v7.n11.2026

SCIENTIFIC JOURNAL - RECIMA21 ISSN 2675-6218

BAYESIAN MODELLING OF SUBWAY TRAIN DOOR FAILURES: RELIABILITY ESTIMATES AND OPERATIONAL IMPACT SCENARIOS

MODELAGEM BAYESIANA DE FALHAS EM PORTAS DE TRENS METROVIÁRIOS: ESTIMATIVAS DE CONFIABILIDADE E CENÁRIOS DE IMPACTO OPERACIONAL

MODELADO BAYESIANO DE FALLOS EN PUERTAS DE TRENES METROPOLITANOS: ESTIMACIONES DE CONFIABILIDAD Y ESCENARIOS DE IMPACTO OPERATIVO

Eduardo Augusto Campos[1], Adriano de Almeida Benassi[2], Luciano Guilherme da Costa Olegário[3], Wolfgang Ludwig Farias Schmidt[4], Vitor Silva Ribas de Oliveira[3], Marcos Rogério Souza da Silva[4], Jurandir Oliveira Rodrigues dos Santos[4], Fabio Cosme Rodrigues dos Santos[5], Edson Melo de Souza[6], André Felipe Henriques Librantz[7]

e7119458

https://doi.org/10.47820/recima21.v7i11.9458

PUBLISHED: 11/2026

ABSTRACT

Subway maintenance requires tools that relate component reliability to the consequences of operational unavailability. This study developed and applied a Bayesian network framework integrating failure estimates, maintenance costs and scenario-based operational impacts for 22 door-system components of one subway operator. Three-parameter Weibull distributions fitted to maintenance records were combined with operational interference probabilities elicited from specialists, used as conditional inputs whose predictive capacity was not validated. Expected failure counts were compared retrospectively with 2019 records. The model estimated 822.6 failures against 940 recorded, an aggregate underestimation of 12.5%, with a mean absolute error of 11.93 failures and a weighted absolute percentage error of 27.9%; 14 of the 22 observed counts fell within 95% Poisson prediction intervals, and dispersion exceeded the Poisson assumption; a previous-year forecast reached a WAPE of 36.5% and a three-year mean, 27.0%. Expected canceled trips per failure, recovered from the implemented network, ranged from 0.87 for depot detections to 6.06 for severe failures detected on the commercial route. Under joint perturbations of up to 50% in these values, the ordering of interventions by investment per avoided trip remained unchanged, whereas the comparison between modernization and a combined lubrication and depot-inspection intervention was not robust. Canceled trips and the monetary proxy were not validated against observed disruptions. The contribution is an exploratory decision-support structure whose managerial use depends on documented parameters and prospective comparison with operational outcomes.

[1] Bachelor's Degree in Electrical Engineering, Master's Degree in Production Engineering, Graduate Program in Informatics and Knowledge Management (PPGI), Nove de Julho University (Uninove), São Paulo, São Paulo, Brazil.

[2] Bachelor's Degree in Control and Automation Engineering, Master's Degree in Production Engineering, PhD candidate in Informatics and Knowledge Management, Graduate Program in Informatics and Knowledge Management (PPGI), Nove de Julho University (Uninove), São Paulo, São Paulo, Brazil.

[3] Bachelor's Degree in Systems Analysis and Development, Master's candidate in Informatics and Knowledge Management, Graduate Program in Informatics and Knowledge Management (PPGI), Nove de Julho University (Uninove), São Paulo, São Paulo, Brazil.

[4] Bachelor's Degree in Computer Science, Master's candidate in Informatics and Knowledge Management, Graduate Program in Informatics and Knowledge Management (PPGI), Nove de Julho University (Uninove), São Paulo, São Paulo, Brazil.

[5] Bachelor's Degree in Electrical Engineering, Bachelor's Degree in Computer Science, PhD in Informatics and Knowledge Management, Cia de Saneamento Básico do Estado de São Paulo (SABESP), São Paulo, São Paulo, Brazil.

[6] Bachelor's Degree in Computer Science, Bachelor's Degree in Production Engineering, PhD in Informatics and Knowledge Management, working in Artificial Intelligence, Intelligent Systems, Decision Making and Knowledge Management, Graduate Program in Informatics and Knowledge Management (PPGI), Nove de Julho University (Uninove), São Paulo, São Paulo, Brazil.

[7] Bachelor's Degree in Physics, PhD in Nuclear Technology, with a solid research record in quantitative methods, especially modeling, computer simulation and multi-criteria decision-making (MCDM) methods, Graduate Program in Informatics and Knowledge Management (PPGI), Nove de Julho University (Uninove), São Paulo, São Paulo, Brazil.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

KEYWORDS: *Bayesian networks. Weibull distribution. Railway maintenance. Reliability. Subway. Decision support.*

RESUMO

A manutenção de sistemas metroviários demanda instrumentos que relacionem a confiabilidade dos componentes às consequências da indisponibilidade operacional. Este estudo desenvolveu e aplicou um *framework* de redes bayesianas que integra estimativas de falhas, custos de manutenção e impactos operacionais baseados em cenários para 22 componentes do sistema de portas de uma operadora metroviária. Distribuições Weibull ajustadas a registros de manutenção foram combinadas com probabilidades de interferência operacional elicitadas de especialistas, cuja capacidade preditiva não foi validada. Comparado retrospectivamente aos registros de 2019, o modelo estimou 822,6 falhas, frente a 940 registradas, uma subestimação agregada de 12,5%, com erro absoluto médio de 11,93 falhas e erro percentual absoluto ponderado de 27,9%. Apenas 14 das 22 contagens situaram-se nos intervalos de predição de Poisson de 95%, e a dispersão excedeu esse pressuposto. O desempenho superou a previsão pelo ano anterior (36,5%) e equiparou-se à média de três anos (27,0%). As viagens canceladas esperadas por falha variaram de 0,87, para detecções em pátio, a 6,06, para falhas severas na via comercial. Sob perturbações conjuntas de até 50% nesses valores, a ordenação das intervenções por investimento por viagem evitada permaneceu inalterada, enquanto a comparação entre modernização e a combinação de lubrificação e inspeção em pátio não se mostrou robusta. Viagens canceladas e indicador monetário não foram validados contra interrupções observadas. A contribuição consiste em uma estrutura exploratória de apoio à decisão, cujo uso gerencial depende de parâmetros documentados e de comparação prospectiva com resultados operacionais.

PALAVRAS-CHAVE: Redes bayesianas. Distribuição Weibull. Manutenção ferroviária. Confiabilidade. Metrô. Apoio à decisão.

RESUMEN

El mantenimiento de sistemas de metro requiere instrumentos que relacionen la confiabilidad de los componentes con las consecuencias de la indisponibilidad operativa. Este estudio desarrolló un marco de redes bayesianas que integra estimaciones de fallos, costos de mantenimiento e impactos operativos basados en escenarios para 22 componentes del sistema de puertas de una operadora de metro. Distribuciones Weibull ajustadas a registros de mantenimiento se combinaron con probabilidades de interferencia operativa obtenidas de especialistas, sin validación predictiva. Frente a los registros de 2019, el modelo estimó 822,6 fallos contra 940, una subestimación agregada del 12,5%, con error absoluto medio de 11,93 fallos y error porcentual absoluto ponderado del 27,9%. Solo 14 de los 22 recuentos quedaron dentro de los intervalos de predicción de Poisson del 95%, y la dispersión superó ese supuesto. El desempeño superó la previsión por el año anterior (36,5%) y se equiparó a la media de tres años (27,0%). Los viajes cancelados esperados por fallo variaron entre 0,87, para detecciones en patio, y 6,06, para fallos graves en la vía comercial. Con perturbaciones conjuntas de hasta el 50% en estos valores, el orden de las intervenciones por inversión por viaje evitado se mantuvo, mientras que la comparación entre modernización y la combinación de lubricación e inspección en patio no fue robusta. Viajes cancelados e indicador monetario no se validaron con interrupciones observadas. La contribución es una estructura exploratoria de apoyo a la decisión, cuyo uso gerencial depende de parámetros documentados y de comparación prospectiva con resultados operativos.

PALABRAS CLAVE: *Redes bayesianas. Distribución Weibull. Mantenimiento ferroviario. Confiabilidad. Metro. Apoyo a la decisión.*

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

INTRODUCTION

Urban rail transport supports mobility and economic activity, and the continuity of its services depends on technical assets. Studies of subway systems show that infrastructure benefits and operational reliability must be considered together (Haddad *et al.*, 2015; Sun *et al.*, 2016; An *et al.*, 2024), so maintenance decisions involve both the cost of restoring equipment and the effects of failures on service provision.

Two distinct problems are involved. Estimating how often components fail can rely on reliability models checked against recorded occurrences, whereas estimating what a failure does to the service depends on severity, detection location and operating conditions and requires disruption records rarely linked to component events. This difference in data and admissible validation structures the objectives, modules and conclusions.

Bayesian networks are well established in dependability, risk and maintenance analysis (Weber *et al.*, 2012), and rail applications have used them to assess equipment risk (Zhu *et al.*, 2026). Related studies have evaluated passenger door failures through FMECA (Dinmohammadi *et al.*, 2016), modeled disruption risk and disruption duration in metro networks (Liu *et al.*, 2021; Louie *et al.*, 2017), optimized preventive maintenance of subway train components from a failure-risk perspective (Tian and Wang, 2022) and allocated maintenance budgets in subway systems (El Hamshary *et al.*, 2022). FMECA scores failure modes ordinally, network disruption models do not start from component reliability, and maintenance optimization rarely translates failure risk into canceled trips. The gap addressed is the joint representation, in one probabilistic structure, of component reliability estimated from maintenance records and of service consequences conditioned on severity, detection location and operating period, with the evidential status of each part made explicit.

The train door system, cycled at every station stop, represents a small share of rolling stock cost but is a recurrent source of service interruptions (Dinmohammadi *et al.*, 2016; Connor, 2018). The research question is how historical failure records and expert judgment can be combined to describe maintenance costs and potential service impacts within one decision-support framework.

The general objective was to develop and apply a Bayesian network framework relating train door component failure probabilities to maintenance costs and to scenario-based operational and socioeconomic consequences. The specific objectives were to estimate mission-specific failure probabilities for 22 components, compare expected counts with 2019 records, make the calculation of operational outputs traceable, and examine how hypothetical changes in reliability,

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

severity and detection location affect the ordering of interventions, for one operator and one subsystem.

Accordingly, Module I produces failure estimates comparable with recorded occurrences, while Module II calculates canceled trips and a monetary proxy conditional on expert judgments and assumptions. No superiority over existing methods and no empirical effectiveness of the simulated interventions is claimed.

1. THEORETICAL BACKGROUND

1.1. Reliability, maintenance and service consequences

Reliability is the probability that an item performs its required function during a specified mission, expressed in time, distance or cycles, under stated conditions (Leemis, 1995; Fogliatto and Ribeiro, 2009). The economic assessment of preventive and corrective interventions depends on the time horizon and consequences considered (Stenström *et al.*, 2016; Galar *et al.*, 2017), and recent reviews favor policies that use operational information to anticipate deterioration (Falamarzi *et al.*, 2019; Bianchi *et al.*, 2025).

Subway operations generate benefits related to travel time, vehicle operating costs, emissions and accidents (Haddad *et al.*, 2015; An *et al.*, 2024), which operators may report in monetary terms (Companhia do Metropolitano de São Paulo, 2023). The aggregate benefit of a transport system and the marginal loss caused by one canceled trip are, however, different quantities. The marginal loss depends on the number of passengers affected, the additional waiting and travel time, the availability of alternative services, the duration and propagation of the disruption and the value attributed to users' time (Sun *et al.*, 2016; Louie *et al.*, 2017; Soza-Parra *et al.*, 2019). None of these variables was measured here. The monetary output is therefore treated as a scenario-based proxy, not as an estimate of social loss.

1.2. Bayesian networks and decision support

A Bayesian network represents a joint probability distribution through a directed acyclic graph and conditional probability tables. Its arcs encode the dependencies specified by the modeler and do not, by themselves, demonstrate causal effects (Koller and Friedman, 2009). Bayes' theorem relates prior and conditional probabilities:

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)} \quad (1)$$

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

In Equation (1), P(A | B) is the probability of A conditional on B, P(B | A) is the probability of B conditional on A, and P(A) and P(B) are marginal probabilities, with P(B) > 0. Bayesian modeling allows empirical frequencies and elicited judgments to be represented together, but the estimates remain conditional on the quality and coherence of both inputs (Fenton and Neil, 2019). Sensitivity analysis is therefore a standard step in network construction (Kjærulff and Madsen, 2013).

Utility nodes, borrowed from influence diagrams (Fenton and Neil, 2019), compute expected counts and costs. Alternatives are compared by changing input probabilities, treated as hypothetical interventions on the model inputs; the differences describe the behavior of the model, not causal effects identified from data.

1.3. Reliability and hazard functions

Let T denote the accumulated mission at failure. The cumulative distribution function F(t) gives the probability of failure by mission t, and the survival function R(t) the probability of remaining functional beyond it:

$$R(t) = P(T > t) = 1 - F(t), \; t \geq 0 \tag{2}$$

The hazard function h(t) is a conditional rate, not a probability. For a continuous lifetime distribution with density f(t), it is defined where R(t) > 0 by:

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)} \tag{3}$$

With missions in kilometers, the hazard has units of inverse kilometers. Decreasing, constant and increasing hazards are associated with early failures, random malfunction and wear-out (Fogliatto and Ribeiro, 2009), but a fitted shape does not identify the physical mechanism.

1.4. Weibull lifetime model and repairable systems

The three-parameter Weibull distribution has shape β > 0, scale η > 0 and location γ. For t > γ, its cumulative distribution, survival, density and hazard functions are:

$$F(t) = 1 - \exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right] \tag{4}$$

$$R(t) = \exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right] \tag{5}$$

$$f(t) = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right] \tag{6}$$

$$h(t) = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} \tag{7}$$

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

For t ≤ γ, F(t) = 0 and R(t) = 1; with γ = 0 the model reduces to the two-parameter form. The scale parameter is a characteristic life, not the mean. Values of β below, equal to and above one correspond to decreasing, constant and increasing hazards (Fogliatto and Ribeiro, 2009). Here, the failure probability of a unit during the annual mission is $p = F(t_0 + m) - F(t_0)$, where t_0 is the accumulated distance of the fleet at the start of the mission and m is approximately 140,000 km (Campos, 2022).

Successive failures at one position of a repairable system form a point process (Ascher and Feingold, 1984; Rausand and Høyland, 2004). Under perfect repair, assumed in the original model (Campos, 2022), the expected number of failures in the mission is the increment of the renewal function; under minimal repair, failures follow a non-homogeneous Poisson process whose expected number is the increment of the cumulative hazard, $\ln R(t_0) - \ln R(t_0 + m)$. The probability p counts only failures of units in service since the beginning of operation and ignores units already replaced, so it is lower than both quantities. In the calculation reported by Campos (2022) for component C02 (β = 0.73, η = 12,218,518 km, t_0 = 730,944 km), $F(t_0)$ = 10.3% and p = 1.77%, whereas the cumulative-hazard increment is 1.99%, that is, 12.6% higher. For decreasing hazards and fleets with many replaced units, the approximation therefore tends to underestimate occurrences.

2. METHODS

2.1. Study design and records

An applied modeling and simulation study was conducted on 22 door-system components of the train fleet of one subway operator, using Minitab, spreadsheets and GeNIe Academic 3.0. The iterative cycle of conception, implementation and analysis (Chwif and Medina, 2013) involved specialists at three points: verification of failure data, definition of interactions among risk factors and assessment of the implemented network.

The maintenance records cover eight years of operation since the fleet entered service, with about 7,800 failure records and 24 variables per record, including component identification, accumulated distance, severity, detection location, date and time (Campos, 2022). Severity was classified as light or severe according to its potential operational consequence, not as a passenger-safety classification, and detection location distinguished the depot from the commercial route. The installed units of each component (Table 3) are consistent with the 1,104 door units of the fleet adopted in the original model, with multiples for components installed more than once per door and fractions for components shared by several doors (Campos, 2022).

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

Failure and censoring counts by component are reported in Supplementary Table S1. The fitting database covers the whole operating period up to the 2019 mission, so the 2019 records were not set aside for validation; the comparison in Section 3.1 is therefore a retrospective, in-sample assessment of model adequacy, not a temporal validation.

2.2. Mission-specific probabilities and empirical inputs

For each component, Weibull models were fitted to the accumulated distance between failures, including right-censored units, with the parametric distribution analysis of Minitab. Adherence was assessed with probability plots for right-censored data, compared across alternative distributions and summarized by the correlation coefficient of the probability plot with 95% confidence bounds; the three-parameter form gave the best fit for most components (Campos, 2022). The original estimates could not be recovered; a later maximum-likelihood re-estimation with records up to 2020 (Supplementary Table S1) found no fit improvement from the three-parameter form (adjusted Anderson-Darling differences below 0.2) and an inestimable covariance for its location parameter. The mission of 140,000 km is the average annual distance per train, a fleet average rather than a component-specific exposure; for mechanisms driven by opening and closing cycles, distance is an adequate proxy only if stops per kilometer are stable across the fleet, which was not tested.

Probabilities of severity and detection location were estimated as empirical proportions in the historical records (Table 1). In GeNIe, the node named Component reliability contains failure and success states; the values reported here are the failure probabilities p, with success represented by 1 − p.

Table 1. Historical proportions of failure severity and detection location by component

ID	Severe	Light	Depot	Route	ID	Severe	Light	Depot	Route
C01	78.5%	21.5%	31.4%	68.6%	C12	96.6%	3.4%	53.6%	46.4%
C02	58.5%	41.5%	37.7%	62.3%	C13	86.2%	13.8%	31.8%	68.2%
C03	89.6%	10.4%	24.3%	75.7%	C14	91.1%	8.9%	32.9%	67.1%
C04	92.3%	7.7%	24.0%	76.0%	C15	91.0%	9.0%	33.2%	66.8%
C05	96.4%	3.6%	19.0%	81.0%	C16	95.3%	4.7%	14.7%	85.3%
C06	92.9%	7.1%	26.1%	73.9%	C17	95.3%	4.7%	14.7%	85.3%
C07	94.9%	5.1%	25.1%	74.9%	C18	24.6%	75.4%	70.5%	29.5%
C08	96.8%	3.2%	53.9%	46.1%	C19	24.6%	75.4%	70.5%	29.5%
C09	95.3%	4.7%	8.9%	91.1%	C20	7.1%	92.9%	34.1%	65.9%
C10	93.2%	6.8%	3.6%	96.4%	C21	84.1%	15.9%	40.2%	59.8%
C11	95.7%	4.3%	51.2%	48.8%	C22	84.1%	15.9%	40.2%	59.8%

Source: Maintenance records of the operator. Route: commercial route. Severity and detection location are marginal proportions and enter the network as independent nodes.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

Severity and detection location are independent root nodes, so their joint probability is the product of the marginal proportions, an assumption adopted because the records were summarized marginally. If severe failures are detected on the route more often than this product implies, the impacts reported below are conservative. Cross-tabulating severity and detection location in the original records would allow this dependence to be represented in future applications.

Operating-period probabilities reflect the share of fleet mileage in weekday peak (38%), weekday lower flow (41%), weekend peak (8%) and weekend lower flow (13%) periods, with peaks from 06:00 to 10:00 and 16:00 to 20:00 and lower flow in the remaining service hours, from 05:00 to 00:40. Mileage and failure distributions across these periods showed similar profiles, with differences attributed to failures recorded at the start and end of operation, when trains undergo door checks in the depot.

2.3. Operational interference and expert judgment

Operational interference probabilities were elicited from ten operation and maintenance specialists of the operator, with 12 to 35 years of professional experience (mean of 24.2 years) in centralized operation, operation and maintenance engineering, operational strategy, traffic supervision and maintenance management (Campos, 2022). Empirical estimation was not possible because interferences are recorded in a separate database and only when they exceed five minutes. After interviews, an online questionnaire presented the 16 combinations of severity, detection location and operating period, and each specialist assigned one category of Table 2 to each combination, considering train retention at the platform as the operational effect. The questionnaire converted the answers into the conditional probabilities of the node (Campos, 2022); with one answer per specialist and combination, each probability corresponds to the proportion of specialists selecting the category, that is, an equal-weight linear opinion pool (O'Hagan *et al.*, 2006). The specialists' profiles and the questionnaire structure are reported in Supplementary Material S2; the complete conditional probability table could not be retrieved, and its effect on the outputs is documented by the values recovered in Table 5. The specialists also reviewed the implemented network, which constitutes face validation, not quantitative validation.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

Table 2. Operational interference categories and canceled-trip weights

Interference category	Canceled-trip weight v_k
None	0
Minor	1
Medium	5
Major	10

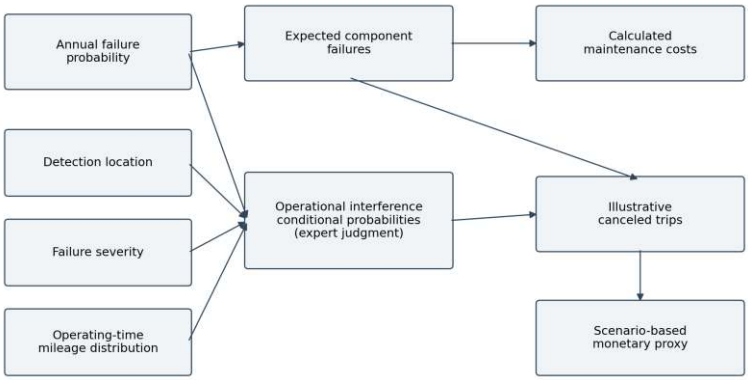
Source: Categories defined with the specialists. The categories classify the effect of train retention at the platform on the supply of trips, from no canceled trip to ten canceled trips (Campos, 2022). They are modeling categories, not empirical estimates; their influence on the outputs is examined in Section 3.3.

The same conditional structure was applied to all components, since every door failure triggers the same standardized operational response (Campos, 2022); component-specific differences could not be tested because interference was not recorded per component event.

2.4. Model structure and output calculations

The framework has two modules (Figure 1). Module I translates annual failure probabilities into expected failure counts and maintenance costs. Module II combines failure, severity, detection location and operating period with expert judgments to calculate canceled trips and the monetary proxy.

Figure 1. Conceptual integration of failure estimates and exploratory impact calculations



Source: Prepared by the authors. Probabilistic inputs and calculated outputs are distinguished; the diagram does not represent causal intervention effects.

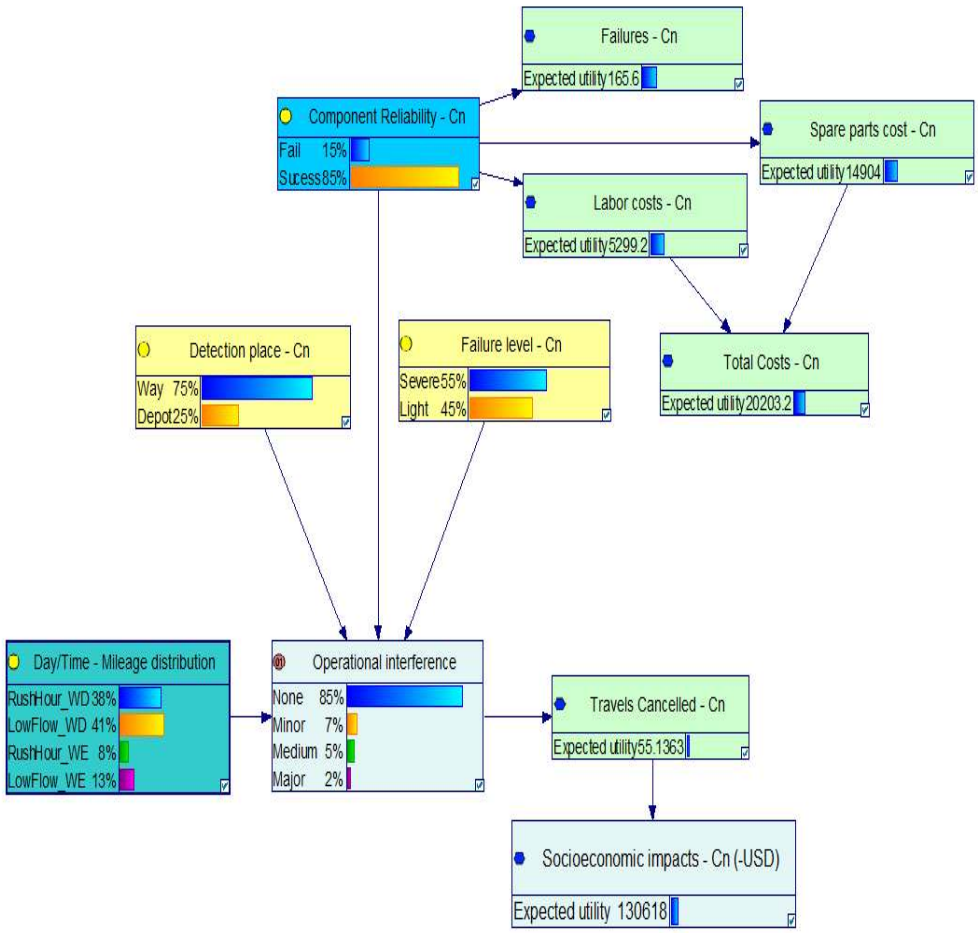
Figure 2 shows the GeNIe network for one component under the baseline scenario. Its chance nodes are Component reliability (fail, success), Detection place (way, depot), Failure level (severe, light), Day/Time mileage distribution (four periods) and Operational interference (none, minor, medium, major), the only node with parents; given success it equals none, so its elicited

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

table has 2 × 2 × 4 = 16 parent configurations. Utility nodes compute expected failures, costs, canceled trips and the monetary proxy.

The system model repeats the component network for C1 to Cn, with n = 22 (Figure 3). The component modules share an operational context, while their expected outputs are aggregated at the system level. Shared operational context does not establish empirical interdependence among component failures; likewise, statistical independence among those failures has not been demonstrated. Summed operational impacts may overcount when simultaneous failures produce a single interruption.

Figure 2. Component network implemented in GeNIe under the baseline scenario



Source: Prepared by the authors in GeNIe Academic 3.0. Bars show marginal probabilities; utility nodes show expected values in failures, USD and canceled trips.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

For each component c, the expected failure count is approximated by multiplying the number of installed units N_c by the mission failure probability p_c:

$$\hat{n}_c = N_c p_c = N_c[F_c(t_0 + m) - F_c(t_0)] \quad (8)$$

As discussed in Section 1.4, Equation (8) counts failures of units in service since the beginning of operation and tends to underestimate occurrences when replaced units fail again. The maintenance cost uses the unit spare-part cost a, the hourly labor cost b and the mean repair time τ:

$$M_c = \hat{n}_c(a + b\tau) \quad (9)$$

In the scenarios, a = USD 90, b = USD 16 per hour and τ = 2 h, giving USD 122 per failure. These values correspond to BRL 450 per part and BRL 160 for two labor hours in the original model (Campos, 2022), converted at BRL 5.00 per USD (2024). They are assumed inputs, not observed expenditure.

For the operational output, let $\pi_{k|sdt}$ denote the elicited probability of interference category k given a failure with severity s, detection location d and operating period t, and v_k the weight in Table 2. The expected canceled trips per failure for each severity and location combination, averaged over the operating periods, are:

$$T_{sd} = \sum_t P(t) \sum_{k=0}^{3} v_k \pi_{k|sdt} \quad (10)$$

In the implemented utility node, canceled trips are obtained by multiplying the marginal probability of each interference category by 100 and by its weight (Campos, 2022). For component c, expected canceled trips and the monetary proxy are therefore:

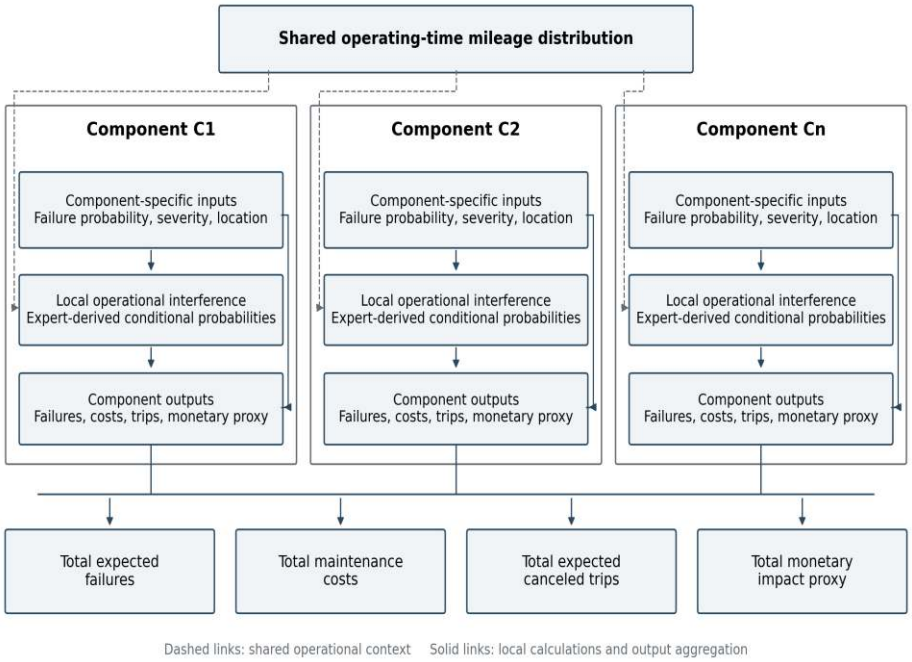
$$C_c = 100\, p_c \sum_s \sum_d P_c(s) P_c(d) T_{sd} \quad (11)$$

$$S_c = VC_c \quad (12)$$

The fixed multiplier expresses canceled trips for a reference population of 100 units rather than for the N_c installed units; scaling by N_c = 1,104 would multiply absolute trips and monetary values by 11.04, without changing comparisons between scenarios with the same N_c. The four values of T_{sd} were recovered from the network outputs, since the four scenarios crossing two severity levels with two detection levels (Section 2.6) provide four linear equations in the four unknowns of Equation (11). Within rounding, they reproduce the modernization scenario (18.38 against 18.4 canceled trips) and the six configurations of the sensitivity analysis reported by Campos (2022), for example 44.76 against 44.77 canceled trips with p = 20%, 45% severe failures and 40% route detection.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

Figure 3. System-level architecture with shared operational context and aggregation of component outputs



Source: Prepared by the authors. C1, C2 and Cn illustrate the repeated architecture; no cross-component failure dependencies are represented.

2.5. Historical comparison and error measures

Expected counts from Module I were compared with the 2019 failure counts by component, by subsystem and in aggregate. No corresponding comparison was possible for Module II. For J components, with observed count y_c and estimate $\hat{n}_c$, the mean absolute error (MAE) and weighted absolute percentage error (WAPE) are:

$$\mathrm{MAE} = \frac{1}{J}\sum_{c=1}^{J}|\hat{n}_c - y_c| \tag{13}$$

$$\mathrm{WAPE} = 100 \times \frac{\sum_{c=1}^{J}|\hat{n}_c - y_c|}{\sum_{c=1}^{J} y_c} \tag{14}$$

The signed relative error is $100(\hat{n}_c - y_c)/y_c$, with negative values indicating underestimation. To express the variability expected even under a correct model, 95% prediction intervals were obtained from a Poisson distribution with mean $\hat{n}_c$, and overdispersion was assessed with the Pearson statistic $X^2 = \Sigma(y_c - \hat{n}_c)^2/\hat{n}_c$ with J degrees of freedom. These intervals ignore the uncertainty of the Weibull estimates and are therefore narrower than full predictive intervals. As

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

naive references, each 2019 count was also forecast by the component's 2018 count and by its mean annual count from 2016 to 2018, taken from the operator's annual failure records since 2011.

2.6. Hypothetical scenarios and sensitivity analysis

The scenarios consider an illustrative component with 1,104 installed units, one per door position, and the inputs in Table 4. Baseline assumptions were p = 15%, 75% detection on the commercial route and 55% severe failures. Investments and probability changes are hypotheses formulated to explore the behavior of the model, not measured effects of implemented interventions; the investments of USD 6,000, 66,000 and 1,200 correspond to BRL 30,000, 330,000 and 6,000 in the original model. The monetary proxy V = USD 2,369 per canceled trip corresponds to BRL 11,846 (Campos, 2022), defined in the original model as the proportional share of the system's socioeconomic benefits attributed to each trip. As an average benefit per trip, it is not a marginal loss (Section 1.1).

Sensitivity analyses addressed the expected canceled trips per failure T_{sd}, which concentrate the expert-derived information and also absorb proportional changes in the weights of Table 2. Each T_{sd} was varied by ±25% one at a time, and the four values were then multiplied jointly by independent factors drawn from a uniform distribution between 0.5 and 1.5 in 20,000 Monte Carlo draws. Since V, the unit costs and the multiplier of Equation (11) act multiplicatively, they rescale all alternatives equally and do not alter comparisons. Interventions were compared by avoided canceled trips and by investment per avoided trip, a ratio meaningful only for ranking that excludes maintenance savings and an evaluation horizon.

3. RESULTS AND DISCUSSION

3.1. Comparison of component failure estimates

Table 3 presents the installed units, annual probabilities, expected and observed counts, Poisson prediction intervals and signed relative errors. The model estimated 822.6 failures against 940 recorded, an aggregate difference of −117.4 failures (−12.5%). The recorded total lies outside the aggregate 95% interval (767 to 879), indicating a systematic underestimation beyond random variation.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

Table 3. Annual failure estimates compared with recorded occurrences in 2019

ID	Component	Sub.	N_c	p_c (%)	Expected	Observed	95% PI	Error (%)
C01	Locking hook	1	414	3.00	12.4	12	6 to 20	+3.3
C02	Door leaf	1	1,104	1.55	17.1	16	10 to 26	+6.9
C03	Bearings	1	2,208	0.91	20.1	17	12 to 29	+18.2
C04	Window glass	1	2,208	2.42	53.4	69†	40 to 68	−22.6
C05	Support rail	1	1,104	3.15	34.7	25	24 to 47	+38.8
C06	Conjugation device	2	1,104	3.41	37.6	35	26 to 50	+7.4
C07	Opening piston	2	1,104	2.67	29.5	41	19 to 41	−28.0
C08	Unlock cylinder	2	1,104	3.56	39.3	90†	28 to 52	−56.3
C09	Transmission axle	2	2,208	3.25	71.7	53†	56 to 89	+35.3
C10	Mechanical lock	2	1,104	4.69	51.8	67†	38 to 66	−22.7
C11	Pneumatic hose	2	4,416	3.76	166.1	143	141 to 192	+16.2
C12	Electrovalve	2	1,104	2.37	26.2	39†	17 to 37	−32.8
C13	Car doors control unit	3	1,104	2.22	24.5	26	15 to 35	−5.8
C14	Network unit	3	1,104	3.43	37.9	42	26 to 50	−9.8
C15	Door unit	3	1,104	2.64	29.1	70†	19 to 40	−58.4
C16	Closing switch	3	1,104	0.75	8.3	34†	3 to 14	−75.6
C17	Locking switch	3	1,104	5.77	63.7	74	49 to 80	−13.9
C18	Service key	3	3,312	1.34	44.3	31†	32 to 58	+42.9
C19	Emergency device	3	1,104	1.67	18.4	20	11 to 27	−8.0
C20	Emergency device cover	3	2,208	0.41	9.1	9	4 to 15	+1.1
C21	Electrical wiring	4	2,208	1.18	26.0	26	17 to 36	+0.0
C22	General control unit	5	138	1.02	1.4	1	0 to 4	+40.0
Total	**All components**				**822.6**	**940†**	**767 to 879**	**−12.5**

Source: Prepared by the authors from the operator's records. Sub.: subsystem. N_c: installed units. PI: Poisson prediction interval for the expected count. † Observed count outside the interval. Subsystems, grouped as in Campos (2022) and labeled by component function: 1, door leaf and guidance (C01 to C05); 2, pneumatic and mechanical drive (C06 to C12); 3, control, switches and emergency devices (C13 to C20); 4, wiring (C21); 5, general control (C22).

The sum of absolute errors was 262.4 failures, giving an MAE of 11.93 failures and a WAPE of 27.9%. Fourteen of the 22 observed counts fell within their intervals, and the Pearson statistic was 240.2 with 22 degrees of freedom, a dispersion about eleven times the Poisson value. The contrast between the aggregate error and the WAPE reveals compensating errors: a satisfactory fleet total does not guarantee comparable precision in spare-part and labor estimates for each component. The previous-year forecast gave an MAE of 15.59 and a WAPE of 36.5% (total 1,185), and the three-year mean an MAE of 11.55 and a WAPE of 27.0% (total 1,055.3): the model outperformed the simplest reference and matched a moving average, while adding the mission structure needed for scenarios.

The largest underestimations occurred for the unlock cylinder (C08: 39.3 expected, 90 observed, −56.3%), the door unit (C15: 29.1 against 70, −58.4%) and the closing switch (C16:

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

8.3 against 34, −75.6%), all far outside their intervals. For C08 and C16, replacements of under-warranty units with low performance were recorded as failures in the maintenance system, inflating the observed counts. For C15 no such explanation was identified; repeated failures at the same positions, which Equation (8) does not count (Section 1.4), are a plausible candidate. The same mechanism, which affects every component with replaced units, is consistent with the aggregate underestimation. Overestimation outside the interval occurred for C09 and C18, and the single failure of C22 shows how relative errors become unstable for very small counts.

At subsystem level, the expected and observed counts were 137.7 and 139 (subsystem 1), 422.2 and 468 (subsystem 2), 235.3 and 306 (subsystem 3), 26.0 and 26 (subsystem 4) and 1.4 and 1 (subsystem 5). Subsystems 2 and 3 fall outside their intervals and concentrate the largest underestimations, so drive and control components deserve priority in data review; agreement in subsystems 4 and 5 rests on very few events.

3.2. Scenario outputs and traceable calculation

Table 4 combines the scenario inputs and the network outputs; depot detection and light failures are the complements of the listed probabilities, and the operating-period distribution was kept constant.

Table 4. Hypothetical scenario inputs and outputs for 1,104 installed units

Scenario	p	Severe	On route	Invest. (USD)	Expected failures	Maint. cost (USD)	Canceled trips	Monetary proxy (USD)
Baseline	15%	55%	75%	None	165.6	20,203.20	55.1	130,618.00
Lubrication improvement	12%	25%	75%	6,000	132.5	16,162.60	35.4	83,898.10
Modernization	5%	55%	75%	66,000	55.2	6,734.40	18.4	43,539.30
Depot inspection improvement	15%	55%	30%	1,200	165.6	20,203.20	29.9	70,796.40
Lubrication and depot inspection	12%	25%	30%	7,200	132.5	16,162.60	20.4	48,398.70

Source: Prepared by the authors from the network outputs. Inputs are hypothetical; outputs were not validated against observed disruptions. Monetary values preserve the precision reported by GeNIe, so products of rounded counts may differ slightly.

The baseline calculation can be followed step by step. Equation (8) gives 1,104 × 0.15 = 165.6 expected failures, and Equation (9) gives 165.6 × 122 = USD 20,203.20. Table 5 lists the expected canceled trips per failure T_{sd} recovered for each severity and location combination. With the baseline weights P(s)P(d), equal to 0.4125 for severe failures on route, 0.1375 for severe failures in the depot, 0.3375 for light failures on route and 0.1125 for light failures in the depot, the expected trips per failure are 3.676. Equation (11) then gives 100 × 0.15 × 3.676 = 55.1 canceled trips, and Equation (12) gives 55.14 × 2,369 = USD 130,618.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

Table 5. Expected canceled trips per failure implied by the implemented network

Detection location	Severe failure	Light failure
Commercial route	6.06	2.84
Depot	0.87	0.87

Source: Prepared by the authors; values T_{sd} of Equation (10), recovered from the scenario outputs and checked against seven configurations not used in their derivation (Section 2.4).

Severe failures detected on the route account for 68% of the baseline canceled trips, and depot detections for about 6%. Hence the depot-inspection scenario reduces canceled trips from 55.1 to 29.9 without changing failures or costs, by moving probability from the most to the least consequential cell of Table 5; whether reinforced inspections achieve this shift in practice was not established.

3.3. Sensitivity of the scenario comparison

In the one-at-a-time analysis, a ±25% change in the value for severe failures on the route moved baseline canceled trips between 45.8 and 64.5, whereas the same change in any other value altered them by less than 7%. Table 6 summarizes the multivariate analysis. Across all draws, the ordering of interventions by investment per avoided trip was unchanged: depot inspection, combined intervention, lubrication and modernization. This stability arises because the assumed investments differ by more than one order of magnitude, while avoided trips vary by a factor of about two.

Table 6. Avoided canceled trips and investment per avoided trip under multivariate perturbation

Intervention	Avoided trips	90% interval	Investment per avoided trip (USD)	90% interval (USD)
Lubrication improvement	19.7	8.7 to 30.7	304	195 to 687
Modernization	36.8	24.3 to 49.1	1,796	1,343 to 2,717
Depot inspection improvement	25.2	13.9 to 36.6	48	33 to 86
Lubrication and depot inspection	34.7	19.8 to 49.6	207	145 to 364

Source: Prepared by the authors. 20,000 draws with each T_{sd} multiplied by an independent factor uniformly distributed between 0.5 and 1.5. Point values use the unperturbed network.

The comparison between the combined intervention and modernization was not robust: the ratio of their avoided trips had a median of 0.95 (90% interval from 0.77 to 1.05), and the combination avoided at least as many trips in 24% of the draws. Its similarity to modernization with about one tenth of the investment is therefore conditional on the elicited values. It also ignores the lower annual maintenance cost of modernization (USD 6,734.40 against USD 16,162.60), whose relevance depends on service life, horizon and discounting, which were not modeled.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

3.4. Interpretation and contribution

Relative to maintenance-cost assessment (Stenström *et al.*, 2016; Galar *et al.*, 2017) and subway budget allocation (El Hamshary *et al.*, 2022), the application organizes and traces inputs with different evidential status, without demonstrating optimal allocation.

Unlike FMECA applied to passenger doors (Dinmohammadi *et al.*, 2016), the framework propagates severity and detection probabilities and expresses consequences in trips and monetary units, although no quantitative comparison supports superior decision quality. Unlike knowledge-graph Bayesian networks (Zhu *et al.*, 2026) and network disruption models (Liu *et al.*, 2021; Wang *et al.*, 2025), it starts from fitted component reliability but omits passenger redistribution, service recovery and overlapping incidents.

For managers, component-level errors and overdispersion are large enough to affect spare-part planning for drive and control components, and in Module II the value for severe failures detected on the route dominates the outputs, so linking incident logs to component events would reduce most of the uncertainty in the scenario comparison.

3.5. Limitations

The application covers one operator and one subsystem. The comparison with 2019 is retrospective and in-sample, the Poisson intervals ignore parameter uncertainty, the overdispersion reveals heterogeneity among positions or recording practices that the model does not represent, and at component level the model did not outperform a three-year moving average. Distance served as exposure for mechanisms that may depend on cycles, the mission probability ignores failures of units replaced before the mission, and the original Weibull estimates and conditional probability table could not be recovered, which limits external reanalysis to the reconstruction and re-estimation reported.

Module II rests on elicited probabilities without quantified uncertainty, a reference population of 100 units, independence between severity and location and one consequence structure for all components, and its outputs were not compared with recorded cancellations. With unverified unit costs, an average benefit per trip and assumed intervention effects, the simulations are not a cost-benefit evaluation.

4. FINAL CONSIDERATIONS

This study integrated failure records and expert judgment in a Bayesian framework relating train door component failures to maintenance costs and scenario-based consequences. Three

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

results were demonstrated. First, a documented probabilistic structure links fitted component reliability, severity, detection location and operating period to costs and canceled trips. Second, the retrospective comparison (822.6 expected against 940 recorded failures; MAE of 11.93; WAPE of 27.9%) outperformed a previous-year forecast but not a three-year mean, and the aggregate total conceals substantial heterogeneity. Third, the scenario calculations were made traceable, showing that the ordering of interventions by investment per avoided trip is robust to large changes in the elicited values, whereas the near-equivalence between modernization and the combined intervention is not.

The framework is a methodological development with exploratory managerial use, not a validated forecasting tool, and its application depends on documented parameters, data quality and prospective comparison with observed outcomes. Further research should adopt opening and closing cycles as exposure, Bayesian updating with new records, temporal validation, explicit models of recurrent failures and recorded canceled trips and interruption times.

DECLARATIONS

Competing interests: The authors declare no competing interests.

Funding: This research did not receive a specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Use of generative artificial intelligence: OpenAI ChatGPT and Anthropic Claude were used to assist with language revision, manuscript organization, consistency checking and formatting, and Claude was used to draft code for recalculating error measures, prediction intervals and the sensitivity analysis. The tools were not used to collect data, elicit expert judgments or fit the Weibull models. All recalculated values were checked against the outputs of the original GeNIe model reported by Campos (2022), and the authors are responsible for the content of the manuscript.

Data and model availability: Component-level counts, estimates and scenario values are reported in the article. The Supplementary Material provides re-estimated Weibull parameters, the specialists' profiles, the recovered trips per failure and annual failure counts; the original estimates and conditional probability table were not available for reanalysis. The raw maintenance records were provided by the operator and are not publicly available.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

REFERENCES

AN, S.; AHN, K.; BAE, J.; SONG, Y. Economic impacts of a subway system: exploring local contexts in a metropolitan area. **Research in Transportation Business & Management**, v. 56, art. 101188, 2024. DOI: https://doi.org/10.1016/j.rtbm.2024.101188.

ASCHER, H.; FEINGOLD, H. **Repairable systems reliability**: modeling, inference, misconceptions and their causes. New York: Marcel Dekker, 1984.

BIANCHI, G.; FANELLI, C.; FREDDI, F.; GIULIANI, F.; LA PLACA, A. Systematic review railway infrastructure monitoring: from classic techniques to predictive maintenance. **Advances in Mechanical Engineering**, v. 17, n. 1, 2025. DOI: https://doi.org/10.1177/16878132241285631.

CAMPOS, E. A. **Modelagem bayesiana da influência do risco de falha em equipamentos nos custos, interferências operacionais e impactos socioeconômicos de uma operação metroferroviária**. 2022. Dissertation (Master in Production Engineering) – Universidade Nove de Julho, São Paulo, 2022.

CHWIF, L.; MEDINA, A. C. **Modelagem e simulação de eventos discretos**. 3. ed. São Paulo: Elsevier, 2013.

COMPANHIA DO METROPOLITANO DE SÃO PAULO. **Relatório integrado 2023**. São Paulo: Metrô, 2023. Available at: https://www.metro.sp.gov.br/wp-content/uploads/2024/03/Relatorio-integrado-2023.pdf. Accessed on: 25 Jul. 2024.

CONNOR, P. R. **Managing technical and operational change**: how understanding a railway's historic evolution can guide future development: a London Underground case study. 2018. Thesis (Doctor of Philosophy) – University of Birmingham, Birmingham, 2018. Available at: https://etheses.bham.ac.uk/id/eprint/8415/. Accessed on: 22 Oct. 2026.

DINMOHAMMADI, F.; ALKALI, B.; SHAFIEE, M.; BÉRENGUER, C.; LABIB, A. Risk evaluation of railway rolling stock failures using FMECA technique: a case study of passenger door system. **Urban Rail Transit**, v. 2, n. 3-4, p. 128-145, 2016. DOI: https://doi.org/10.1007/s40864-016-0043-z.

EL HAMSHARY, O.; ABOUHAMAD, M.; MARZOUK, M. Integrated maintenance planning approach to optimize budget allocation for subway operating systems. **Tunnelling and Underground Space Technology**, v. 121, art. 104322, 2022. DOI: https://doi.org/10.1016/j.tust.2021.104322.

FALAMARZI, A.; MORIDPOUR, S.; NAZEM, M. A review of rail track degradation prediction models. **Australian Journal of Civil Engineering**, v. 17, n. 2, p. 152-166, 2019.

FENTON, N. E.; NEIL, M. **Risk assessment and decision analysis with Bayesian networks**. 2. ed. Boca Raton: CRC Press, 2019.

FOGLIATTO, F. S.; RIBEIRO, J. L. D. **Confiabilidade e manutenção industrial**. Rio de Janeiro: Elsevier, 2009.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

GALAR, D.; SANDBORN, P. A.; KUMAR, U. **Maintenance costs and life cycle cost analysis**. Boca Raton: CRC Press, 2017.

HADDAD, E. A.; HEWINGS, G. J. D.; PORSSE, A. A.; VAN LEEUWEN, E. S.; VIEIRA, R. S. The underground economy: tracking the higher-order economic impacts of the São Paulo Subway System. **Transportation Research Part A: Policy and Practice**, v. 73, p. 18-30, 2015.

KJÆRULFF, U. B.; MADSEN, A. L. **Bayesian networks and influence diagrams**: a guide to construction and analysis. 2. ed. New York: Springer, 2013.

KOLLER, D.; FRIEDMAN, N. **Probabilistic graphical models**: principles and techniques. Cambridge: MIT Press, 2009.

LEEMIS, L. M. **Reliability**: probabilistic models and statistical methods. Englewood Cliffs: Prentice Hall, 1995.

LIU, K.; ZHU, J.; WANG, M. An event-based probabilistic model of disruption risk to urban metro networks. **Transportation Research Part A: Policy and Practice**, v. 147, p. 93-105, 2021. DOI: https://doi.org/10.1016/j.tra.2021.03.010.

LOUIE, J.; SHALABY, A.; HABIB, K. N. Modelling the impact of causal and non-causal factors on disruption duration for Toronto's subway system: an exploratory investigation using hazard modelling. **Accident Analysis & Prevention**, v. 98, p. 232-240, 2017. DOI: https://doi.org/10.1016/j.aap.2016.10.008.

O'HAGAN, A.; BUCK, C. E.; DANESHKHAH, A.; EISER, J. R.; GARTHWAITE, P. H.; JENKINSON, D. J.; OAKLEY, J. E.; RAKOW, T. **Uncertain judgements**: eliciting experts' probabilities. Chichester: Wiley, 2006.

RAUSAND, M.; HØYLAND, A. **System reliability theory**: models, statistical methods, and applications. 2. ed. Hoboken: Wiley, 2004.

SOZA-PARRA, J.; RAVEAU, S.; MUÑOZ, J. C.; CATS, O. The underlying effect of public transport reliability on users' satisfaction. **Transportation Research Part A: Policy and Practice**, v. 126, p. 83-93, 2019. DOI: https://doi.org/10.1016/j.tra.2019.06.004.

STENSTRÖM, C.; NORRBIN, P.; PARIDA, A.; KUMAR, U. Preventive and corrective maintenance: cost comparison and cost-benefit analysis. **Structure and Infrastructure Engineering**, v. 12, n. 5, p. 603-617, 2016.

SUN, H.; WU, J.; WU, L.; YAN, X.; GAO, Z. Estimating the influence of common disruptions on urban rail transit networks. **Transportation Research Part A: Policy and Practice**, v. 94, p. 62-75, 2016.

TIAN, Q.; WANG, H. Optimization of preventive maintenance schedule of subway train components based on a game model from the perspective of failure risk. **Sustainable Cities and Society**, v. 81, art. 103819, 2022. DOI: https://doi.org/10.1016/j.scs.2022.103819.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

WANG, Y.; LI, Y.; CHEN, X.; ZHOU, Y. Urban rail transit disruption management: a decade of research on network resilience, internal, and external strategies. **Intelligent Transportation Infrastructure**, v. 4, art. liaf014, 2025. DOI: https://doi.org/10.1093/iti/liaf014.

WEBER, P.; MEDINA-OLIVA, G.; SIMON, C.; IUNG, B. Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. **Engineering Applications of Artificial Intelligence**, v. 25, n. 4, p. 671-682, 2012. DOI: https://doi.org/10.1016/j.engappai.2010.06.002.

ZHU, L.; LIU, S.; HUANG, Y.; WANG, F.; LIU, Z. Constructing Bayesian networks from knowledge graphs for risk assessment and causal inference of urban rail transit equipment. **Accident Analysis & Prevention**, v. 233, art. 108555, 2026. DOI: https://doi.org/10.1016/j.aap.2026.108555.

This article is published in open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.