Published online on the journal's webpage: **https://journals.telkomuniversity.ac.id/index.php/dinda**

Journal of Dinda

Data Science, Information Technology, and Data Analytics

Vol. 6 No. 2 (2026) 95 - 105 E-ISSN: 2809-8064

Bayesian Network and Croston Analysis for Axle Counter Maintenance at DAOP 7 Madiun

Hani'a Tsabita Fajriah Kansa[1*], Yuliani Puji Astuti[2]

[1,2*]Department of Data Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya

[1*]hania.23073@mhs.unesa.ac.id, [2]yuliani@unesa.ac.id

Abstract

Failures in railway signaling systems are critical issues affecting the safety and operational efficiency of train services. A key component is the axle counter, which detects train presence by counting its axles. This study aims to identify the dominant factors and causes of axle counter failures and predict future failure frequencies to support predictive maintenance. Data were obtained from signaling equipment failure records maintained by the Signaling and Telecommunications (Sintelis) Unit of PT Kereta Api Indonesia (Persero) DAOP 7 Madiun (January 2024–June 2026) and expert judgment obtained through a structured interview with the Assistant Manager of the Sintelis Unit. Failure factors and causes were analyzed using a Bayesian Network based on a Directed Acyclic Graph (DAG), Conditional Probability Tables (CPT), Joint Probability Distributions (JPD), and posterior probabilities derived from Bayes' Theorem. Components with the highest failure frequencies were then forecast using Croston's Method to predict intermittent failures. The results show that the Head component is mainly affected by external factors (71.43%), particularly metallic objects (59.26%). The Wheel Detection Equipment (WDE) is predominantly influenced by internal factors (66.67%), with deterioration-related damage as the dominant cause (52.70%). Similarly, the Evaluator Module is mainly influenced by internal factors (81.25%), with deterioration-related damage as the dominant cause (40.84%). Forecasting produced estimated failure frequencies of 0.11, 0.35, and 0.23 failures per monthly observation period for the Head, WDE, and Evaluator Module, respectively, indicating that the WDE is expected to experience the highest average failure frequency during the next monthly observation period. The integration of the Bayesian Network and Croston's Method supports predictive maintenance by prioritizing maintenance activities and improving component planning to enhance signaling system reliability.

Keywords: axle counter, Bayesian Network, Croston's Method, predictive maintenance, railway signaling.

1. Introduction

The railway signaling system functions to provide signals to the driver, consisting of various components such as mechanical switches, electrical switches, wire lines, level crossings, and power supplies. This system is useful as an indicator of the condition of the rail track ahead, whether the track is still occupied by another train or is safe to use [1]. This signaling system plays a crucial role in maintaining safety and regulating the smooth operation of train travel by controlling stopping, speed, and train crossings [2]. As the volume of train travel increases, signaling technology continues to evolve from manual signaling systems to automated systems that utilize sensors and computers. These developments aim to improve travel safety while reducing the risk of accidents so that train operations can run more efficiently [3].

One important device in the modern signaling system is the axle counter, a train detection system that works by counting the number of axles entering and leaving a detection track [4]. Once the number of outgoing axles equals the number of incoming axles, the track is declared clear again, allowing the next train to safely pass through the line. This system operates using sensors placed at two detection points so that each passing axle can be accurately counted. Therefore, the accuracy of axle calculations is crucial, as miscalculations can lead to operational disruptions and train failures.

The importance of analyzing failures in signaling systems has been widely discussed in various studies. Signaling devices can exhibit failure mechanisms characterized by gradual degradation over time, as components such as sensors, cables, and Evaluator Modules experience aging that can be affected by environmental factors [5]. However, most research still focuses on analyzing failures or risks in specific infrastructure, such as railways [6]. The application of Bayesian Network methods to handle uncertainty and the implementation of predictive analytics that integrate

Received: 27-07-2026 | Revised: 13-08-2026 | Published: 21-08-2026

asset data with real-time failure records on axle counters are still relatively limited. Axle counters are vulnerable to various failures, both from external factors, such as extreme weather, human error, and overcounting, and internal factors, such as component degradation.

To predict these failures, a probabilistic approach is required that can represent the relationships between variables and handle the uncertainty in failure data. Bayesian Networks were chosen because they can model probabilistic dependencies between variables and perform probabilistic inference under uncertainty based on available evidence [7]. In the context of axle counter systems, this method allows various factors causing disruptions, such as weather conditions, rail cleanliness, and component conditions, to be modeled in a Directed Acyclic Graph (DAG) so that the probabilistic dependencies among variables can be systematically analyzed [8]. Given the axle counter's crucial role in ensuring the safety and smooth operation of trains, this analysis was conducted to predict disruptions that could lead to train failures.

In this study, inference using Bayesian Networks was conducted in two stages. The first stage aimed to identify the dominant disruption factors in each component, namely internal or external factors. The results of this inference were then used as the basis for the second stage of analysis to determine the most dominant causes of disruptions based on actual historical incidents in the field. This stepwise approach is expected to produce more specific maintenance recommendations because it not only identifies the source category of disruptions but also determines the root cause of disruptions in each axle counter component.

In addition to analyzing the causes of disruptions, this study also applies time series forecasting using Croston's Method to predict the frequency of disruptions on priority components identified through a Bayesian Network. The Croston Method was chosen because it is capable of handling intermittent disruption data, or disruptions that occur irregularly with many uninterrupted periods. The prediction results are used to estimate the frequency of disruptions in future periods, thus providing a basis for planning maintenance activities and spare parts procurement.

By integrating Bayesian Network and Croston's Method, this study is expected to not only be able to identify the dominant factors and causes of disruptions, but also predict the frequency of disruptions that could potentially occur in the next period. The combination of these two methods provides a complementary approach, where Bayesian Network is used to analyze causal relationships and determine the most dominant causes of disruptions, while Croston's Method is used to estimate the frequency of disruption events based on historical patterns of intermittent disruptions. Thus, the results of this study are expected to be the basis for determining maintenance priorities, developing predictive maintenance strategies, and planning more precise and efficient component inventories, so that the reliability of the signaling system and the operational safety of train travel can be maintained.

2. Research Methods

This study used a quantitative approach to collect and analyze failure data on the axle counter signaling system at the Sintelis Unit of PT Kereta Api Indonesia (Persero) DAOP 7 Madiun. The quantitative approach was chosen because it provides objective information and allows for measurement of failure frequency. The study spanned five months, from January 2026 to June 2026.

2.1. *Research Stages*

Failure data on the axle counter signaling system were obtained from historical incident records documented in maintenance logbooks maintained by the Sintelis Unit of DAOP 7 Madiun over the period January 2024 to June 2026. The information collected included failure location, time of occurrence, duration, root cause, and the type of trackside equipment component affected. This historical data was then combined with structured interviews with one Assistant Manager of the Sintelis Unit to obtain expert judgment. Expert knowledge was elicited through structured interviews discussing historical failure records, operational experience, and the relationships between failure factors and axle counter components. The integration of historical data and expert knowledge was used to construct a Directed Acyclic Graph (DAG) structure in a Bayesian Network model based on probabilistic relationships between failure factors and axle counter components [9], [10].

The Directed Acyclic Graph (DAG) was manually developed based on the expert's knowledge of the axle counter system. The DAG structure reflects the operational characteristics of the system, in which each component may be influenced by two categories of failure factors. Internal factors represent failures originating from the component itself, such as deterioration, aging, or performance degradation. External factors represent failures caused by external conditions, including lightning, metallic objects, maintenance activities (MTT), and other environmental failures. Therefore, directed edges were assigned from both internal and external factor nodes to each axle counter component. The proposed DAG structure was subsequently reviewed and validated by the Assistant Manager to ensure that it accurately represented actual operational conditions.

Based on the Bayesian Network structure, each variable is represented in the form of a probability obtained from the failure data processing. These probability values are used to describe the level of influence of each failure factor on the axle counter components and as the basis for the probabilistic inference process to identify the most dominant failure factors and causes. The inference results are then used to determine priority components that require further attention in the prediction stage.

Next, the Bayesian Network analysis results are used as the basis for selecting components to be predicted using the time series method. However, because the failure data is intermittent, and not all components have a sufficient number of occurrences for forecasting, this study uses only the components with the highest failure frequency. These components are selected based on historical data, which is more representative than other components, so the prediction results are expected to be more accurate and reflect the failure pattern.

The forecasting process is conducted using Croston's Method, a time series forecasting method specifically designed for intermittent demand data, where failure events do not occur in every observation period. The forecasting results in the form of an estimate of the frequency of failures in the following period are then used as supporting information in preparing maintenance priorities and planning component inventory requirements in the axle counter system.

2.2. *Data Processing Method*

2.2.1. *Data Preparation and Cleaning*

The data preparation and cleaning phase was carried out to ensure that the data used in the study was of good quality, allowing for accurate analysis. The data that had undergone this process was then used as input for Bayesian Network analysis and forecasting using Croston's Method.

a. Historical Data Collection.

The data used in this study consisted of historical records of axle counter signaling device failures obtained from the Signal, Telecommunication, and Electricity Unit (Sintelis) of PT Kereta Api Indonesia (Persero) DAOP 7 Madiun. The dataset covered the period from January 2024 to June 2026 and included information on the incident date, affected component, failure factor, failure cause, and downtime duration. In addition to the historical records, interviews with Sintelis technicians and the Assistant Manager were conducted to provide expert judgment for the development of the Bayesian Network model. A total of 46 failure records were initially collected from the maintenance logbooks. During the data cleaning process, 12 records were removed because they were duplicate entries or did not contain clearly identified failure causes. Consequently, 34 validated failure records were retained for subsequent Bayesian Network analysis.

The cleaned dataset consisted of 34 validated failure records distributed across six axle counter components: Head (7), WDE (6), Evaluator Module (16), Modem (3), Cable Connector (1), and Button (1). The failure data exhibited intermittent characteristics, with many monthly observation periods containing no recorded failures. Over the 30 monthly observation periods (January 2024–June 2026), the Head, WDE, and Evaluator Module experienced 23, 24, and 14 zero-failure months, respectively. These three components were selected for forecasting because they had the highest failure frequencies among all components while still exhibiting substantial zero-failure periods. The large number of zero periods justifies the use of Croston's Method, which is specifically designed for forecasting intermittent events.

b. Data Cleaning

The data cleaning phase was conducted to improve data consistency and quality. This process includes removing duplicate data, improving date formats for uniformity, standardizing component naming, standardizing categories of factors and causes of disruptions, and removing incomplete or irrelevant data.

c. Data Transformation

The cleaned data is then transformed into a format suitable for the analysis needs. Transformation is performed by grouping data based on components, disruption factors (internal and external), disruption causes, and the period of occurrence.

2.2.2. *Bayesian Disruption Factor Analysis Stage*

a. Determining Parameters, Downtime, and Prior Probabilities

Each first stage begins with determining the parameters used in the Bayesian Network model. Each variable is represented in the form of probabilities obtained from historical disruption data and expert judgment. The downtime value for each disruption is used as the basis for weighting to describe the level of disruption impact on system reliability.

Next, a prior probability is calculated, indicating the initial probability of the disruption factor occurring before considering the evidence. The prior value is obtained based on the proportion of internal and external factors occurring in historical data. In this study, the prior value is obtained based on the accumulated calculation of historical disruption data based on influencing factors [11].

b. Evidence History Calculation Based on Components

Evidence history is actual historical data that describes component failure events. This data serves as evidence in the Bayesian Network inference process to indicate component conditions based on internal and external failure factors. The evidence value is obtained from the frequency of failure events recorded in the failure data from January 2024 to June 2026 and is used as input in the posterior probability calculation.

c. Compiling a Conditional Probability Table (CPT)

The Conditional Probability Table (CPT) was constructed using the historical frequency of failures for each axle counter component and refined using expert judgment from the Assistant Manager of the Sintelis Unit. The conditional probability of each component given a failure factor was calculated by dividing the number of failures associated with a particular factor by the total number of failures under that factor. Expert judgment was used to verify that the resulting probability distribution was consistent with actual operational conditions and maintenance experience.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \tag{1}$$

To ensure a valid Bayesian Network, the conditional probabilities in each CPT were normalized so that the probabilities associated with each parent node summed to one. Normalization was performed by dividing the frequency of each component by the total frequency of failures under the corresponding parent node. Consequently, the conditional probability distribution for each parent node satisfies (2):

$$\sum P(Component|Factor) = 1 \tag{2}$$

Example:

$$P(Head|External) = \frac{5}{10} = 0.50$$

$$P(Head|Internal) = \frac{2}{24} = 0.083$$

d. Calculation of Joint Probability Distribution (JPD) and Posterior Probability

Joint Probability Distribution (JPD) is used to represent the joint probability of two or more variables occurring simultaneously in a Bayesian Network model. The JPD value is obtained by multiplying the prior probability by the conditional probability according to the relationships between the variables established in the DAG structure. The posterior probability of each failure factor was subsequently calculated using Bayes' Theorem, as expressed in Equation (3).

$$P(A \cap B) = P(A) \times P(B|A) \tag{3}$$

For the Head component,

Prior:

$$P(Eksternal) = \frac{10}{34} = 0.294$$

Evidence:

$$P(\,Head \mid External\,) = 0.50$$

Joint probability:

$$P(Head, External) = 0.50 \times 0.294 = 0.147$$

Total evidence:

$$P(Head) = 0.147 + 0.059 = 0.206$$

Posterior:

$$P(External \mid Head) = 0.2060.147 = 0.714$$

e. Probabilistic Inference Calculation

Probabilistic inference is the final stage in Bayesian Network analysis, which aims to determine the most dominant failure factor based on posterior probability values. In this study, the inference process is carried out through a diagnostic reasoning mechanism by comparing the posterior values of external and internal factors for each component. The factor with the highest posterior value is determined as the most dominant cause of the failure.

2.2.3. *Bayesian Phase II Failure Cause Analysis*

The second stage is carried out using the same procedure as Phase I, but the variables analyzed are the failure causes. Prior probabilities are compiled based on the historical frequency of each failure cause, while historical evidence is formed based on the relationship between the failure cause and the axle counter component. Next, a Conditional Probability Table (CPT) is compiled, Joint Probability Distribution (JPD) is calculated, posterior probabilities are calculated, and probabilistic inference is used to determine the most dominant failure cause for each component.

2.2.4. *Time Series Calculation Using Croston's Method*

After obtaining the priority components based on the Bayesian Network analysis results, the next stage is predicting the failure frequency using Croston's Method. This method was chosen because the axle counter component failure data has intermittent characteristics, namely the failure does not occur in every observation period so that there are many periods without events (zero demand). In contrast to conventional time series methods, Croston's Method performs a smoothing process (exponential smoothing) separately for the amount of demand (demand) and the interval between requests (inter-demand interval) [12], [13].

a. Calculating the Time Period

The first step is to compile historical disruption data ased on the chronological order of occurrence. The data is then grouped into monthly periods to determine the number of disruptions that occurred in each period. This value is used as the demand size in the Croston's Method calculation process.

b. Calculating the Interval Between Demands

Next, the interval between demands is calculated, indicating the time interval between consecutive disruption events. The interval value is calculated based on the number of periods from the occurrence of the previous disruption to the occurrence of the next disruption. This interval is used to describe the pattern of irregular (intermittent) disruption occurrences.

c. Determining the Smoothing Constant

The smoothing constant (α) is a parameter that determines the influence of recent observations on the forecasting results. Candidate α values ranging from 0.1 to 0.9 were evaluated using a backtesting procedure. For each candidate value, Croston's Method was applied to the historical failure data, and forecasting accuracy was evaluated using the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The α value with the lowest forecasting error was selected for the final forecasting model.

d. Performing Croston's Calculation

The calculation is performed by separating the smoothing process into the demand magnitude and the interval between demands [14]. The smoothed demand value is calculated using equations 4 and 5.

$$S_t = \alpha X_t + (1-\alpha) S_{t-1} \tag{4}$$

$$I_t = \alpha q + (1-\alpha) I_{t-1} \tag{5}$$

e. Forecasting

After obtaining the smoothed demand and smoothed interval values, the forecast value is calculated using equation 6.

$$M_t = \frac{S_t}{I_t} \tag{6}$$

The final forecast value represents the expected number of failures per monthly observation period, obtained as the ratio between the smoothed demand and the smoothed inter-demand interval. These forecast values are used to estimate the average failure frequency in the next month and to support maintenance prioritization and spare parts planning.

3. Results and Discussion

3.1. *Constructing the Bayesian Network Structure*

Based on historical data of failures for the period January 2024 – June 2026 and the results of expert judgment of Sintelis technicians, this structure is designed to represent probabilistic dependencies among failure factors, failure causes, and axle counter components, which serve as the basis for probabilistic inference [15], [16]. The node consists of failure factors, namely external and internal as parent nodes that affect component nodes, namely Head, WDE, Evaluator Module, Modem, Transmission Cable and several other axle counter components. This relationship shows that each component can experience failures influenced by external and internal factors.

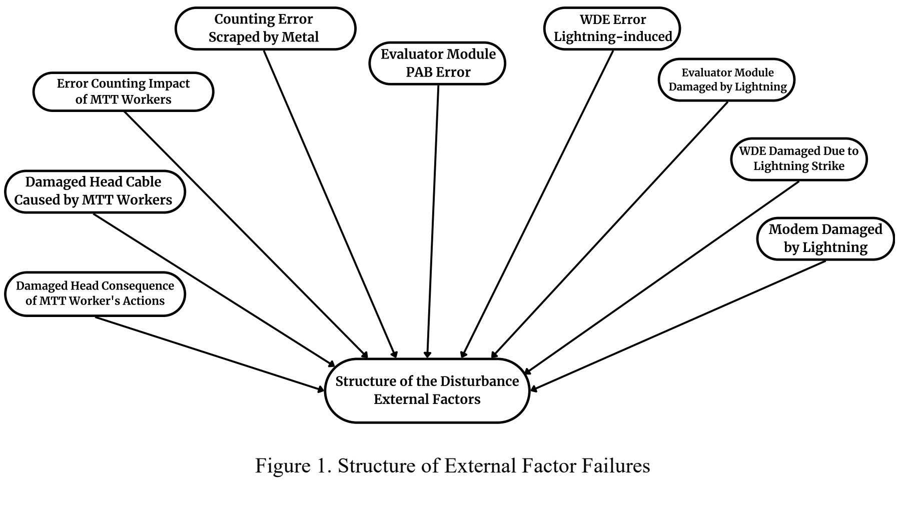

Figure 1. Structure of External Factor Failures

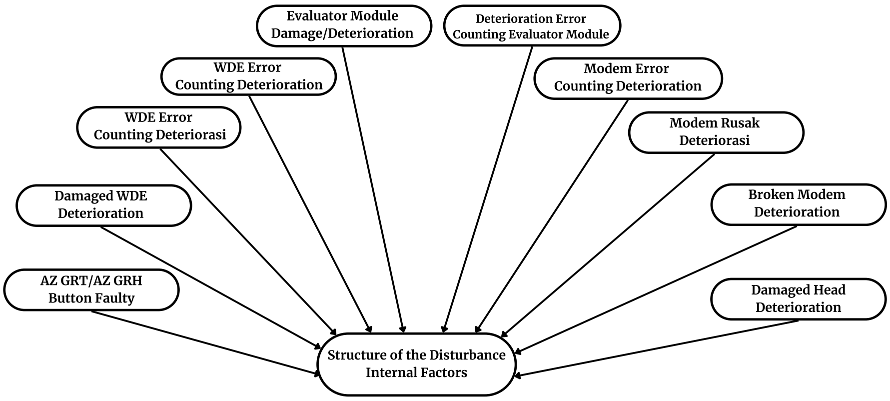

Figure 2. Structure of Internal Factor Failures

3.2. *Determining Downtime Parameters and Weighting*

Each disruption represented within the Bayesian Network structure has a parameter value representing downtime. These weights are derived from field data and interviews with Sintelis Unit staff. The parameter determination process begins by assigning an impact weight to each type of disruption based on the extent of its influence on the resulting downtime.

Table 1. Classification of Weight Ranges Based on Downtime

Name of Failure	Impact Weight
WDE Damaged Due to Lightning	0.70
Head Damaged Due to MTT Work	0.80
Evaluator Module Damaged Due to Lightning	0.60
Modem Damaged Due to Lightning	0.60
Head Cable Damaged Due to MTT Work	0.80
WDE Error Due to Lightning	0.70
Evaluator Module Error Due to Level Crossing (PAB)	0.60
Counting Error Due to MTT Work	0.30
Counting Error Due to Metallic Objects	0.30
Modem Damaged Due to Lightning	0.60
Transmission Cable Damaged	0.90
AZ GRT/AZ GRH Button Damaged	0.30
DB-9 Cable Connector Damaged	0.50
Modem Counting Error Due to Deterioration	0.50
WDE Counting Error Due to Deterioration	0.70
Evaluator Module Counting Error Due to Deterioration	0.60
WDE Damaged Due to Deterioration	0.70
Modem Damaged Due to Deterioration	0.60
Head Damaged Due to Deterioration	0.80
Evaluator Module Damaged Due to Deterioration Transmission Cable Damaged	0.40
Evaluator Module Damaged Due to Deterioration Transmission Cable Damaged	0.90

3.3. *Determination of Prior Probability*

Prior probability is calculated based on the frequency proportions of internal and external disruption factors in historical data, serving as an initial probability before considering evidence. In this study, the prior value is derived from accumulated calculations of historical disruption data, categorized by the influencing factors. The prior probability serves as the foundation for the inference process within the Bayesian Network model and is updated once evidence of actual disruptions is incorporated into the model. The results of the prior probability calculations for each factor group are presented in Table 2.

Table 2. Prior Probabilities of External and Internal Factors

Component	Number of Failures	Frequency
External Factors	10	0.294
Internal Factors	24	0.706

3.4. *Evidence History*

Evidence history represents the record of actual failures occurring in the components of the axle counter signaling system. The evidence values indicate the sensitivity level of each component to failures arising from either external or internal factors. The evidence history results for each component are presented in Table 3.

Table 3. Axle Counter Component Evidence History

Component	External Factors	Internal Factors	External Prior	Internal Prior
Head	5	2	0.50	0.08
WDE	2	4	0.20	0.17
Cable Connector	0	1	0.00	0.04
Evaluator Module	3	13	0.30	0.54
Modem	0	3	0.00	0.13
Button	0	1	0.00	0.04

By integrating the prior probability value and evidence history using Bayes' Theorem, the posterior probability value is obtained. This value represents the updated probability after incorporating evidence of field-based failures into the Bayesian Network model. The calculated posterior probability results for each component are presented in Table 4.

Table 4. Posterior Probability Calculation Results

Component	Calculation Result	External	Internal
Head	0.206	0.147	0.059
WDE	0.176	0.059	0.118
Cable Connector	0.029	0.000	0.029
Evaluator Module	0.471	0.088	0.382
Modem	0.088	0.000	0.088
Button	0.029	0.000	0.029

3.5. *Probabilistic Inference*

The probabilistic inference stage is the core of the Bayesian Network analysis used to determine the most dominant cause of failure for each component. Handling priorities are established by comparing the posterior probability values of external and internal factors. The factor with the highest probability value is identified as the dominant state. The results of this inference are presented in Table 5.

Table 5. Probabilistic Inference Calculation Results

Component	External	Internal	Dominant	Percent
Head	0.7143	0.2857	External	71.43%
WDE	0.3333	0.6667	Internal	66.67%
Cable Connector	0.0000	1.0000	Internal	100.00%
Evaluator Module	0.1875	0.8125	Internal	81.25%
Modem	0.0000	1.0000	Internal	100.00%
Button	0.0000	1.0000	Internal	100.00%

3.6. *Prior Probabilities of Failure Causes Based on Bayesian Stage II*

The prior probabilities were determined based on the historical frequencies of each failure cause within the internal and external factor groups, using the same procedure as in the first stage. The results of the prior probability calculations are presented in Tables 6 and 7.

Table 6. Prior Probability of External Failure Causes

External Factor	Number of Occurrences	Prior
Lightning	5	0.50
Metallic Objects	4	0.40
MTT Work	1	0.10

Table 7. Prior Probability of Internal Failure Causes

Internal Factor	Number of Occurrences	Prior
Deterioration Damage	13	0.54
Deterioration Error	11	0.46

These values indicate the prior probability of each cause of the failure before considering evidence from each component.

3.7. *Evidence History of Failure Causes*

The next step involves compiling an evidence history based on the historical distribution of failure causes for each axle counter component. Evidence values are calculated from the proportion of each cause's occurrence relative to the total number of failures affecting the respective component. This evidence history reflects actual field conditions, thereby revealing failure characteristics specific to each component. The results of the evidence history are presented in Table 8.

Table 8. Dominant Failure Causes Based on Posterior Probability

Component	Lightning	Metallic Objects	MTT Work	Damage	Error
Head	0	0.57	0.14	0.14	0.14
WDE	0.33	0	0	0.5	0.17
Cable Connector	0	0	0	1	0
Evaluator Module	0.19	0	0	0.38	0.44
Modem	0	0	0	0.33	0.67
Button	0	0	0	1	0

The results of the evidence history serve as the basis for calculating posterior probabilities in the second stage to identify the most dominant failure cause for each axle counter component.

3.8. *Posterior Probability Inference*

Based on prior probabilities and evidence history, an inference process using a Bayesian Network model is performed to obtain posterior probability values for each failure cause. The posterior probability value indicates the level of confidence regarding each failure cause after considering historical data for each component. Subsequently, the dominant state is determined based on the highest posterior probability value for each component. The dominant state represents the failure cause with the highest level of confidence after accounting for prior probabilities and evidence history. The results of the dominant state determination are presented in Table 9.

Table 9. Dominant Failure Causes Based on Posterior Probability

Component	Dominant Failure	**Percent**
Head	Metallic Objects	59.26%
WDE	Damage Due to Deterioration	52.70%
Cable Connector	Damage Due to Deterioration	100.00%
Evaluator Module	Damage Due to Deterioration	40.84%
Modem	Error Due to Deterioration	62.86%
Button	Damage Due to Deterioration	100.00%

The inference results in Table 9 indicate that the dominant cause of failures varies across the components of the axle counter system.

Table 10. Summary of Bayesian Network Inference Results

Component	Dominant Factor	Dominant Cause
Head	External 71.43%	Metallic Objects 59.26%
WDE	Internal 66.67%	Damage Due to Deterioration 52.70%
Cable Connector	Internal 100.00%	Damage Due to Deterioration 100.00%
Evaluator Module	Internal 81.25%	Damage Due to Deterioration 40.84%
Modem	Internal 100.00%	Error Due to Deterioration 62.86%
Button	Internal 100.00%	Damage Due to Deterioration 100.00%

For the Evaluator Module, the probability of deterioration-related damage is 40.84%, only slightly higher than the 40.31% probability of deterioration-related errors. This indicates that both types of failures contribute almost equally to the failure of the Evaluator Module. Meanwhile, the Modem exhibits the highest probability for deterioration-related errors at 62.86%, leading to the conclusion that performance degradation is a more dominant factor than physical component damage. To obtain more comprehensive maintenance recommendations, the results from the first and second stages of the Bayesian Network modeling were integrated, as presented in Table 10.

3.9. *Time Series Calculation Using Croston's Method*

After identifying priority components based on the Bayesian Network analysis results, failure frequency forecasting was conducted using Croston's Method. This method was selected because the failure data exhibits an intermittent demand pattern meaning failures do not occur in every observation period. In this study, forecasting was performed for the three components with the highest failure frequencies: the Head, Wheel Detection Equipment (WDE), and the Evaluator Module.

3.10. *Determination of Intervals and Demand*

The interval between demands is calculated based on the time elapsed between two consecutive failure events to characterize the failure occurrence pattern. Calculations for the intervals between demands are presented in Tables 11, 12, and 13.

Table 11. Intervals Between Demands for the Head Component

Demand	Interval
1	1
1	2
2	7
1	3
1	3
1	10

Table 12. Intervals Between Demands for WDE Components

Demand	Interval
1	1
1	1
1	3
1	7
1	1
1	3

Table 13. Intervals Between Demands for Evaluator Module Components

Demand	Interval
1	1
1	1
4	2
1	3
1	1
1	3
2	1
1	5
2	1
1	8
1	4

3.11. *Determination of the α Value*

The smoothing constant α is determined as a parameter in the exponential smoothing process. In this study, a value of α = 0.9 is used.

3.12. *Forecast Validation and Selection of α*

To determine the most appropriate smoothing constant, Croston's Method was evaluated using α values ranging from 0.1 to 0.9. A backtesting procedure was performed on the historical failure data, and forecasting performance was assessed using the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The evaluation results are presented in Table 14.

Table 14. Forecast Validation Results for Different α Values Using Croston's Method

Standart Deviasi	MAE	RMSE
0,1	0,43	0,43
0,2	0,35	0,39
0,3	0,31	0,38
0,4	0,29	0,37
0,5	0,27	0,37
0,6	0,26	0,37
0,7	0,25	0,37
0,8	0,25	0,38
0,9	0,24	0,38

Based on the evaluation results, $\alpha = 0.9$ produced the lowest MAE (0.24), while maintaining a low RMSE (0.38). Therefore, $\alpha = 0.9$ was selected as the smoothing constant for the final Croston forecasting model.

To further evaluate forecasting performance, Croston's Method was compared with the Naïve forecasting method, which uses the previous observation as the forecast for the next period. The comparison results are presented in Table 15.

Table 15. Forecast Performance Comparison Between Croston's and Naïve Methods

Method	MAE	RMSE
Naïve	0.48	0.79
Croston	0.24	0.38

The comparison indicates that Croston's Method outperformed the Naïve forecasting method. Croston reduced the MAE from 0.48 to 0.24 and the RMSE from 0.79 to 0.38. This result confirms that Croston's Method is more suitable for forecasting intermittent failure data because it separately models failure occurrence and inter-failure intervals.

3.13. *Croston's Method Calculation*

Calculations are performed by separately smoothing the demand values and the intervals between demands using the Croston's Method equations. The calculations for demand and intervals between demands using Croston's Method are presented in Tables 16, 17, and 18.

Table 16. Croston's Method Calculation for the Head Component

Period	Demand (S)	Interval (I)	S	I
February 2024	1	1	1,00	1,00
April 2024	1	2	1,00	1,90
November 2024	2	7	1,90	6,49
February 2025	1	3	1,09	3,35
May 2025	1	3	1,01	3,03
February 2026	1	10	1,00	9,30

Table 17. Inter-demand Interval for WDE Components

Period	Demand (S)	Interval (I)	S	I
February 2024	1	1	1,00	1,00
March 2024	1	1	1,00	1,00
June 2024	1	3	1,00	2,80
January 2025	1	7	1,00	6,58
February 2025	1	1	1,00	1,56
May 2025	1	3	1,00	2,86

Table 18. Inter-demand Interval for Evaluator Module Components

Period	Demand (S)	Interval (I)	S	I
January 2024	1	1	1,00	1,00
February 2024	1	1	1,00	1,00
April 2024	4	2	3,70	1,90
July 2024	1	3	1,27	2,89
August 2024	1	1	1,03	1,19
November 2024	1	3	1,00	2,82
December 2024	2	1	1,90	1,18
May 2025	1	5	1,09	4,62
June 2025	2	1	1,91	1,36
February 2026	1	8	1,09	7,34
June 2026	1	4	1,01	4,33

3.14. *Forecasting*

Forecast values were calculated as the ratio of the smoothed demand to the smoothed inter-demand interval using Croston's Method. The resulting values represent the expected average number of failures per monthly observation period rather than the probability of failure. These estimates are used to support predictive maintenance scheduling and spare parts planning. The forecast values are presented in Table 19.

Table 19. Forecasted Failure Frequency per Monthly Observation Period

Component	Forecasted Failure Frequency (failures/month)
Head	0.11
WDE	0.35
Evaluator Module	0.23

Subsequently, the forecasting results are integrated with the Bayesian Network inference results to derive more comprehensive maintenance priorities. The Bayesian Network is used to identify the most dominant factors and causes of disruptions, while Croston's Method is employed to estimate the frequency of disruptions in the upcoming period. The results of integrating these two methods are presented in Table 20.

Based on the integration results, the Head component is predominantly affected by external factors primarily involving metal objects whereas the WDE and Evaluator Module components are mainly influenced by internal factors, primarily due to deterioration. Among the analyzed components, the WDE exhibits the highest forecasted failure frequency per monthly observation period, followed by the Evaluator Module and the Head. This indicates that, on average, the WDE is expected to experience failures more frequently than the other components in future monthly observation periods.

Table 20. Results of Integrating the Bayesian Network and Croston's Method

Component	Dominant Factor	Dominant Cause	Forecasting
Head	External	Metallic Objects	0.11
WDE	Internal	Damage Due to Deterioration	0.35
Evaluator Module	Internal	Damage Due to Deterioration	0.23

Overall, the study demonstrates that integrating the Bayesian Network and Croston's Method provides more comprehensive information to support predictive maintenance. The Bayesian Network identifies the dominant factors and causes of failure, while Croston's Method estimates the frequency of failures in future periods. Consequently, maintenance recommendations are not based solely on potential failure causes but also account for the likelihood of failure occurrence for each component, thereby facilitating more effective prioritization of inspections, component replacements, and spare parts planning.

4. Conclusion

Based on the analysis using Bayesian Networks and Croston's Method, it can be concluded that each axle counter component exhibits distinct failure characteristics. The Head component is predominantly affected by external factors, with a posterior probability of 71.43%; the most probable failure cause inferred by the Bayesian Network is metallic objects, with a probability of 59.26%. In contrast, the Wheel Detection Equipment (WDE) and Evaluator Module components are primarily influenced by internal factors, showing posterior probabilities of 66.67% and 81.25%, respectively. For both of these components, the dominant inferred failure cause is deterioration-related damage, with posterior probabilities of 52.70% for the WDE and 40.84% for the Evaluator Module.

Forecasting results using Croston's Method indicate relatively low estimated failure frequencies for the upcoming period: 0.11 for the Head component, 0.35 for the WDE, and 0.23 for the Evaluator Module. Although the forecast values are below one incident per period, the WDE component shows the highest forecasted failure frequency per monthly observation period, indicating that it is expected to experience failures more frequently than the Head and Evaluator Module in future observations.

Overall, integrating Bayesian Networks and Croston's Method provides comprehensive information to support predictive maintenance. Bayesian Networks were used to identify the dominant failure factors and infer the most probable failure causes, while Croston's Method estimates failure frequencies for the next period. Combining these methods helps establish maintenance priorities, enhances the effectiveness of component inspections, and supports spare parts planning, thereby ensuring the continued reliability of the axle counter signaling system at PT Kereta Api Indonesia (Persero) DAOP 7 Madiun.

Acknowledgments

The authors wish to express their sincere gratitude to PT Kereta Api Indonesia (Persero) DAOP 7 Madiun—specifically the Signals, Telecommunications, and Electrical (Sintelis) Unit for providing historical maintenance data, technical guidance, and valuable insights throughout this research. The authors also extend their thanks to the Assistant Manager and Sintelis technicians for their expert assessments and extensive discussions, which contributed significantly to the development and validation of the Bayesian Network model.

References

[1] Wibawanto, B.S., Hapsari, J.P., Suprajitno, A. and Arifianto, T., 2022. Analysis of Railway Signaling Equipment Using the Silsafe4000 Electronic Signaling System at Lempuyangan Station, Yogyakarta. Indonesian Railway Journal, 6(2), pp.42–48.

[2] Aptana, Arsyad, M.F. and Prasetyo, P.H., 2023. Comparative Analysis of Conventional Signaling Systems and CBTC on Railway Headway. Journal of Transportation Engineering, 2(1), pp.68–85.

[3] Rosyidi, M. and Santoso, D.B., 2023. Analysis of Axle Counter System Failures in the Railway Network at PT Kereta Api Indonesia (Persero), Sintel Resort 1.16 Karawang, Operational Area 1 Jakarta. JATI (Journal of Informatics Engineering Students), 7(2), pp.1317–1321.

[4] Nugraha, G.W.S., Sunardi and Arifianto, T., 2023. Development of an Axle Counter Grounding System Based on Soil Resistivity and Grounding Depth. Journal of Telecommunication, Electronics, and Control Engineering (JTECE), 5(1), pp.1–13.

[5] Nguyen, V. A., 2023. Predictive Maintenance for Axle Counters of Urban Railway Train Control Signaling Systems. Hanoi, Vietnam: Hanoi University.

[6] Putri, V.N.A., Usman, K., Kustiani, I. and Siregar, A.M., 2024. Risk Analysis of Rail Fracture Causes on Railway Tracks in Divre IV Tanjung Karang. Journal of Civil and Environmental Engineering, 9(2), pp.283–292.

[7] Sari, I.P., Elvitaria, L. and Yoelanda, I., 2024. Bayesian Network Method for Determining the Probability of Bipolar Disorder Indications. JEKIN – Journal of Informatics Engineering, 4(2), pp.196–208.

[8] Fajri, M.N. and Handayani, T., 2024. Posterior Probability Analysis Using the Bayesian Network Method for Predictive Maintenance of Industrial Assets. Indonesian Journal of Data Science (JSDI), 4(1), pp.15–26.

[9] Siregar, A.M. and Adam, R., 2020. Quantitative Approach Analysis in Measuring the Reliability of Public Transportation Systems. Journal of Information Systems and Technology, 2(3), pp.85–92.

[10] Hartati, S. and Nugroho, A., 2021. Probabilistic Computation Using the Bayesian Network Model for an Expert System in Machine Failure Diagnosis. Journal of Informatics and Computer (JIKO), 6(2), pp.121–129.

[11] Prasetyo, B. and Utami, S.S., 2023. Probabilistic Inference Based on Bayes' Theorem for Early Detection of Railway Electronic Component Failures. Indonesian Railway Technology Journal, 7(2), pp.202–211.

[12] Nuraini, S.S., Anugrah, I.G. and Witra, W.P.P., 2026. Pengendalian Persediaan Intermittent Demand Berbasis FSN dan Croston-TSB. INTECOMS: Journal of Information Technology and Computer Science, 9(3), pp.543–557.

[13] Xu, Q., Wang, N. and Shi, H., 2012. Review of Croston's Method for Intermittent Demand Forecasting. Proceedings of the 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pp.1456–1460.

[14] Febrianto, D.T. and Bertuah, E., 2024. Perancangan Model Peramalan Permintaan dalam Menentukan Persediaan Paling Ekonomis. Jurnal Cahaya Mandalika, 5(2), pp.1125–1139.

[15] Hasniati, Arianti and Philip, W., 2018. Application of the Bayesian Network Model to Calculate the Probability of Infant Dyspnea. Journal of Information Technology Engineering (JURTI), 2(1), pp.62–68.

[16] Ramadan, T. and Setiawan, B., 2019. Application of the Bayesian Network Model to Calculate the Probability of Infant Dyspnea. Journal of Informatics Education and Research (JEPIN), 5(2), pp.174–181.