International Journal of GEOMATE, Feb., 2024 Vol.26, Issue 114, pp.34-41
ISSN: 2186-2982 (P), 2186-2990 (O), Japan, DOI: https://doi.org/10.21660/2024.114.4084
Geotechnique, Construction Materials and Environment

THE DEVELOPMENT OF THE TRAIN ACCIDENT MODEL TO THE INFRASTRUCTURE FACTORS IN INDONESIA

*Dicky Arisikam[1,2], Harun Al Rasyid Lubis[1], Aine Kusumawati[1], Desiderius Viby Indrayana[1]

[1]Faculty of Civil and Environment Engineering, Bandung Institute of Technology, Bandung 40132, Indonesia
[2]Directorate of Infrastructure, PT Kereta Api Indonesia (Persero), Bandung 40132, Indonesia

*Corresponding Author, Received: 26 July 2023, Revised: 01 Jan. 2024, Accepted: 05 Jan. 2024

ABSTRACT: The majority of Train Accidents (TA) in Indonesia from 2015 to 2020 were caused by infrastructure factors, namely railway tracks, bridges, and signals. To mitigate these TAs, infrastructure maintenance is required, prioritizing locations with a high risk of TA. The prioritization of these locations can be accomplished through a model that estimates TA based on infrastructure risk factors. Locations with the highest TA estimates will be prioritized for infrastructure maintenance. This model illustrates the associative relationship between KKA as the dependent variable and exposure (train frequency and track length) and infrastructure risk factors (railway tracks, bridges, and signals) as independent variables. The model is constructed using the Generalized Linear Model (GLM), considering Poisson Regression (PR), Negative Binomial (NB), Zero Inflated Poisson (ZIP), and Zero Inflated Negative Binomial (ZINB) models. TA data from the Operational Areas (OA) of Jakarta, Bandung, and Cirebon during 2015-2020 were used to build the model, with the model entity being the segment between two train stations. The selection of the best model is based on tests of dispersion value, goodness-of-fit test, and Vuong test. Modeling results indicate that the NB model is the most suitable for illustrating the associative relationship between TA and infrastructure factors in the Indonesian Railway. The variables are train frequency (trains/day), track length (km), train speed (km/h), length of curves with a radius of 500 m to ≤ 1000 m (km), number of vulnerable areas (points), length of electricity network (km), and track type (single or double).

Keywords: Railway, Railway Accident, Generalized Linear Model (GLM), Negative Binomial (NB)

1. INTRODUCTION

The highest risk of TA is derailment [1], and one of the factors for this is infrastructure such as tracks, bridges, and signals. Data from the Indonesian Ministry of Transportation indicates that the TA from 2015 to 2019, 52% were caused by infrastructure, 23% by equipment, 12% by human resources, 7% by external factors, and 6% by natural factors. Several factors that can influence railway safety include (1) bridges [2]; (2) types of level crossing safety devices [3]; (3) train speed [3]; (4) train frequency [3]; (5) the number of level crossings [4]; (6) signal systems [5]; (7) turn-out [6]; (8) vulnerable areas [7]; (9) curves [8]; (10) track length [8]; (11) concrete sleepers [9]; (12) rails [10]. Comprehensive safety risk management and evaluation are needed to reduce the TA, accomplished by managing data such as infrastructure data, inspection data, maintenance data, etc. One of these is to understand the relationship between TA and infrastructure factors through a model. Accident analysis models and methods explain and even predict accident mechanisms, which help to choose and implement effective and efficient countermeasures [11].

Initially, linear regression was widely employed for accident modeling, where accident data was assumed to follow a normal distribution [12] with constant variance [13]. However, this distribution fails to accurately represent accident data, particularly those stemming from random and infrequent time and location events [14]. Accidents are random occurrences in which the number of vehicles involved in accidents on a specific road segment during a given period is probabilistic in nature [15]. Linear regression models must be used cautiously in safety studies because accident figures cannot be negative, and accident data exhibit non-constant variance [16]. Moreover, a normal distribution allows for negative values, especially in cases of low traffic volume. Linear regression models are unsuitable for predicting accident numbers due to their discrete [17], random [18], non-negative, and sporadic nature [19]. Hence, they cannot conform to a normal distribution, and their variance is also non-constant. The limitations of linear regression models in accident prediction have led to the development of Generalized Linear Models (GLM) [20], where non-linearity in GLM arises due to the inclusion of link functions. Over the past two decades, there has been extensive research in the field of transportation focusing on accident prediction using models like Poisson Regression (PR), Negative Binomial (NB), Zero-Inflated Poisson (ZIP), and Zero-Inflated Negative Binomial (ZINB) [21]. GLM is widely utilized for constructing accident prediction models, deviating

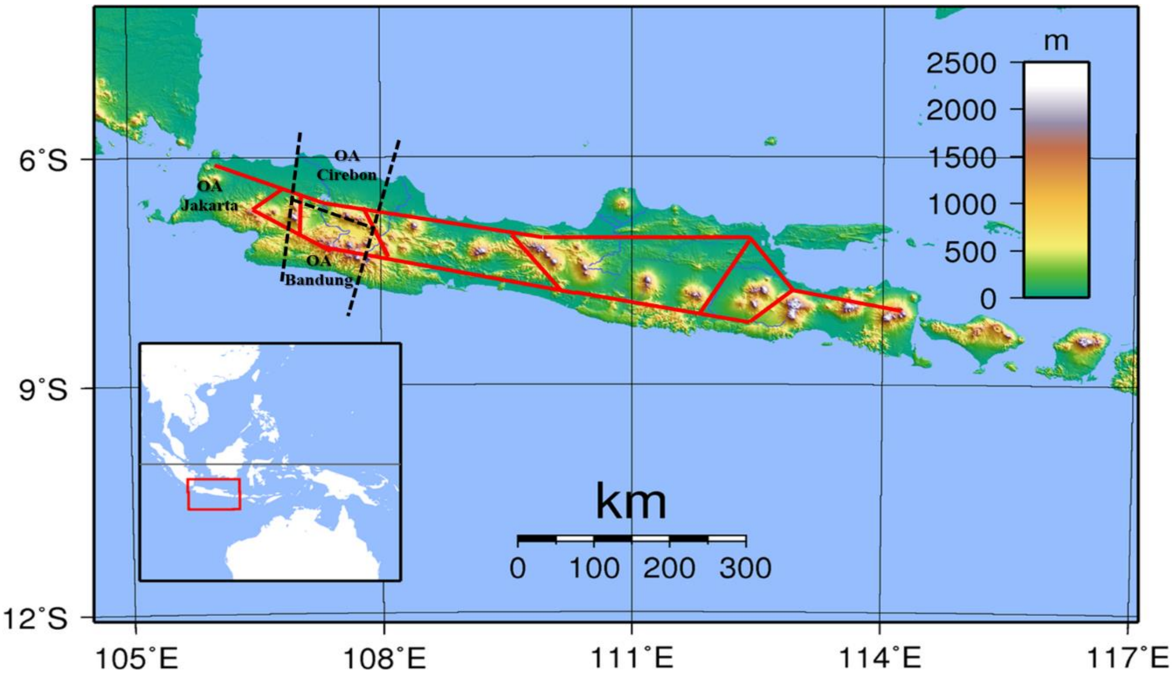


Fig. 1. Map of Research Location

from the assumption of a normal distribution [22]. Instead, it employs a set of Independent Variables (IV) to establish a relationship with the Dependent Variable (DV), which represents accident rates through link functions.

Rakhmat [2012] developed an accident prediction model using GLM, comprising PR, NB, ZIP, and ZINB for the case of the Purbaleunyi highway accidents in Indonesia. In their model, 13 parameters believed to influence accidents on the highway were considered, and the results indicated that the most appropriate model for the case was NB [23]. Silla [2012] conducted a study on railway safety in Finland using GLM. The parameters used included the number of casualties, the number of derailments and accidents, the number of level crossings, the number of accidents, and the number of vehicles passing through level crossings. The research showed that the suitable model was PR. This study focused solely on level crossings and not on lines (tracks) between stations [24]. Hence, it becomes essential to understand the impact of infrastructure on TA in Indonesia through a model that demonstrates the relationship between infrastructure factors and TA. Currently, such a model does not exist, necessitating further research to develop it. Research related to accident modelling has been conducted extensively, but primarily in the context of road accidents. In the case of TA, research is limited, making it a challenge to gather references and determine the appropriate variables and methods for creating a TA model.

This research involves identifying various infrastructure factors considered of TA in Indonesian Railway cases, followed by analysis to create a TA model. This model will illustrate the associations relationship between infrastructure factors (IV) and estimates of TA (DV). The results of this research, in the form of a TA model, will represent a significant contribution. With this model, it is expected that the relationship between infrastructure factors and TA can be understood, serving as a tool for decision-making to reduce of TA.

This journal will cover the introduction, research significance, methodology, data collection, model development analysis, and conclusions.

2. RESEARCH OBJECTIVE AND SCOPE

The objective of this research is to develop a predictive model for TA based on infrastructure variables associated with TA in Indonesia, focusing on the analysis unit of railway segments or lines (tracks) between railway stations.

The scope of this research encompasses the following aspects:

1. This research is conducted in Operational Area 1 Jakarta, Area 2 Bandung, and Area 3 Cirebon of the Indonesian Railway, with the analysis unit being the length of railway segments or lines (tracks) between railway stations.
2. A total of 379 segments are considered (229 in Operational Area 1 Jakarta, 78 in Area 2 Bandung, and 72 in Area 3 Cirebon).
3. The types of TA considered include derailments and collisions, both involving passenger and freight trains.
4. The variables under consideration include infrastructure variables such as tracks, bridges,

signals, and electrical.

5. The model development involves the use of GLM, comprising PR, NB, ZIP, and ZINB.

Fig. 2. Train Collisions

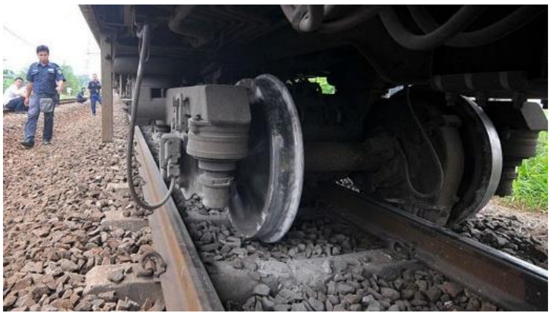
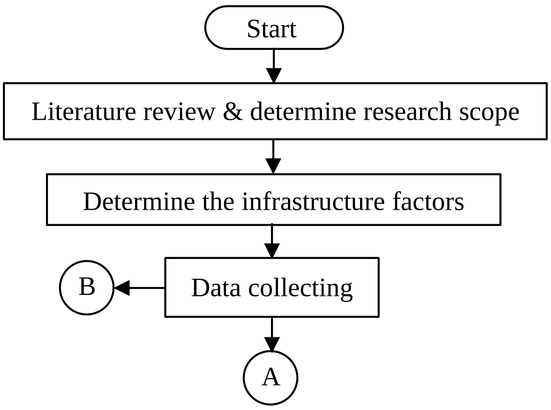


Fig. 3. Train Derailment

3. RESEARCH SIGNIFICANCE

Currently, there is no research related to a model that demonstrates the associative relationship between TA and infrastructure factors in Indonesian Railway. Hence, this represents a novelty in our study. The model is utilized to estimate TA based on infrastructure factors, which constitute the majority of TA causes. The output from this model (TA estimates) can be used to prioritize maintenance at locations with a high risk of TA, where those with the highest TA will be the primary focus for infrastructure upkeep. The findings of this research are highly beneficial for enhancing railway safety and reducing TA in Indonesia.

4. RESEARCH METHODOLOGY



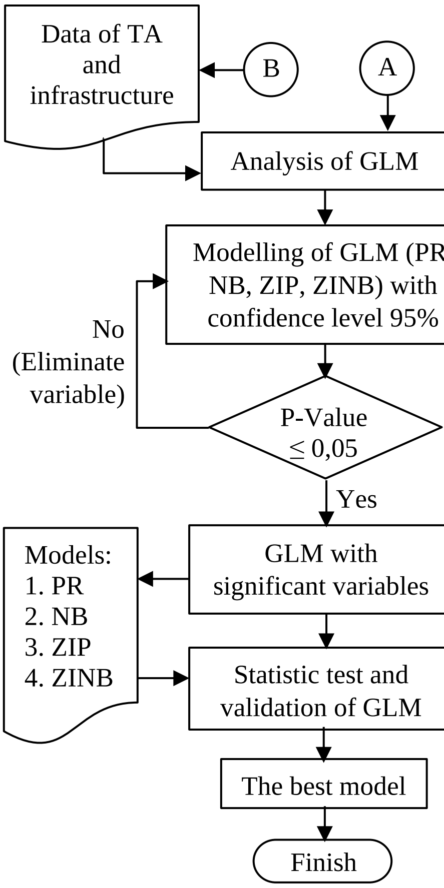


Fig. 4. Research Methodology

5. MODEL DEVELOPMENT

The model is constructed by establishing a relationship between TA and infrastructure variables (exposures and risks). Model development is using GLM. The specification of the basic model is as follows:

$$\mu = k\,(V.2190)^{\alpha} \times L^{\beta} \times \exp(\gamma_1 X_1 + \gamma_2 X_2 + \cdots) \quad (1)$$

Where:
- μ = Accident expectation
- k = Constant
- α, β, γ = Coefficient
- V = Train frequency (Train/day)
- L = Length of track line (Km)
- $X_1, X_2, ..$ = Risk variables

The model will be constructed by four distributions are PR, NB, ZIP, and ZINB, with a 95% confidence level. This means that the variable in the model must be significant, with a P-value ≤ 0,05. Variables with P-values > 0,05 will be eliminated one by one until all variables are significant.

5.1 Generalized Linear Model (GLM)

5.1.1 Poisson Regression (PR)

If Y_i is the PR distribution variable, μ_i is the expected TA, and y_i is the actual TA at location i, then its probability function is as follows:

$$\mathbf{f(Yi = y_i) = \frac{\mu_i{}^{y_i} exp^{-\mu_i}}{y_i!}} \quad (2)$$

If Xi is the independent variable and β is the constant, then the expected TA E(Yi) is:

$$\mathbf{E(Yi) = \mu_i = exp(X_i\beta) = exp\left(\textstyle\sum_1^p x_{ij}\beta_j\right)} \quad (3)$$

In this model, the variance (Var(Yi)) is equal to the mean of TA. This can be a problem because in many cases, the variance is greater than the mean, especially for data with a wide spread. Therefore, as an alternative is NB or ZIP models.

5.1.2 Negative Binomial (NB)

If Yi follows a NB distribution, where α > 0, the probability function is as follows:

$$\mathbf{f(Yi = y_i) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma(y_i+1)\,\Gamma\left(\frac{1}{\alpha}\right)}\left(\frac{1}{1+\alpha_i\mu_i}\right)^{\frac{1}{\alpha}}\left(\frac{\alpha_i\mu_i}{1+\alpha_i\mu_i}\right)^{y_i}} \quad (4)$$

The expected TA (E(Yi)) are the same as Equation (2), and the variance is given by:

$$\mathbf{Var(Yi) = \mu_i + \alpha_i\mu_i{}^2} \quad (5)$$

When α > 0, the model follows a NB distribution. In cases many of zero accidents, the PR and NB can become biased and unsuitable. As an alternative, the ZIP and ZINB models are used, specifically for data with a significant number of zero values.

5.1.3 Zero-Inflated Poisson (ZIP)

If Yi follows a ZIP distribution, the probability function is as follows:

$$\mathbf{f(Yi = y_i) = q_i + (1 - q_i)e^{-\lambda_i}} \quad (6)$$

$$\mathbf{f(Yi = y_i) = (1 - q_i)\frac{e^{-\lambda_i}\,\lambda_i{}^{y_i}}{y_i!}} \quad (7)$$

$$\mathbf{log\left(\frac{q_i}{1-q_i}\right) = \tau \textstyle\sum_1^p x_{ij}\beta_j} \quad (8)$$

$$\mathbf{\lambda_i = exp(X_i\beta) = exp\left(\textstyle\sum_1^p x_{ij}\beta_j\right)} \quad (9)$$

Where Xi is an IV, and β is a constant. The expected TA (E(Yi)) and the variance (Var(Yi)) are given by:

$$\mathbf{E(Yi) = \mu_i = (1 - q_i)\lambda_i} \quad (10)$$

$$\mathbf{Var(Yi) = \mu_i + \left(\frac{q_i}{1-q_i}\right)\mu_i{}^2} \quad (11)$$

When qi = 0, the ZIP model is identic to PR. When 0 < qi < 1, the variance of Yi will exceed its mean. Thus, the ZIP model accommodates data that is highly dispersed and includes many zero values.

5.1.4 Zero-Inflated Negative Binomial (ZINB)

If Yi follows a ZINB distribution, the probability function is as follows:

$$\mathbf{f(Yi = y_i) = q_i + (1 - q_i)\left[\frac{1}{1+\alpha_i\mu_i}\right]^{1/\alpha}} \quad (12)$$

$$\mathbf{f(Yi = y_i) = (1 - q_i)\left[\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)\left(\frac{1}{1+\alpha_i\mu_i}\right)^{1/\alpha}\left(\frac{\alpha_i\mu_i}{1+\alpha_i\mu_i}\right)^{y_i}}{\Gamma\left(\frac{1}{\alpha}\right)y_i!}\right]^{1/\alpha}} \quad (13)$$

$$\mathbf{log\left(\frac{q_i}{1-q_i}\right) = \tau \textstyle\sum_1^p x_{ij}\beta_j} \quad (14)$$

The expected TA (E(Yi)) is the same as Equation (4), and the variance (Var(Yi)) is the same as Equation (5), with the TA expectation (λ_i) given by:

$$\mathbf{\lambda_i = exp(X_i\beta) = exp\left(\textstyle\sum_1^p x_{ij}\beta_j\right)} \quad (15)$$

5.2 Test of Statistic Model

The statistical model is tested using Dispersion Parameter, Goodness-of-Fit Test, and Vuong Test. The initial step in determining the best GLM is to examine the dispersion parameter (α) of the PR model. If α = 1, the variance equals the mean, making the PR is the best model. If α > 1, over-dispersion occurs, and the best model is likely NB or ZIP. To determine whether the appropriate model is NB (indicating over-dispersion in the TA data) or ZIP (indicating over-dispersion due to many zero occurrences in the data), a Goodness-of-Fit Test is conducted.

5.2.1 Goodness of Fit Test

The Goodness-of-Fit Test conducted in this research includes the Pearson X^2 (Chi-Square Test) and the Scaled Deviance G^2 (G Square Test). The model is considered appropriate if the G^2 value ($E(G^2)$) approaches (n-p), where n is the number of data points, and p is the number of estimated parameters. Mathematically, Pearson X^2 and Scaled Deviance G^2 can be expressed as follows:

$$\mathbf{Pearson\ \chi^2 = \textstyle\sum_{i=1}^n \frac{(y_i - E[y_i])}{VAR[y_i]}} \quad (16)$$

$$\mathbf{Scaled\ Deviance\ G^2 = 2\big(l(y|y) - l(\mu|y)\big)} \quad (17)$$

Where:

l(y|y) = Log-likelihood value for the maximum model if the generated model fits the entire data

l(μ|y) = Log-likelihood value for the reduced model

If the Goodness-of-Fit Test results indicate that the ZIP model is accepted, the testing is concluded. However, if the results favour the NB model, further testing is conducted by comparing the NB and ZINB models to determine whether over-dispersion occurs in the TA data or if it's due to the prevalence of zero occurrences. This test is using Vuong Test.

5.2.2 Vuong Test

If N is total sample, $\bar{\mathrm{m}}$ is mean, Sm is standard deviation, the Vuong statistic (V) is as follows:

$$\mathbf{V = N^{1/2}\ \bar{m}/S_m} \tag{18}$$

Table 1. Vuong Statistic [25]

	t Statistic of α	
Vuong Statistic	< 2	> 2
V < -1.96	PR	NB
V > 1.96	ZIP	ZINB

After one of the best models between NB and ZINB is selected, the next step is model validation to assess its accuracy using the Mean Absolute Percentage Error (MAPE).

5.3 Validation Model

MAPE is the percentage of average absolute error. It's a statistical measurement of prediction or estimation accuracy. MAPE provides information about the extent of prediction or estimation errors compared to the actual values. A smaller percentage error value in MAPE indicates a more accurate prediction or estimation. If y_i is observed value, $\hat{y}_i$ is estimated value, and n is number of data points, the MAPE is as follows:

$$\mathrm{MAPE} = \sum_{t=1}^{n} \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| \times 100\% \tag{19}$$

MAPE values can be interpreted in four categories: if MAPE < 10%, the model is very good; if MAPE ranges from 10-20%, the model is good; if MAPE ranges from 20-50%, the model is reasonable; and if MAPE > 50%, the model is inaccurate or unsuccessful [26].

6. DATA COLLECTING

The research use data from OA Jakarta, Bandung, and Cirebon, with the unit of analysis is the segments between two railway stations. The variables and data used are described in Table 2. The study covers a total of 379 segments (229 segments from OA Jakarta, 78 segments from OA Bandung, and 72 segments from OA Cirebon). There are 20 variables considered to influence TA, resulting in a total of 7.580 data.

Table 2. Data Variable

No	Notation	Variable	Unit	Variable Role
1	X_1	Train frequency	Train/ day	Exposure
2	X_2	Track length	Km	Exposure
3	X_3	Train speed	Km/ hour	Risk
4	X_4	Track length with rail type of R54	Km	Risk
5	X_5	Track length with concrete sleeper	Km	Risk
6	X_6	Turnouts number	Unit	Risk
7	X_7	Level crossing number	-	Risk
8	X_8	Signalling type	-	Risk
9	X_{91}	Length of bridge with the age >100 years	Km	Risk
10	X_{92}	Length of bridge with the age <100 years	Km	Risk
11	X_{101}	Length of curve with the radius ≤250 m	Km	Risk
12	X_{102}	Length of curve with the radius 250 m to 500 m	Km	Risk
13	X_{103}	Length of curve with the radius 500 m to 1000 m	Km	Risk
14	X_{104}	Length of curve with the radius >1000 m	Km	Risk
15	X_{11}	Number of vulnerable areas	-	Risk
16	X_{12}	Length of catenary line	Km	Risk
17	X_{131}	Length of track with the slope ≤10 permille for incline and ≤-10 permille for descendant	Km	Risk
18	X_{132}	Length of track with the slope >10 permille for incline and >-10 permille for descendant	Km	Risk
19	X_{14}	Double or single track	-	Risk
20	X_{15}	Passing tonnage	Million ton / year	Risk

7. ANALYSIS AND DISCUSSION

7.1 Model Development

The model is constructed using Equation 1 and analyzed using the PR, NB, ZIP, and ZINB models with a 95% confidence level through multiple iterations to identify significant variables (P-value ≤ 0,05). The analysis results in Table 3.

Table 3. Modeling Results

Model	Equation
PR	$\mu = 1{,}654\ (2190.X_1)^{-0{,}021} \times X_2^{-0{,}149} \times \exp(-0{,}027.X_3 + 0{,}288.X_4 - 0{,}763.X_{102} - 1{,}227.X_{103} + 0{,}313.X_{11} - 0{,}267.X_{12} - 1{,}411.X_{14})$ (20)
NB	$\mu = 0{,}927\ (2190.X_1)^{-0{,}021} \times X_2^{0{,}121} \times \exp(-0{,}032.X_3 - 1{,}591.X_{103} + 0{,}395.X_{11} - 0{,}308.X_{12} - 2{,}163.X_{14})$ (21)
ZIP	$\mu = 0{,}383\ (2190.X_1)^{-0{,}248} \times X_2^{0{,}362} \times \exp(-0{,}035.X_3 - 0{,}364.X_{101} - 0{,}355.X_{12} + 0{,}272.X_{15})$ (22)
ZINB	$\mu = 0{,}383\ (2190.X_1)^{-0{,}248} \times X_2^{0{,}036} \times \exp(-0{,}035.X_3 - 0{,}364.X_{101} - 0{,}355.X_{12} + 0{,}272.X_{14})$ (23)

7.2 Statistic Model Test

Based on the analysis for the PR model, dispersion value (α) is 1,44. This indicates over-dispersion in the data, suggesting that the PR model is not suitable, and the more appropriate models could be NB or ZIP. To determine whether the suitable model is NB (over-dispersion occurs in the TA data) or ZIP (over-dispersion happens due to the abundance of zero occurrences in the TA data), a Goodness-of-Fit Test was conducted, using Equations 16 and 17 as shown in Table 4.

Tabel 4. Goodness-of-Fit Test

Model	NB	ZIP
Data Count	379	379
Parameter Count	7	6
Degrees of Freedom	371	372
X^2	315,41	267,59
G^2	375,53	235,66
$X^2_{(0,95;7)}$, $X^2_{(0,95;6)}$	416,91	417,85
	Accepted	Rejected

Based on the analysis results presented in Table 4, it was determined that the suitable model for this case is the NB model. Subsequently, a Vuong Test was conducted on the NB and ZINB models to ascertain whether over-dispersion occurred in the TA data or if it was due to the abundance of zero occurrences in the TA data. Analysis yielded a t-statistic value of 2,44, which is greater than 2,00, indicating that the suitable model could be either NB or ZINB. Additionally, a Vuong Statistic value of -10,94 < -1,96, suggests that the most appropriate model for the TA case is the NB model. Further validation of the model was carried out using MAPE.

7.3 Model Validation

Model validation was using MAPE as per Equation 19. For the NB model, a MAPE value of 20,10% was obtained, falling within the range of 20% to 50%. This percentage error, indicated by the MAPE value, demonstrates that the TA estimation results using this equation are reasonable.

7.4 Discussion

Based on the modeling and analysis results, out of the 20 infrastructure variables considered, only seven significantly influence train accidents in Indonesia. These variables are the frequency of trains, the length of railway tracks, train speed, the length of curves with a radius of 500 m to ≤ 1000 m, the number of vulnerable locations, the length of catenary networks, and single-track lines.

1. The frequency of trains has a positive impact on TA, meaning that the higher the frequency of trains, the higher the TA.
2. The length of railway tracks has a positive impact on TA, meaning that the longer the railway tracks, the higher the TA.
3. Train speed has a negative impact on TA, meaning that the higher the train speed, the lower the TA. This is related to the frequency of railway track maintenance. High-speed railway tracks will certainly affect the increase in railway track maintenance.
4. The length of curves with a radius of 500 m to ≤ 1000 m has a negative impact on TA, meaning that the longer the curve, the lower the TA. Curves with this radius are safer to use compared to curves with a radius < 500 m.
5. The number of vulnerable areas has a positive impact on TA, meaning that the more vulnerable areas, the higher the TA.
6. The length of catenary networks has a negative impact on TA, meaning that the longer the catenary networks, the lower the TA. Generally, catenary networks are associated with commuter transportation, where the railway tracks are not long, thereby reducing the TA.
7. Single-track lines have a positive impact on TA; single-track lines can influence an increase in TA compared to double-track lines, and this is related to their operational pattern.

These variables will affect the expected number of accidents. The higher the TA on a railway track, the more dangerous the track, requiring attention and intervention.

8. CONCLUSION

1. The suitable GLM for the railway accidents model in Indonesian Railway case is the NB.

2. The factors associated with TA in Indonesia as per this model encompass:
   a. Train frequency (X_1 in trains/day) as exposure. Train frequency plays a crucial role in shaping the likelihood of TA. A higher frequency, or the greater number of trains traversing a specific rail segment per day, escalates the probability and risk of TA.
   b. Length of track (X_2 in km) as exposure. The extent of the railway track directly affects TA. Longer tracks translate to extended travel durations for trains, which can also influence the probability of TA.
   c. Train speed (X_3 in km/hour) as a risk factor. Train speed significantly affects the incidence of TA. Elevated train speeds on a railway track elevate the likelihood and risk of TA.
   d. Length of the curve with a radius of 500 m to ≤ 1000 m (X_{103} in km) as a risk factor. The length of a curve with a radius ranging from 500 m to ≤ 1000 m exerts an influence on TA. Considering the available data, this radius category is prevalent within the study area. The greater the number or length of curves within this range, the higher the risk of TA. Mitigating this risk to railway safety necessitates the reduction of such curves.
   e. Number of vulnerable areas (X_{11} in points) as a risk factor. The quantity of vulnerable areas is a pertinent factor in determining TA. An escalation in the count of vulnerable areas correlates with an augmented risk of TA. To alleviate risks to railway safety, reducing the number of vulnerable areas becomes imperative.
   f. Length of catenary line (X_{12} in km) as a risk factor. The extent of the catenary line directly influences TA. A longer catenary line along a track line amplifies the risk of TA. The length of the catenary line typically aligns with the length of the track line in sections featuring catenary line, such as in OA Jakarta.
   g. Single track (X_{14}) as a risk factor. Whether a railway track is single or double-track has a significant impact on the occurrence of TA. Single-track sections present a heightened risk of TA compared to double-track sections, primarily due to operational patterns. Operational procedures on single tracks allow trains to cross on the same track, substantially heightening the risk of TA.

9. ACKNOWLEDGMENTS

The author would like to thank PT Kereta Api Indonesia (Persero) for supporting this research.

10. REFERENCE

[1] Profillidis V. A, Botzoris G. N, A Comparative Analysis of Performances of Econometric Fuzzy and Time-Series Models for The Forecast of Transport Demand, IEEE Internation Conference on Fuzzy Systems, 2007.

[2] Jia C, Xu W, Wang H, Study of Management Information System of Railway Permanent Way Safety Risks and Comprehensive Evaluation, Procedia Engineering, 15, 1293-1297, 2011.

[3] Pasha J, Dulebenets M. A, Abioye O. F, Kavoosi M, Moses R, Sobanjo J, Ozguven E. E, A Comprehensive Assessment of the Existing Accident and Hazard Prediction Models for the Highway-Rail Grade Crossings in the State of Florida. Sustainability-Switzerland, 12(10), 2020.

[4] Blagojević A, Stević Ž, Marinković D, Kasalica S, Rajilić S, A Novel Entropy-Fuzzy Piprecia-DEA Model for Safety Evaluation of Railway Traffic, Symmetry, 12(9), 2020.

[5] Wemaker W. D, Jack A, Schmid F, Establishing the Relationship Between Railway Safety and Operational Performance, International Journal of Transport Development and Integration, 2(1), 98-114, 2018.

[6] Shabani S, Zakeri J. A, Tabar V. R, Ghanbarpour A, A Risk Based Model for the Switch Derailmaent Using Bayesian Network, IJRR, Vol 6, No 2, 103-112, 2019.

[7] Lu C, Cai C, Challenges and Counter Measures for Construction Safety During the Sichuan-Tibet Railway Project. Engineering, 5(5), 833-838, 2019.

[8] Ishida H, Miyamoto T, Maebashi E, Doi H, Iida K, Furukawa A, Safety Assessment For Flange Climb Derailment of Trains Running at Low Speeds on Sharp Curves, Quarterly Report of RTRI, 47(2), 65–71, 2006.

[9] Amanzholovich R. M, Mukhamedievna R. G, Ahmetovna S. Z, Modification of Concrete Railway Sleepers and Assessment of Its Bearing Capacity, International Journal of Geomate, Vol 20, Issue 77, 2021, pp.40-48.

[10] Phanyakit T, Satiennam T, Fuzzy Multi-Attribute Decision Making for the Selection of a Suitable Railway Track Maintenance Plan: A Case Study in Thailand, International Journal of Geomate, Vol 17, Issue 60, 2019, pp.96-104.

[11] Rad M. A, Lefsrud L. M, Hendry M. T, Application of Systems Thinking Accident Analysis Methods: A Review for Railways, Safety Science, Vol 16, 2023.

[12] Kim S. H, Chung S. B, Song K. H, Chon K. S, Development of An Accident Prediction Model Using GLIM (Generalized Log-linear Model) and EB Method: A Case of Seoul, Journal of the Eastern Asia Society for Transportation Studies, 6, 3669-3682, 2005.
[13] Rakha H, Arafeh M, Salam A. G, Guo F, Flintsch A. M, Linear Regression Crash Prediction Models: Issues And Proposed Solutions Virginia Transportation Research Council, 510401, 2010.
[14] Harnen S, Umar R. S, Wong S. V, Hashim W. I, Development of Prediction Models for Motorcycle Crashes at Signalized Intersections on Urban Roads in Malaysia. Journal of Transportation and Statistics, 7(2–3), 27–39, 2005.
[15] Ackaah W, Salifu M, Crash Prediction Model for Two-Lane Rural Highways in the Ashanti Region of Ghana, IATSS Research, 35(1), 34-40, 2011.
[16] Jovanis P. P, Chang H. L, Modeling the Relationship of Accidents to Miles Traveled. Transportation Research Record, 1986, pp. 42-51.
[17] Maycock G, Hall R. D, Accidents at 4-Arm Round About, US Transport and Road Research Laboratory, 1984.
[18] Granà A, How to Formulate an Accident Prediction Model for Urban Intersections, WIT Transactions on the Built Environment, 107, 589–597, 2009.
[19] Ma W, Yuan Z, Analysis and Comparison of Traffic Accident Regression Prediction Model, Icectt, 364-369, 2018.
[20] Yannis G, Dragomanovits A, Laiou A, Richter T, Ruhl S, La T. F, Domenichini L, Graham D, Karathodorou N, Li H, Use of Accident Prediction Models in Road Safety Management-An International Inquiry, Transportation Research Procedia, 14, 4257-4266, 2016.
[21] Lord D, Persaud B, Estimating the Safety Performance of Urban Road Transportation Networks, Accident Analysis and Prevention, 36(4)(416), 609-620, 2004.
[22] Machsus, Sulistio H, Wicaksono A, Djakfar L, The Effect of Access Points on Motorcycle Accident Rates on Surabaya Arterial Roads, Australian Journal of Basic and Applied Sciences, 38-43, 2014.
[23] Rakhmat L. A, Kusumawati A, Frazila R. B, Hendarto S, Pengembangan Model Prediksi Kecelakaan Lalu Lintas pada Jalan Tol Purbaleunyi, Jurnal Teknik Sipil, 19(3), 277, 2012.
[24] Silla A, Kallberg V. P, The Development of Railway Safety in Finland, Accident Analysis and Prevention, 45, 737–744, 2012.
[25] Kumara S. S. P, Chin H. C, Modeling Accident Occurrence at Signalized Tee Intersections With Special Emphasis on Excess Zeros. Traffic Injury Prevention, 4(1), 53-57, 2003.
[26] Lewis C. D, Industrial and Business Forecasting Methods, London: Butterworths, 1982.