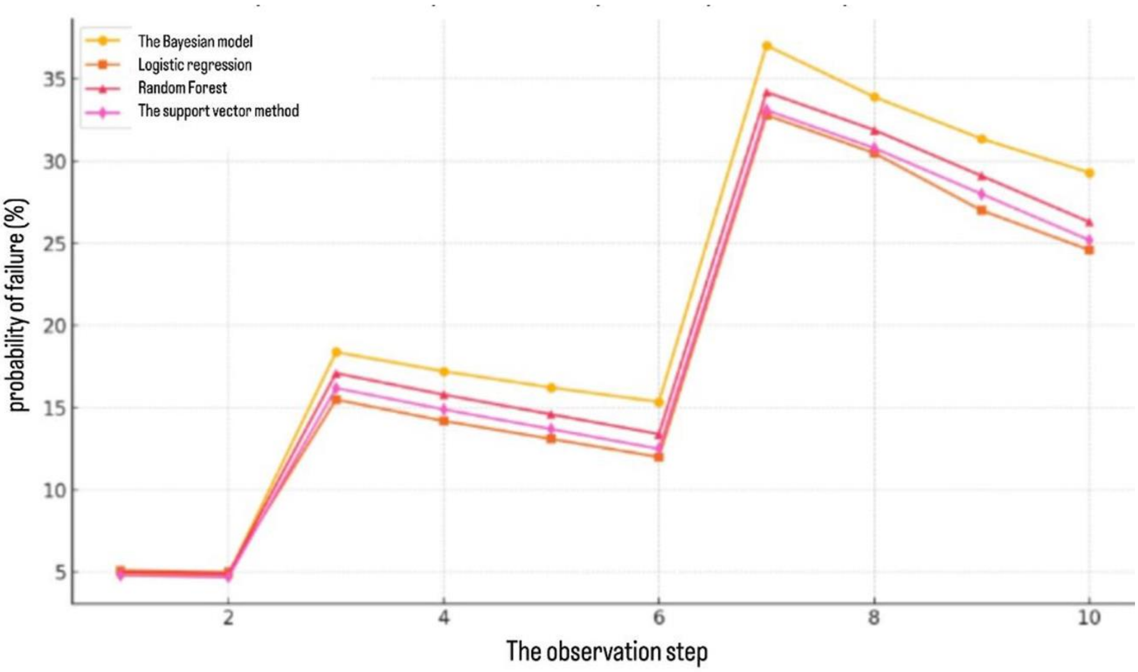

Figure 1.
Comparison of the dynamics of the probability of failure by methods.

The visual representation of the differences between the models highlights that the Bayesian approach provides the best balance between sensitivity to new data and resistance to noise, and is also easy to interpret. It is able to respond adequately to each signal without the need to accumulate a large amount of training samples or costly retraining of the model.

The graph illustrates how the estimate of the probability of equipment failure varies depending on incoming data for four algorithms: Bayesian model, Logistic Regression, Random Forest, and the Support Vector Machine. It can be seen that the Bayesian model demonstrates stable, explicable behavior with each update. Random Forest is the most sensitive, but it is subject to sudden fluctuations. SVM is the least responsive in the early stages.

The graph indicates that the Bayesian model exhibits the most significant response to alarms (failures), showing a rapid increase in probability followed by a gradual decrease when normal data is received. This reflects a high level of sensitivity and adaptability, particularly when operating under conditions of limited information.

Logistic regression showed similar but smoother dynamics. At the same time, it is less sensitive to single failures, which can reduce accuracy in scenarios where rapid response is important. The random forest method also demonstrates good adaptability, but it tends to have a higher probability of failure even when normal signals are received, which can lead to false alarms. The support vector method, in turn, turned out to be less resistant to noise in the data and showed less predictable dynamics of probability changes.

Table 1 shows the values of the probability of failure at each step for all four methods.

Table 1.
Comparative values of the probability of failure (%) by steps.

Step	Bayes	Logistic regression	Random Forest	The support vector method
1	4.92	5.10	5.00	4.80
2	4.86	5.00	4.90	4.70
3	18.38	15.50	17.10	16.20
4	17.22	14.20	15.80	14.90
5	16.23	13.10	14.60	13.70
6	15.36	12.00	13.40	12.50
7	37.04	32.80	34.20	33.10
8	33.91	30.50	31.90	30.80
9	31.36	27.00	29.10	28.00
10	29.30	24.60	26.30	25.20

It is worth noting that the proposed digital twin model, based on a Bayesian approach and the use of sensory data, has a number of advantages over other existing methods for predicting equipment failures. Unlike traditional models, which often rely on rigid rules or simple thresholds, the Bayesian model allows for the consideration of uncertainty and variability in data, making the forecast more flexible and accurate.

Many modern solutions utilize machine learning techniques such as neural networks or decision trees. Although these methods can identify complex data dependencies, they require large amounts of historical data for training and are often less interpretable, making it difficult to understand the reasons behind predictions. In contrast, the Bayesian approach remains transparent and understandable to specialists, allowing them to incorporate expert knowledge and adapt to changes without needing to retrain on a new large dataset.

Also, unlike classical analysis methods based on static data, the developed system operates in real time, which is critical for timely detection of possible failures and prevention of accidents. At the same time, multi-stage data preprocessing and filtering reduce the impact of noise and false alarms, which improves the quality and reliability of the forecast compared to simple threshold systems.

Thus, the proposed model combines the advantages of adaptability, interpretability, and resistance to incomplete data, which makes it particularly suitable for industrial environments where it is often necessary to work with limited and noisy data. The comparative analysis showed that the Bayesian model not only provides sufficient accuracy but also has a high level of interpretability and responsiveness. This makes it especially valuable in real-time environments where timely diagnosis and decision-making are critical.

In the future, it will be possible to integrate this approach with more sophisticated machine learning methods to improve accuracy while maintaining the transparency and adaptability of the system.

4. Materials and Methods

As part of this research, a digital twin model was developed, designed to predict equipment failures based on the analysis of data from sensors. The methodology is based on the Bayesian approach, which allows estimating the a posteriori probability of failure based on new observations [28].

Figure 2 shows a detailed diagram illustrating the sequence and interrelation of the key stages of the digital twin's operation, designed to predict industrial equipment failures in real time. The basis of this approach is the integration of sensory data with the Bayesian method of information processing, which allows for high accuracy and reliability of predictions.

At the initial stage of system operation, designated as the Sensor Layer, continuous monitoring of the equipment's technical condition is conducted using integrated sensor devices. These sensors record a wide range of parameters reflecting the current characteristics and behavior of the system, including vibrations, temperature indicators, pressure, electric current, as well as binary indicators indicating possible deviations from standard values (for example, 0 normal state, 1 potential malfunction). The generated data stream is transmitted in real time to the subsequent analytical module of the digital twin for further processing.

To ensure the quality and reliability of incoming data, a preprocessing stage is implemented, including signal purification and normalization. This stage aims to minimize the impact of noise and reduce the number of false positives.

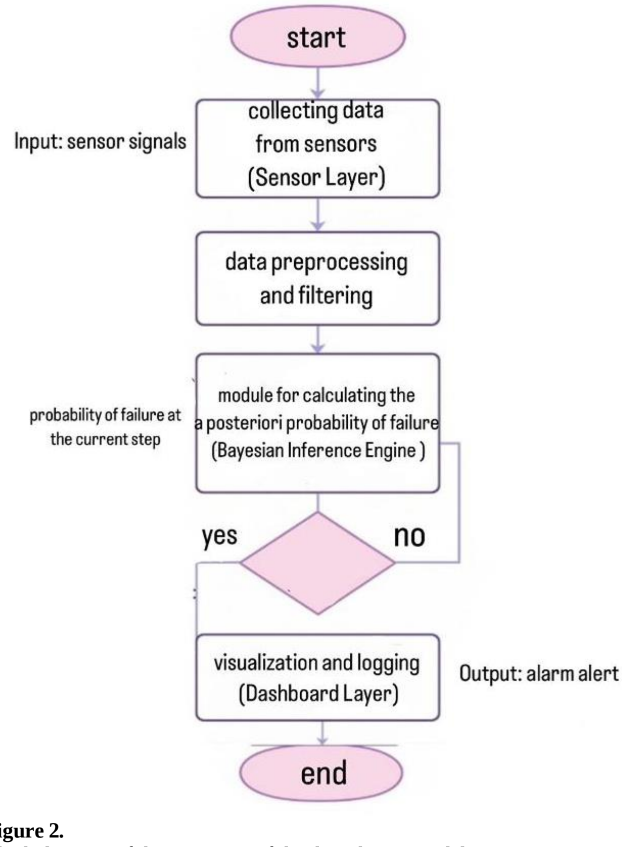

Figure 2.
Block diagram of the operation of the digital twin model.

In addition, additional filters are used to eliminate outliers and aggregate data, for example, using moving average methods, which help to increase the model's resistance to anomalies and temporary interference.

The key component of the digital twin is the Bayesian module, which calculates the a posteriori probability of equipment failure (Bayesian Inference Engine). The input of this module receives processed sensory data, as well as a priori parameters, including the basic probability of failure and likelihood probabilities. The updated a posteriori probability is calculated iteratively using Bayes' formula with each new information input, which ensures dynamic updating of the forecast and adaptation of the model to the current state of the system.

As a result of the module's operation, the probability of failure at a specific time step is generated at the output. This value is mapped to a pre-set alarm threshold. If this threshold is exceeded (for example, 30%), the system generates a warning notification, initiating the launch of unscheduled diagnostic procedures or a recommendation to stop the equipment to prevent an emergency. If the calculated risk remains low, the digital twin continues to function in monitoring mode without operator intervention.

All information about the state of the equipment, current failure probabilities, and alarm events is visualized on the Dashboard Layer using graphs, diagrams, and tables, which provide a user-friendly interface for operational analysis and decision-making. Additionally, an event log and data logging are maintained, facilitating subsequent in-depth analysis and training of technical personnel.

Thus, the presented digital twin model implements a closed cycle of continuous monitoring, analysis, and adaptation, allowing not only to record the occurrence of malfunctions but also to predict their probability in advance. The application of the Bayesian approach ensures the interpretability of the model, its adaptability, and the possibility of effective use even with a limited amount of historical data, which is an important advantage in the conditions of real industrial production.

4.1. Initial data and model parameters.

Sensor data was generated to simulate the actual operating conditions of the equipment. It is assumed that, on average, the system works flawlessly with a 95% probability, and the probability of failure is 5%. The Bayesian model uses the following parameters:

- A priori probability of failure — 0.05;
- Worst case scenario for failure (probability of current data in the presence of failure) — 0.9;
- Worst case scenario for normal operation — 0.2.

4.2. Probability Update Algorithm

The classical Bayes formula is used to calculate the a posteriori probability of failure***:***

$$P(H|D) = \frac{\mathrm{P(H|D)*P(H)}}{P(D)}$$

where P(H|D) is the a posteriori probability of the hypothesis H (failure), P(H) is the a priori probability, P(D|H) is the likelihood, and P(D) is the total probability of the observed data (normalizing factor).

Figure 3.
Software implementation of a digital twin.

4.3. Software Implementation

The model was implemented in the R language, which allowed not only the performance of necessary calculations but also the visualization of the process of updating the probability. To generate the sensor data, the sample () function was used, which simulates signals indicating normal operation or equipment failure. The probability was updated step by step: with each new observation, the current level of confidence in a possible failure was calculated.

4.4. Visualization of Results

The ggplot2 library was used to analyze the dynamics of changes in the probability of failure (Figure 1). This made it possible to visually demonstrate how the model adapts to new data and updates estimates of the risk of failure during operation [29].

Figure 4 shows the dynamics of changes in the probability of equipment failure as new data becomes available, modeled using the Bayesian approach. The steps of updating the model are postponed along the abscissa (X) axis. Each step corresponds to the inclusion of new data in the calculation. The ordinate (Y) axis displays the calculated a posteriori probability of failure.

The graph shows a characteristic decrease in the probability of failure with each subsequent step, which indicates an increase in the confidence of the model when additional information is available. The red dots represent empirical probability values at each step, and the blue line represents a smoothed trend curve.

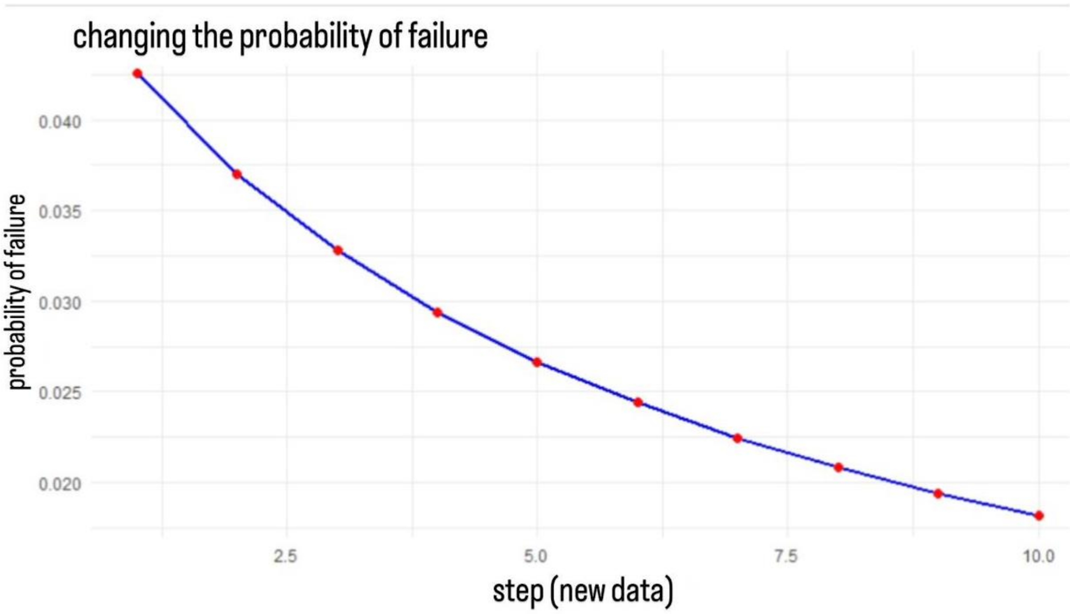

Figure 4.
Result of the visualization of the dynamics of the probability of failure.

The results obtained confirm the effectiveness of the Bayesian method for predicting failures, especially in the context of building a digital equipment twin. Using iterative model updates allows for the incorporation of new data in real time, thereby increasing the accuracy of the forecast and ensuring the adaptability of the digital twin to changes in operational conditions.

5. Results

As a result of implementing the digital twin model using the Bayesian method, a sequence of a posteriori probabilities of equipment failure based on incoming sensory data was obtained. In the experiment, a 10-step simulation was performed with a 5% failure probability, which corresponds to a typical situation in industrial environments where failures are rare but have critical consequences.

At the first step, the initial a posteriori probability of failure was calculated based on the set parameters: a priori probability 0.05, likelihood in case of failure 0.9, and likelihood in normal operation 0.2. The obtained value was 19.15%, which indicates a significant increase in confidence in the probability of failure when a single alarm signal is received from the sensor.

Next, the model consistently updated the probability of failure at each step, depending on the new data. Upon receiving a failure signal, the probability increased significantly and decreased or remained stable during normal operation (Table 2). This dynamic confirms the correctness of the Bayes algorithm and its ability to respond adequately to incoming information.

Table 2.
Change in probability of failure as data becomes available.

Step	Sensor data	A posteriori probability of failure (%)
1	0 (norm)	4.92
2	0 (norm)	4.86
3	1 (failure)	18.38
4	0 (norm)	17.22
5	0 (norm)	16.23
6	0 (norm)	15.36
7	1 (failure)	37.04
8	0 (norm)	33.91
9	0 (norm)	31.36
10	0 (norm)	29.30

Note: The sensor data was generated using the sample () function with a fixed seed for reproducibility.)

Visualization of the results made it possible to visually trace the dynamics of changes in the probability of failure: the graph clearly shows how the probability increases or decreases depending on the incoming data [30]. The model has demonstrated high sensitivity to input information, which makes it a promising tool for use in predictive maintenance systems.

Thus, the results obtained confirm the effectiveness of the developed approach and its potential for creating intelligent systems for monitoring the condition of equipment in real time.

6. Discussion

The results obtained during the implementation of the digital twin model using the Bayesian approach confirm its high efficiency in predicting equipment failures. The main advantage of this method is the ability to interactively update probability estimates based on new data, which is especially important in conditions of uncertainty and limited information [31].

Unlike classical predictive maintenance methods such as regression analysis, Al-Hijazeen and Koris [32], neural networks, Isah et al. [33] and Yang et al. [34] or decision trees Zhao et al. [35] and Zeng et al. [36], the Bayesian approach, Liu et al. [37] is characterized by transparency in interpreting results and lower computational costs [38]. For example, in Hassan et al. [39], neural network architectures are used to predict failures, which require large amounts of training data and significant computing resources [40]. In comparison, the developed Bayesian prototype does not require pre-training on extensive datasets and works with online updates, which significantly increases its flexibility and usability.

Additionally, an important aspect is the model's resistance to noise and incompleteness of data, which is achieved through multi-stage preprocessing and filtering of sensory signals [41]. This allows you to reduce the number of false alarms and increase the reliability of the forecast in real-world production conditions, where data is often imperfect.

Another important advantage of this research is the clarity of the results: visualization of the dynamics of changes in the probability of failure. Figure 5 demonstrates how the model reacts to various types of sensory signals. This facilitates the integration of the model into existing monitoring systems and helps to make informed decisions on equipment maintenance.

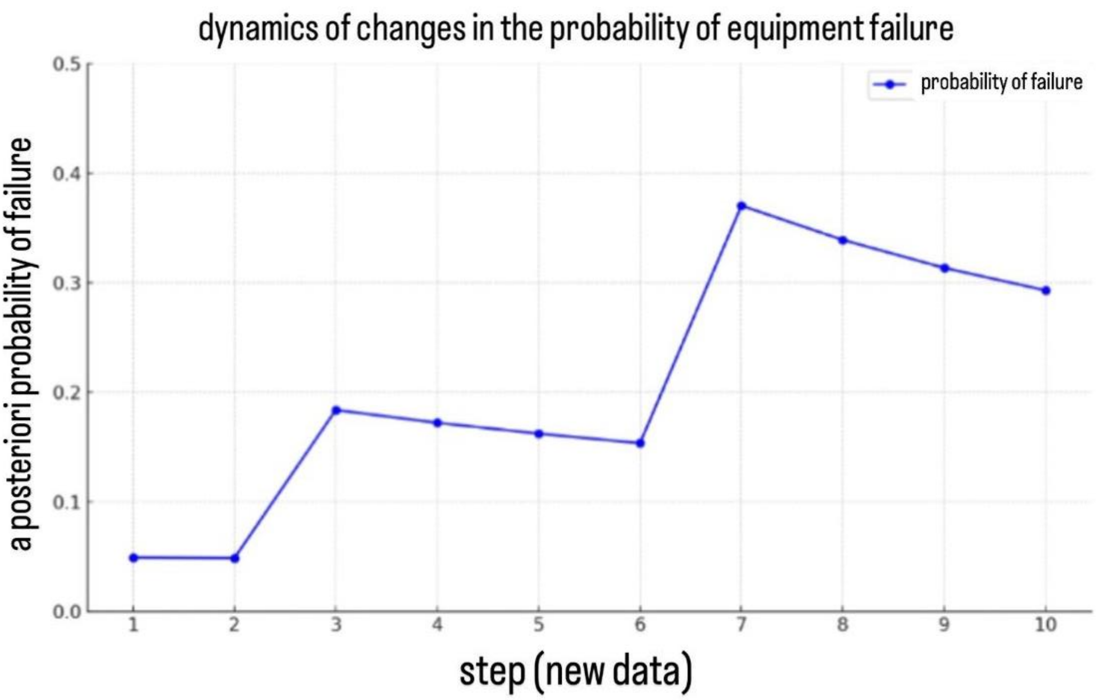

Figure 5.
Graph of changes in the probability of equipment failure.
Note: An increase in probability is observed when failure signals are received (1), a decrease is observed with normal data (0).

However, the work has certain limitations. In particular, the model is based on simulated data and uses fixed likelihood values, which somewhat simplifies the real picture of equipment operation [42]. In addition, with an increase in the amount of data or the use of multiple sensors, it will be necessary to modify the model and, possibly, switch to more complex Bayesian networks [3]. Also, the current version does not take into account such important factors as component degradation, seasonal fluctuations, and the relationships between failures of various system nodes.

It should be noted that the integration of additional data sources, such as operational logs and environmental conditions, can significantly expand the capabilities of the model and improve the quality of forecasting. The analysis of such factors will make it possible to take into account more fully the complex interdependencies and dynamics of the equipment condition.

In the future, it is planned to expand the model to process several types of data, create a hierarchical Bayesian structure, and test it on real production data. This will improve the accuracy of forecasting and the adaptability of the system in a variety of operating conditions.

7. Conclusion

As part of this research, a prototype of a digital twin for industrial equipment was developed, employing a Bayesian method to assess and update the probability of failure in real time based on sensor data.

Simulation results demonstrated that the Bayesian model effectively adapts to incoming data: when failure signals are received, the estimated probability increases sharply, while under normal readings it decreases thus confirming the validity of the chosen approach.

Implementing the model in the R environment enabled visualization of changes in failure probability, which improved decision-making transparency and made the results more interpretable for engineering and technical personnel.

Limitations. The current version of the model uses a limited number of sensor inputs and assumes independence among observations. Additionally, the simulated data do not fully capture the complexity of real-world industrial conditions, which may limit the generalizability of the results. The absence of real industrial case studies at this stage also prevents a full assessment of the model's robustness to noise and incomplete data.

Future research directions. Future work will focus on extending the model to incorporate data from multiple sensor types, developing multidimensional Bayesian networks, and validating the approach using real-world industrial data. These enhancements are expected to improve the accuracy and generalizability of failure predictions.

The proposed approach is of high practical relevance, as it can be seamlessly integrated into existing industrial equipment monitoring systems and applied within the framework of predictive maintenance strategies.

References

[1] M. Attaran and B. G. Celik, "Digital twin: Benefits, use cases, challenges, and opportunities," *Decision Analytics Journal*, vol. 6, p. 100165, 2023. https://doi.org/10.1016/j.dajour.2023.100165

[2] M. Cheng, Z. Geng, and H. O. Gao, "Performance-oriented system digital twin: A Bayesian network-based framework to customize Sub-digital twins," *Structure and Infrastructure Engineering*, vol. 29, pp. 35-45, 2025. https://doi.org/10.1080/15732479.2025.2504535

[3] K. Lepenioti, A. Bousdekis, D. Apostolou, and G. Mentzas, "Prescriptive analytics: Literature review and research challenges," *International Journal of Information Management*, vol. 50, pp. 57-70, 2020. https://doi.org/10.1016/j.ijinfomgt.2019.04.003

[4] J. Yu, Y. Song, D. Tang, and J. Dai, "A digital twin approach based on nonparametric Bayesian network for complex system health monitoring," *Journal of Manufacturing Systems*, vol. 58, pp. 293-304, 2021. https://doi.org/10.1016/j.jmsy.2020.07.005

[5] K. Haddad, "A comprehensive review and application of bayesian methods in hydrological modelling: Past, present, and future directions," *Water (Switzerland)*, vol. 17, no. 7, p. 1095, 2025. https://doi.org/10.3390/w17071095

[6] S. Barbetta, G. Coccia, T. Moramarco, and E. Todini, "Real-time flood forecasting downstream river confluences using a Bayesian approach," *Journal of Hydrology*, vol. 565, pp. 516-523, 2018. https://doi.org/10.1016/j.jhydrol.2018.08.043

[7] M. Ganesh et al, "Digital Twin Application in Various Sectors," *Journal Transforming Industry using Digital Twin Technology*, vol. 35, pp. 219-237, 2024. https://doi:10.1007/978-3-031-58523-4_11

[8] T. Sun, X. He, and Z. Li, "Digital twin in healthcare: Recent updates and challenges," *DIGITAL HEALTH*, vol. 9, p. 20552076221149651, 2023. https://doi.org/10.1177/20552076221149651

[9] C. Cronrath, A. R. Aderiani, and B. Lennartson, "Enhancing digital twins through reinforcement learning," presented at the IEEE 15th International Conference on Automation Science and Engineering (CASE). IEEE, 2019, pp. 293–298, 2019.

[10] G. N. Schroeder, C. Steinmetz, C. E. Pereira, and D. B. Espindola, "Digital twin data modeling with automationml and a communication methodology for data exchange," *IFAC-PapersOnLine*, vol. 49, no. 30, pp. 12-17, 2016.

[11] A. Robertas and T. Iešmantas, "Application of Bayesian methods for age-dependent reliability analysis," *Quality and Reliability Engineering International*, vol. 30, no. 1, pp. 121-132, 2014. https://doi.org/10.48550/arXiv.1210.5095

[12] M.-W. Ho, "On bayes inference for a bathtub failure rate via S-paths," *Annals of the Institute of Statistical Mathematics*, vol. 63, no. 4, pp. 827-850, 2011. https://doi.org/10.1007/s10463-009-0253-1

[13] D. Pagano, "A predictive maintenance model using long short-Term Memory Neural Networks and Bayesian inference," *Decision Analytics Journal*, vol. 6, p. 100174, 2023. https://doi.org/10.1016/j.dajour.2023.100174

[14] M. N. Schmidt and M. Morup, "Nonparametric Bayesian modeling of complex networks: An introduction," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 110-128, 2013.

[15] M. Kraus and S. Feuerriegel, "Forecasting remaining useful life: Interpretable deep learning approach via variational Bayesian inferences," *Journal Machine Learning*, vol. 2, pp. 46-56, 2019. https://doi.org/10.48550/arXiv.1907.05146

[16] S. Weike, A. R. C. Paiva, P. Xu, A. Sundaram, and R. D. Braatz, "Fault detection and identification using Bayesian recurrent neural networks," *Computers & Chemical Engineering*, vol. 141, p. 106991, 2020. https://doi.org/10.1016/j.compchemeng.2020.106991

[17] C. Ruah, O. Simeone, and B. Al-Hashimi, "A bayesian framework for digital twin-based control, monitoring, and data collection in wireless systems," *Journal Electrical Engineering and Systems Science*, vol. 3, pp. 1-14, 2022. https://doi.org/10.48550/arXiv.2212.01351

[18] C. Li, S. Mahadevan, Y. Ling, S. Choze, and L. Wang, "Dynamic bayesian network for aircraft wing health monitoring digital twin," *Aiaa Journal*, vol. 55, no. 3, pp. 930-941, 2017. https://doi.org/10.2514/1.J055201

[19] E. Rabiei, E. L. Droguett, and M. Modarres, "A prognostics approach based on the evolution of damage precursors using dynamic Bayesian networks," *Advances in Mechanical Engineering*, vol. 8, no. 9, p. 1687814016666747, 2016. https://doi.org/10.1177/1687814016666747

[20] E. A. Micu, F. Asgari, M. Khouadjia, L. Connolly, and K. Zavitsas, "A probabilistic digital twin model for inland waterway transportation systems using Bayesian networks," in *CSuM 2024-Climate Crisis and Resilient Transportation Systems*, 2024.

[21] V. Karkaria *et al.*, "Towards a digital twin framework in additive manufacturing: Machine learning and bayesian optimization for time series process optimization," *Journal of Manufacturing Systems*, vol. 75, pp. 322-332, 2024. https://doi.org/10.1016/j.jmsy.2024.04.023

[22] A. Chong, W. Xu, S. Chao, and N.-T. Ngo, "Continuous-time Bayesian calibration of energy models using BIM and energy data," *Energy and Buildings*, vol. 194, pp. 177-190, 2019. https://doi.org/10.1016/j.enbuild.2019.04.017

[23] F. Tao, M. Zhang, Y. Liu, and A. Y. C. Nee, "Digital twin driven prognostics and health management for complex equipment," *CIRP Annals*, vol. 67, no. 1, pp. 169-172, 2018. https://doi.org/10.1016/j.cirp.2018.04.055
[24] H. Khaled, "A comprehensive review and application of bayesian methods in hydrological modelling: Past, present, and future directions," *Water*, vol. 17, no. 7, doi: 10.3390/w17071095.
[25] O. V. Sushko, Kolodiy, and O. S. Penov, "Individual forecasting of technical condition of machines and development of method for determining the conditional function of distributing their residual resource," *Journal Electrical Engineering*, vol. 10, no. 4, pp. 63-69, 2019.
[26] M. Prochnau, M. Holzbrink, W. Wang, M. Holters, J. Stollenwerk, and P. Loosen, "Measurement methods to build up the digital optical twin," in *Components and Packaging for Laser Systems IV*, 2018, vol. 10513: SPIE, pp. 247-256.
[27] R. Zhang and S. Song, "Bayesian network approach for dynamic fault tree with common cause failures and interval uncertainty parameters," *Eksploatacja i Niezawodność – Maintenance and Reliability*, vol. 26, no. 4, pp. 1-17, 2024 2024. https://doi.org/10.17531/ein/190379
[28] S. Park, S. Maliphol, J. Woo, and L. Fan, "Digital twins in industry 4.0," *Electronics*, vol. 13, no. 12, pp. 1-5doi: https://doi.org/10.3390/electronics13122258
[29] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for Industry 4.0-based manufacturing systems," *Manufacturing Letters*, vol. 3, pp. 18-23, 2015. https://doi.org/10.1016/j.mfglet.2014.12.001
[30] T. Akyaz and D. Engın, "Machine learning-based predictive maintenance system for artificial yarn machines," *IEEE Access*, vol. 12, pp. 125446-125461, 2024. https://doi.org/10.1109/ACCESS.2024.3454548
[31] N. Ungureanu, A. Petrovan, M. Ungureanu, and M. Alexandrescu, "Functioning condition monitoring of industrial equipment," in *IOP Conference Series: Materials Science and Engineering*, 2017, vol. 174, no. 1: IOP Publishing, p. 012068.
[32] A. Al-Hijazeen and K. Koris, "Implementation of digital twin and multiple linear regression in SHM of bridges: A case study," *SHM*, vol. 1, pp. 45-57, 2023.
[33] A. Isah, H. Shin, I. Aliyu, R. M. Sulaiman, and J. Kim, "Graph neural network for digital twin network: A conceptual framework," in *2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 19-22 Feb. 2024 2024, pp. 1-5.
[34] Y. Yang, Y. Wei, and T. Shen, "A review of graph neural networks for recommender applications," in *2021 IEEE International Conference on Unmanned Systems (ICUS)*, 2021: IEEE, pp. 602-607.
[35] R. Zhao *et al.*, "Decision tree based parameter identification and state estimation: Application to Reactor Operation Digital Twin," *Nuclear Engineering and Technology*, vol. 57, no. 7, p. 103527, 2025. https://doi.org/10.1016/j.net.2025.103527
[36] J.-Q. Zeng, H.-X. Zhang, H.-L. Gong, and Y.-T. Luo, "Ensemble Bayesian method for parameter distribution inference: application to reactor physics," *Nuclear Science and Techniques*, vol. 34, no. 12, pp. 199-209, 2023.
[37] Y. Liu, X. Sun, and N. T. Dinh, "Validation and uncertainty quantification of multiphase-CFD solvers: A data-driven Bayesian framework supported by high-resolution experiments," *Nuclear Engineering and Design*, vol. 354, p. 110200, 2019. https://doi.org/10.1016/j.nucengdes.2019.110200
[38] H. Zhou, S. Cheng, and R. Arcucci, "Multi-fidelity physics constrained neural networks for dynamical systems," *Computer Methods in Applied Mechanics and Engineering*, vol. 420, p. 116758, 2024. https://doi.org/10.1016/j.cma.2024.116758
[39] I. U. Hassan, K. Panduru, and J. Walsh, "Review of data processing methods used in predictive maintenance for next generation heavy machinery," *Data*, vol. 9, no. 5, pp. 1-38doi: https://doi.org/10.3390/data9050069.
[40] E. I. A. El-Latif, "Detection and identification drones using long short-term memory and Bayesian optimization," *Multimedia Tools and Applications*, vol. 84, no. 14, pp. 13983-13999, 2025. https://doi.org/10.1007/s11042-024-19469-1
[41] G. Fenza, M. Gallo, V. Loia, F. Orciuoli, and E. Herrera-Viedma, "Data set quality in machine learning: Consistency measure based on group decision making," *Applied Soft Computing*, vol. 106, p. 107366, 2021. https://doi.org/10.1016/j.asoc.2021.107366
[42] ISO 13379-1, "Condition monitoring and diagnostics of machines – Data processing," 2012.