Factors Affecting Traffic Accident Severity on Expressways: Evidence from Vietnam

Tien Dung CHU[a], Xuan Quy LE[b]

[a, b] *Department of Highway and Traffic Engineering, Faculty of Civil Engineering, University of Transport and Communications, No. 3 Cau Giay Street, Lang Thuong Ward, Dong Da District, Hanoi, Vietnam*
[a] *E-mail: dungchu@utc.edu.vn*
[b] *E-mail: lexuanquy@utc.edu.vn*

Abstract: This study analyzes 108 road traffic accidents (RTAs) on Vietnam's expressway network, focusing on key accident characteristics, including time of occurrence, vehicle involvement, and contributing factors. Results indicate that most accidents occurred between 0:00–12:00, with the highest proportion between 12:00–18:00. Failure to maintain a safe following distance was the leading cause, while head-on collisions were exclusive to two-lane expressways without median barriers. Severity analysis, based on the Severity Index (SI) and Ordered Probit (OP) models, shows that higher posted speed limits, vehicle types (coaches and trucks), and collision types (head-on, rollover, run-off-road) increase accident severity. Given the study's limited dataset (two years), future research should expand data collection and apply advanced models. The findings highlight the need for stricter traffic enforcement, infrastructure improvements, and public awareness campaigns to enhance expressway safety.

Keywords: Road Traffic Accidents, Severity Index, Ordered Probit, Expressway, Safety

1. INTRODUCTION

Road traffic accidents (RTAs) remain a critical global concern. According to the World Health Organization (WHO, 2023), an estimated 1.19 million fatalities occurred due to RTAs in 2021, representing a 5% decline compared to 2010. The global fatality rate stood at 15 per 100,000 people, marking a 16% reduction over the same period. This modest decrease happened despite the global motor vehicle fleet more than doubling, road networks expanding significantly, and the world's population increasing by nearly one billion. However, the target of halving traffic fatalities set under the Decade of Action for Road Safety 2011–2020 was not met. At the current pace, achieving the 50% reduction goal outlined in the Global Goals for Sustainable Development remains unlikely.

Furthermore, WHO's 2023 report highlights that road traffic fatalities and injuries continue to present a major public health and development challenge worldwide. As of 2019, RTAs were the leading cause of death among individuals aged 5 to 29 and ranked 12th across all age groups. Nearly two-thirds of fatalities involved individuals of working age (18–59 years), resulting in significant health, economic, and social consequences. Motorcyclists and users of two- or three-wheeled motorized vehicles accounted for 30% of deaths, followed by occupants of four-wheeled vehicles at 25%. Pedestrians represented 21%, cyclists 5%, and those traveling in heavy trucks or vehicles carrying more than ten passengers made up 19%. Approximately 90% of traffic-related deaths occurred in low- and middle-income countries, with the highest risk found in low-income nations. In terms of regional distribution, Southeast Asia recorded the highest share of fatalities (28%), followed by the Western Pacific (25%), Africa (19%), the Americas (12%), the Eastern Mediterranean (11%), and Europe (5%).

Vietnam, located in Southeast Asia, is among the countries with the highest share of traffic fatalities in the region. It ranks second after Thailand in terms of RTAs and related fatalities. According to the Vietnam National Traffic Safety Committee (NTSC, 2023), there were 22,067 traffic accidents in 2023, resulting in 11,628 deaths and 15,292 injuries. These statistics highlight the urgent need for comprehensive measures to reduce RTAs. One critical area requiring immediate attention is the occurrence of RTAs on Vietnam’s expressway network.

Vietnam’s expressway system began construction in late 1998, and by 2025, its total length is estimated to reach approximately 3,000 km. The network is expected to expand to approximately 5,000 km by 2030. Some expressways have 4 to 6 lanes, fully equipped with median barriers and emergency stopping lanes, with posted speed limits of 100 to 120 km/h. However, due to budget constraints, many expressways are still being developed in phases. In the initial phase, some routes have only 2 to 4 lanes, lacking emergency stopping lanes and/or median barriers, with speed limits set at 80 to 90 km/h (see Figure 1).

Initial-phase expressway investment includes a 2×3.5m carriageway, a 2×2m treated shoulder, and no median barrier.

Initial-phase -phase expressway investment includes a 4×3.5m carriageway with a median barrier but without emergency stopping lanes.

Full-phase expressway investment includes a 4×3.75m carriageway with a median barrier and 2×3m emergency stopping lanes

Full-phase expressway investment includes a 6×3.75m carriageway with a median barrier and 2×3m emergency stopping lanes

Figure 1. Typical cross-section types in Vietnam's expressway network

As reported by the Ministry of Public Security (MPS, 2024), in the first seven months of 2024, 112 RTAs occurred on Vietnam's expressway network (0.79%), resulting in 46 fatalities (0.75%) and 82 injuries (0.75%). Although these figures represent a small proportion, most RTAs were severe or extremely serious.

To reduce RTAs and their severity, this paper utilizes data from 108 RTAs collected on Vietnam's expressway network to identify factors influencing RTA severity using the Ordered

Probit (OP) model. While RTA databases typically classify severity based on the most severe outcome within an incident, RTAs can involve multiple severity levels and more than one victim, meaning the recorded severity level may not fully reflect the overall impact. Additionally, previous studies have commonly treated severity levels independently (e.g., fatal, serious injury, minor injury). However, we contribute to the literature by incorporating a Severity Index, which applies a severity weighting system to account for all severity levels within an RTA.

The remainder of this paper is structured as follows: Section 2 presents the Literature Review, followed by Section 3, which discusses Data Collection. Section 4 covers Model Development, and finally, Section 5 concludes the paper with key findings and recommendations.

2. LITERATURE REVIEW

Predicting traffic accident severity has long been a crucial focus in transportation safety research. Studies on this topic can be divided into two main approaches: Machine Learning and the OP model. The following sections provide a brief discussion of each approach.

2.1 Machine Learning Approach

Pei *et al.* (2024) introduced an interpretable framework integrating Extreme Gradient Boosting (XGBoost), SHapley Additive exPlanations (SHAP), and Attention-based Interactive Spatial-Temporal Graph Convolutional Network (AISTGCN), achieving high prediction accuracy while providing meaningful insights into key influencing variables. Similarly, Dong *et al.* (2024) refined feature selection and hyperparameter optimization in a CatBoost model, reaching a prediction accuracy of 96.63% on U.S. accident data. These studies highlight the potential of machine learning to improve accident severity forecasting through enhanced feature engineering and model interpretability.

Recognizing the multifaceted nature of accident severity, Wang *et al.* (2024) categorized influencing factors into human and non-human domains, applying a two-level fuzzy comprehensive evaluation method to improve predictive precision for freeway accidents. Meanwhile, Alhaek *et al.* (2024) leveraged deep learning by integrating Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) networks, capturing spatial-temporal dependencies and demonstrating superior accuracy in severity classification.

Several studies have also emphasized the role of driver behavior, road conditions, and environmental factors. Cicek *et al.* (2023) employed decision trees, neural networks, and Shapley values to identify key contributors such as seatbelt usage, alcohol consumption, and speed violations. Pérez-Sala *et al.* (2024) developed a convolutional neural network-based model that underscored the importance of data preprocessing and model scalability in severity prediction.

Despite their predictive power, machine learning models often face challenges in interpretability. Li *et al.* (2024) addressed this by integrating a multinomial logit model with adaptive sparse group lasso, effectively filtering out irrelevant factors while identifying critical determinants of maritime accident severity. Similarly, Sufian *et al.* (2024) combined machine learning with econometric techniques such as the Generalized Method of Moments (GMM) and Autoregressive Integrated Moving Average (ARIMA), incorporating Explainable AI (XAI) to enhance model transparency.

2.2 Ordered Probit Model Approach

Wang *et al.* (2022) examined maritime accident injury severity using a zero-inflated OP (ZIOP) model, effectively distinguishing between injury-free and injury-prone states. Their findings emphasized that factors such as capsizing, hull damage, adverse sea conditions, and crew characteristics significantly influenced injury severity. The ZIOP model outperformed traditional OP approaches by better capturing the dual nature of accident severity outcomes.

Focusing on driver characteristics, Nickkar *et al.* (2024) applied a heteroskedastic OP (HETOP) model to account for variability in injury severity distribution. Their study revealed that male drivers and older individuals were more susceptible to severe injuries, with environmental factors such as lighting conditions and road surfaces further influencing severity levels.

In another application, Fang *et al.* (2024) explored risk factors in angle crashes using a random parameter bivariate OP (RPBOP) model, which uncovered significant roles of driver demographics, road characteristics, and crash conditions. Their analysis highlighted unobserved heterogeneity in factors like driver age, collision points, and road grade, demonstrating the value of Bayesian inference and marginal effects analysis in severity estimation.

To address latent heterogeneity, Karabulut and Ozen (2023) employed a latent class OP (LCOP) model, categorizing crashes into four clusters before applying the probit framework. Their results identified motorcycles, fixed objects, and run-off-road crashes as major contributors to severe injuries. Behavioral factors such as speeding, alcohol impairment, and traffic violations also significantly increased injury severity.

Environmental conditions have also been a focus of OP models. Hyodo and Hasegawa (2021) investigated accident severity in cold and snowy weather, revealing that icy and snowy roads generally resulted in less severe crashes, whereas low visibility and extreme temperatures increased the risk of severe multi-vehicle accidents.

Expanding on heterogeneity modeling, Atombo *et al.* (2023) developed random-parameter OP (RPOP) models with heterogeneity in means and variances to analyze both motorized and non-motorized crashes. Their study found that old-age occupants, high-speed crashes, and weekend accidents significantly increased injury severity in motorized crashes, while male gender, lack of helmet use, and urban settings were key contributors to severe outcomes in non-motorized crashes. These findings underscore the necessity of tailored countermeasures for different road user groups.

3. DATA COLLECTION

The dataset used in this study consists of 108 RTAs on Vietnam's expressway network from January 1, 2023, to February 20, 2025. The data were collected and compiled from online news reports on traffic accidents. For each accident, key details were recorded, including the expressway name, date and time of occurrence, number of fatalities, number of injuries (serious and light), number of vehicles involved, and the cause of the accident. These details were then cross-verified with other sources to ensure data accuracy. Additional information, such as the number of lanes, lane width, presence of a median barrier or emergency stopping lanes, and posted speed limits, was obtained from design consultants and management agencies. Some detailed information on the RTA dataset is summarized in Table 1.

Table 1. Descriptive statistics of RTA data

Type	Investment phase	Number of lanes	Lane width	Posted speed limits	Emergency stopping lanes	Median barrier	RTAs	Fatalities	Serious injuries	Light injuries
1	Initial phase	2	3.5	80	No*)	No	30	39	66	57
2	Initial phase	4	3.5	90	No	No	18	12	3	0
3	Full phase	4	3.75	100	Yes	Yes	54	12	60	39
4	Full phase	4	3.75	120	Yes	Yes	3	3	0	3
5	Full phase	6	3.75	120	Yes	Yes	3	6	12	18
						Total	**108**	**72**	**141**	**117**

Note: Emergency stopping lanes do not exist, but a 2x2m treated shoulder is available

Some key characteristics of the dataset are shown in Figures 2 to 5. Figure 2 indicates that more than half (55%) of the accidents occurred between 0:00 - 12:00, with 30% happening between 0:00 - 6:00 and 25% between 6:00 - 12:00. The time period from 12:00 - 18:00 accounted for the highest percentage at 31%, followed by 18:00 - 24:00 with 14%.

Regarding the day of occurrence (Figure 3), 65% of accidents happened on weekdays (Monday to Friday), while the remaining 36% occurred on weekends (Saturday and Sunday). However, when considering accident frequency per day (accidents/day), the weekend had a higher accident rate at 59%, compared to 41% on weekdays.

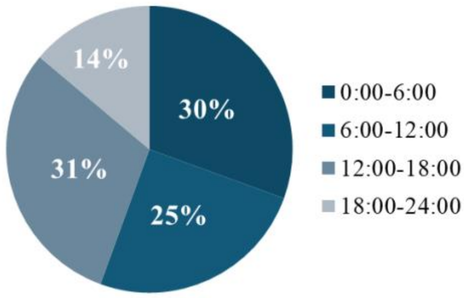

Figure 2. Time of occurrence of the RTA

a) Proportion of RTA on weekend and weekday

a) Proportion of RTA on weekend and weekday (weighted by number of days)

Figure 3. Date of occurrence of the RTA

As shown in Figure 4, failure to maintain a safe following distance was the leading cause of accidents, accounting for 62%. Head-on collisions followed at 14%, all of which occurred on two-lane expressways without a median barrier. Next were accidents involving stationary vehicles on expressways (12%) and those caused by rollovers or vehicles running off the road

(12%). Notably, all accidents involving stationary vehicles occurred on phased-investment expressways without emergency stopping lanes.

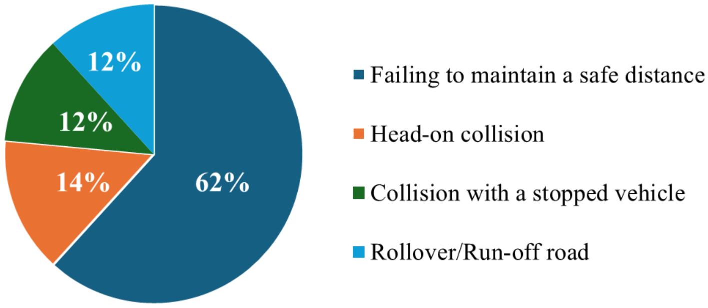

Figure 4. Primary causes of RTAs

Regarding vehicle types involved in accidents (Figure 5), coaches were involved in 39 out of 108 cases, while trucks were involved in 72 out of 108 cases. Trucks and coaches typically carry heavy loads or a large number of passengers. In the event of an accident, they often result in significant casualties due to the high number of passengers on board. Additionally, smaller vehicles are generally at a disadvantage when colliding with or being struck by these larger vehicles.

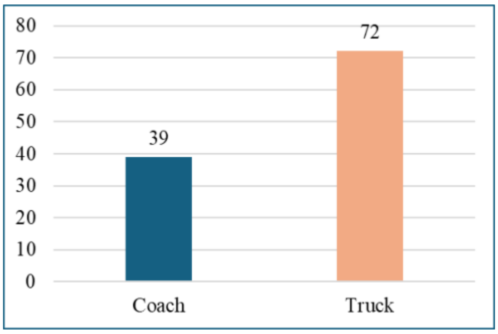

Figure 5. Vehicles involved in RTAs

4. MODEL DEVELOMENT

4.1 Traffic Accident Severity

RTAs vary in severity based on the level of injury and property damage. The U.S. Highway Safety Manual (HSM, 2010) defines accident severity based on the most serious injury sustained in an incident. While a single accident may result in multiple injuries of varying severity, its classification is determined by the most severe outcome. The KABCO scale is commonly used to categorize accident severity into five levels. A fatal injury (K) occurs when the accident results in death. An incapacitating injury (A) refers to severe injuries that prevent the affected individual from performing normal activities such as walking or driving. A non-incapacitating evident injury (B) includes visible injuries, such as bruises or cuts, that do not cause major disability. A possible injury (C) is one that is reported or suspected but not visibly apparent at the scene. Lastly, a property damage only (PDO) accident (O) involves no injuries, only vehicle or infrastructure damage. This classification system is widely used in accident data analysis to assess safety risks and prioritize road safety improvements.

However, a key limitation of the KABCO scale is that it only considers the most severe outcome of an accident, overlooking the full extent of injuries sustained. This can lead to an underestimation of the total impact, as accidents with multiple serious but non-fatal injuries may be classified at the same level as those with minor injuries. Additionally, the subjectivity in classifying injuries, particularly for possible injuries (C) and non-incapacitating injuries (B), can introduce inconsistencies across different jurisdictions and reporting agencies. The lack of standardized medical assessment further contributes to potential misclassification. Moreover, KABCO does not directly account for long-term consequences, such as disability or medical costs, limiting its ability to fully capture accident severity.

Recognizing these limitations, researchers have explored alternative methods, such as severity weighting systems. The current study adopted the Severity Index (SI) approach proposed by Geurts et al. (2004). Their research focused on identifying and ranking hazardous road locations, commonly referred to as black spots, in Flanders, Belgium. Using historical accident data from 1997 to 1999, they analyzed 1,014 sites classified as dangerous based on recorded incidents. To prioritize these locations, they applied a severity weighting system, assigning values of 1 for light injuries, 3 for serious injuries, and 5 for fatal injuries. This approach provides a more quantitative and structured assessment of accident severity, addressing the limitations of categorical classifications like KABCO.

Specifically, the Severity Index (SI) for each accident can be calculated as follows:

$$SI = N_{light} \times 1 + N_{serious} \times 3 + N_{fatal} \times 5 \quad (1)$$

As shown in Figure 6, the Severity Index (SI) is categorized into three levels: (1) Low (SI ≤ 3), (2) Moderate (3 < SI ≤ 8), and (3) High (SI > 8). Notably, SI = 0 indicates that the accident resulted in property damage only. This categorical variable of SI is later used in the OP model as a dependent variable to identify the factors affecting the severity level of traffic accidents on expressways in Vietnam. A brief summary of the OP model will be provided in the following section.

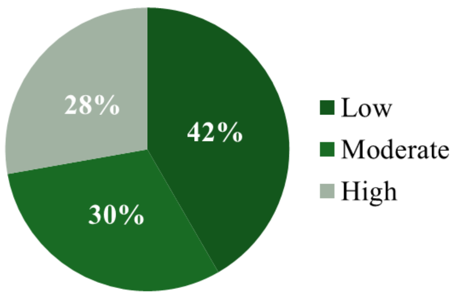

Figure 6. The proportion of each category in the Severity Index (SI)

4.2 Ordered Probit Model

As mentioned in section 2.2, the OP model is widely used to analyze highway accident severity due to its ability to handle ordinal dependent variables. Accident severity categories—such as minor injury, severe injury, and fatal accidents—have a meaningful ranking but are not measured on an interval scale. Unlike linear regression models, OP captures the categorical and hierarchical nature of severity outcomes, providing more reliable estimates of influencing factors.

OP assumes that an unobserved latent severity level influences observed categories, enabling probabilistic estimation based on explanatory variables. It effectively addresses issues such as non-linearity and threshold effects in severity classification. Compared to the Ordered Logit model, OP performs better when the assumption of normally distributed error terms aligns with the data.

This study applies the OP model, which is suitable when the dependent variable takes ordered values such as 1, 2, 3, ..., J. It is derived from an unobserved latent variable, formulated as follows:

$$y_i^* = \beta x_i' + \varepsilon_i \tag{2}$$

In this equation:

- β is a vector of unknown parameters,
- x_i' is a vector of explanatory variables,
- ε is the error term.

The observed dependent variable takes discrete values 1, 2, 3,…, J based on threshold values as follows:

$$y_i = \begin{cases} 1 \text{ if } \mu_0 < y_{\mathrm{i}}^* \leq \mu_1 \\ 2 \text{ if } \mu_1 < y_{\mathrm{i}}^* \leq \mu_2 \\ \quad \dots \\ \mathrm{J} \text{ if } \mu_{\mathrm{J}-1} < y_{\mathrm{i}}^* \leq \mu_{\mathrm{J}} \end{cases} \tag{3}$$

Here, μ_{J} are unknown thresholds that need to be estimated, and they must always satisfy the condition $\mu_{\mathrm{J}-1} < \mu_{\mathrm{J}}$. It is noted that $\mu_0 = -\infty$ và $\mu_{\mathrm{J}} = +\infty$. The probability that $y_i = j$ is given by:

$$\Pr(y_i = j) = \Pr(\mu_{\mathrm{j}-1} < y_{1\mathrm{i}}^* \leq \mu_{\mathrm{j}}) \tag{4}$$

Since ε follows a standard normal distribution, the probability can be computed as:

$$\Pr(y_{\mathrm{i}} = j) = \Phi(\mu_{\mathrm{j}} - \beta x_{\mathrm{i}}') - \Phi(\mu_{\mathrm{j}-1} - \beta x_{\mathrm{i}}') \tag{5}$$

Where Φ represents the cumulative distribution function (*CDF*) of the standard normal distribution. The log-likelihood function for an individual observation can be expressed as:

$$\ln(L_{\mathrm{i}}) = \sum_{i=1}^{J} \mathrm{I}(y_{\mathrm{i}} = j) \ln(\Pr(y_{\mathrm{i}} = j)) \tag{6}$$

Assuming that all observations are independent, the log-likelihood function for the entire sample (N) is:

$$\ln(L_{\mathrm{i}}) = \sum_{\mathrm{i}=1}^{\mathrm{N}} \sum_{i=1}^{J} \mathrm{I}(y_{\mathrm{i}} = j) \ln(\Pr(y_{\mathrm{i}} = j)) \tag{7}$$

The unknown parameters (β, μ_j) are estimated using the Maximum Likelihood Estimation (*MLE*) method with the *oprobit* command in Stata.

4.3 Variables Considered in The Model

This paper considers several variables, including the time and day of accident occurrence, the posted speed limit, the type of vehicle involved in the accident (coach, truck, or passenger car), the cause of the accident (failing to maintain a safe distance, wrong-way driving, unsafe stopping and parking, etc.), expressway characteristics (number of lanes, presence of a median barrier or emergency lane), and the type of accident (head-on collision, rear-end collision, rollover, run-off-road).

The time of accidents is categorized into four dummy variables (0:00–6:00, 6:00–12:00, 12:00–18:00, and 18:00–24:00), while the day of the accident is classified into two dummy variables (weekend and weekday).

To address multicollinearity, the Variance Inflation Factor (VIF) and the Pearson correlation coefficient between independent variables were calculated. Generally, a VIF above 5 or an absolute Pearson correlation coefficient greater than 0.7 indicates high multicollinearity, and such variables are excluded from the OP model. The final list of independent variables included in the model is presented in Table 2.

Table 2. Descriptive statistics of independent variables

Independent variables	Mean	Std.dev
Time of accident, dummy: 12:00 - 18:00 = 1, otherwise = 0	0.31	0.46
Date of accident, dummy: Weekend = 1, Weekday = 0	0.36	0.48
Failing to maintain a safe distance, dummy: 1 if true, 0 otherwise	0.58	0.50
Posted speed limit	93.89	10.66
Type of vehicle involved, dummy: Coach = 1, otherwise = 0	0.36	0.48
Type of vehicle involved, dummy: Truck = 1, otherwise = 0	0.67	0.47
Head-on collision, dummy: 1 if true, 0 otherwise	0.14	0.35
Rollover/Run-off road, dummy: 1 if true, 0 otherwise	0.11	0.32
Number of involved vehicles	2.50	0.99

4.4 Model Result and Discussion

The model estimation results are displayed in Table 3. Importantly, six out of the nine independent variables are statistically significant at the 0.05 level, while the other three are significant at the 0.1 level. The results demonstrate a strong fit to the data, as indicated by the rho-squared (0.77) and adjusted rho-squared (0.68) values. Typically, a rho-squared value above 0.2–0.3 suggests a reasonable model fit in discrete choice modeling, so the obtained values indicate a high level of explanatory power. The adjusted rho-squared, which accounts for model complexity, remains strong at 0.68, reinforcing the model's reliability. Additionally, the log-likelihood at convergence (-90.43) suggests that the model has effectively maximized the likelihood function given the data. These results imply that the selected independent variables significantly contribute to explaining accident severity, making the OP model a suitable approach for this analysis.

The findings in Table 3 indicate that accidents occurring between 12:00 and 18:00 and on weekends tend to be less severe. One possible reason is that travel speeds during this period are lower than at other times, such as nighttime and early morning. Increased leisure travel on weekends also contributes to reduced traffic speed. Moreover, lower travel speeds may help mitigate accident severity. Similarly, accidents caused by failure to maintain a safe following distance often result in rear-end collisions, which generally have less severe consequences. However, these assumptions should be further investigated and validated.

By contrast, accidents tend to be more severe with higher posted speed limits, the type of vehicle involved (coach, truck), the type of collision (head-on, rollover, run-off-road), and the number of vehicles involved. When expressway speed limits are higher, operating speeds and speed fluctuations increase, which can lead to more severe accidents. Additionally, some aggressive drivers may exceed the speed limit, further increasing the risk of severe casualties in an accident. Therefore, traffic management agencies and law enforcement should strengthen efforts to enforce speed regulations and penalize violations.

Accidents involving coaches and/or trucks (coach vs. truck, truck vs. truck, or truck vs. passenger car) often result in greater severity. This can be explained by several factors. Trucks and coaches, especially trucks, typically carry heavy loads. At high speeds, their significant kinetic energy makes stopping more difficult. Furthermore, coaches frequently transport many passengers, increasing the risk of fatalities and injuries in crashes.

Head-on collisions, rollovers, and run-off-road accidents are highly dangerous and can result in severe consequences. Head-on collisions frequently occur on undivided expressway sections in the first phase of investment, where there is a single lane per direction and no median barriers. Thus, investing in a median barrier system is crucial for preventing such accidents. Additionally, vehicles running off the road and falling from embankments increase injury severity. When designing expressways, sections with high embankments should be reviewed to ensure proper guardrails prevent vehicles from falling.

Lastly, multi-vehicle collisions involving more than two vehicles are also associated with a higher severity level. In Vietnam, although the government has issued Circular No. 38/2024/TT-BGTVT, dated November 15, 2024, guiding drivers to maintain a safe distance when following other vehicles on expressways, many aggressive drivers do not comply with this rule. As a result, multi-vehicle collisions often occur when the leading vehicle suddenly brakes or stops abruptly. Therefore, launching a campaign to raise awareness among highway drivers about maintaining a safe distance is crucial to minimizing injuries and fatalities in accidents. Enforcement cameras can also be installed in locations where RTAs frequently occur to monitor whether drivers maintain proper safe following distances. These cameras can also detect and penalize speeding violations or unsafe stops on the expressway, especially on sections without emergency stopping lanes.

Table 3. Model estimation results

Independent variables	Coeff.	Std. err	z	P>|z|
Time of accident: 12:00 - 18:00 = 1, otherwise = 0	-0.77	0.31	-2.45	0.014
Date of accident: Weekend = 1, Weekday = 0	-0.57	0.30	-1.90	0.057
Failing to maintain a safe distance: 1 if true, 0 otherwise	-1.15	0.26	-4.39	0.000
Posted speed limit	0.02	0.01	2.35	0.019
Type of vehicle involved: Coach = 1, otherwise = 0	1.44	0.29	4.98	0.000
Type of vehicle involved: Truck = 1, otherwise = 0	1.57	0.46	3.37	0.001
Head-on collision: 1 if true, 0 otherwise	0.93	0.49	1.91	0.056
Rollover/Run-off road: 1 if true, 0 otherwise	1.56	0.56	2.77	0.006
Number of involved vehicles	0.26	0.14	1.91	0.056
/cut1 (μ_1)	3.16	1.24		
/cut1 (μ_2)	4.29	1.23		
Number of estimated parameters	11			
Number of observations	108			
Initial log likelihood	-116.95			
Log likelihood at convergence	-90.43			
Rho squared	0.77			
Adjusted rho squared	0.68			
AIC	202.86			
BIC	232.36			

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This study analyzed 108 road traffic accidents (RTAs) on Vietnam's expressway network, using data collected from online news reports and cross-verified with official sources. Key accident characteristics—including time of occurrence, vehicle involvement, and contributing factors—were examined. The results reveal that more than half of the accidents occurred between 0:00 - 12:00, with the highest proportion between 12:00 - 18:00. Although most accidents happened on weekdays, the accident rate per day was higher on weekends.

Failure to maintain a safe following distance was identified as the leading cause of accidents, followed by head-on collisions, accidents involving stationary vehicles, and run-off-road incidents. Notably, head-on collisions occurred exclusively on two-lane expressways without median barriers, emphasizing the need for safety improvements. Similarly, crashes involving stationary vehicles were concentrated on expressways lacking emergency stopping lanes.

Accident severity analysis, based on the Severity Index (SI) and Ordered Probit (OP) model, highlighted key influencing factors. While accidents occurring between 12:00 - 18:00 and on weekends tended to be less severe due to lower travel speeds, severity increased with higher posted speed limits, vehicle types (coaches and trucks), and specific collision types (head-on, rollover, and run-off-road). Multi-vehicle collisions also posed a greater risk of severe outcomes, particularly when involving large commercial vehicles.

5.2 Limitations and Future Research Directions

The dataset used in this study is limited in size and time span, covering only two years of accident data. Future research should incorporate a more extended dataset spanning 5 to 10 years to improve the robustness of the findings. Additionally, while this study employed the OP model for severity analysis, future research could explore advanced analytical techniques, such as machine learning, Latent Class Ordered Probit (LCOP), Random Parameters Binary Ordered Probit (RPBOP), and Random Parameters Ordered Probit (RPOP) models. These approaches may provide deeper insights into accident severity patterns and improve predictive accuracy.

5.3 Recommendations

These findings underline the need for enhanced safety measures, including the installation of median barriers on undivided expressways, improvements to roadside infrastructure to prevent run-off-road incidents, and stricter enforcement of regulations on maintaining safe following distances and preventing speeding violations. Enforcement cameras should be installed in locations with a high incidence of RTAs to monitor and penalize violations related to safe following distances, speeding, or unsafe stops on expressways, especially on sections without emergency stopping lanes. Additionally, public awareness campaigns should be launched to educate drivers on safe driving practices, particularly on high-speed expressways. Strengthening these measures will help reduce accident severity and improve overall traffic safety on Vietnam's expressway network.

ACKNOWLEDGEMENTS

This research is funded by the Ministry of Education and Training (MoET) under Grant No. B2025-GHA-01

REFERENCES

World Health Organization. (2023). Global status report on road safety 2023. (Accessed February 17, 2025).

National Traffic Safety Committee (2023). Vietnam status report on road safety 2023.

(Accessed February 17, 2025) (in Vietnamese).
Ministry of Public Security (2024). Road Safety Report on Expressways – First Seven Months of 2024 (Accessed February 20, 2025) (in Vietnamese).
Pei, Y., Wen, Y., & Pan, S. (2024). Traffic accident severity prediction based on interpretable deep learning model. *Transportation Letters*, 1-15. https://doi.org/10.1080/19427867.2024.2398336
Dong, L., Gong, R., Wang, Z., Chen, Z., Li, Y., & Ren, W. (2024, August). An accident severity prediction framework with consideration of features interaction. In Proceedings of the Institution of Civil Engineers-Transport (pp. 1-12). Emerald Publishing Limited. https://doi.org/10.1680/jtran.24.00050
Wang, G., Li, J., Shen, L., Ding, S., Shi, Z., & Zuo, F. (2024). Towards efficient and accurate prediction of freeway accident severity using two-level fuzzy comprehensive evaluation. *Heliyon*, 10(16). https://doi.org/10.1016/j.heliyon.2024.e36396
Alhaek, F., Liang, W., Rajeh, T. M., Javed, M. H., & Li, T. (2024). Learning spatial patterns and temporal dependencies for traffic accident severity prediction: A deep learning approach. *Knowledge-Based Systems*, 286, 111406. https://doi.org/10.1016/j.knosys.2024.111406
Cicek, E., Akin, M., Uysal, F., & Topcu Aytas, R. (2023). Comparison of traffic accident injury severity prediction models with explainable machine learning. *Transportation letters*, 15(9), 1043-1054. https://doi.org/10.1080/19427867.2023.2214758
Pérez-Sala, L., Curado, M., Tortosa, L., & Vicent, J. F. (2023). Deep learning model of convolutional neural networks powered by a genetic algorithm for prevention of traffic accidents severity. *Chaos, Solitons & Fractals*, 169, 113245. https://doi.org/10.1016/j.chaos.2023.113245
Li, B., Lu, J., Ji, Y., & Wang, S. (2024). Exploring the critical factors influencing the severity of maritime accidents via multinomial logit model with adaptive sparse group lasso penalty. *Ocean Engineering*, 313, 119355. https://doi.org/10.1016/j.oceaneng.2024.119355
Sufian, M. A., Varadarajan, J., & Niu, M. (2024). Enhancing prediction and analysis of UK road traffic accident severity using AI: integration of machine learning, econometric techniques, and time series forecasting in public health research. *Heliyon*, 10(7). https://doi.org/10.1016/j.heliyon.2024.e28547
Wang, H., Liu, Z., Wang, X., Huang, D., Cao, L., & Wang, J. (2022). Analysis of the injury-severity outcomes of maritime accidents using a zero-inflated ordered probit model. Ocean Engineering, 258, 111796. https://doi.org/10.1016/j.oceaneng.2022.111796
Nickkar, A., Pourfalatoun, S., Miller, E. E., & Lee, Y. J. (2024). Applying the heteroskedastic ordered probit model on injury severity for improved age and gender estimation. Traffic injury prevention, 25(2), 202-209. https://doi.org/10.1080/15389588.2023.2286429
Fang, Z., Yuan, R., & Xiang, Q. (2024). An exploratory investigation into the influence of risk factors on driver injury severity in angle crashes: A random parameter bivariate ordered probit model approach. Traffic injury prevention, 25(1), 70-77. https://doi.org/10.1080/15389588.2023.2244104
Karabulut, N. C., & Ozen, M. (2023). Exploring driver injury severity using latent class ordered probit model: A case study of Turkey. KSCE Journal of Civil Engineering, 27(3), 1312-1322. https://doi.org/10.1007/s12205-023-0473-6
Hyodo, S., & Hasegawa, K. (2021). Factors affecting analysis of the severity of accidents in cold and snowy areas using the ordered probit model. Asian Transport Studies, 7,

100035. https://doi.org/10.1016/j.eastsj.2021.100035
Atombo, C., Turkson, R. F., & Akple, M. S. (2023). Estimating injury severity for motorized and non-motorized vehicle-involved crashes: insights from random-parameter ordered probit model with heterogeneity in means and variances. IATSS research, 47(4), 455-467. https://doi.org/10.1016/j.iatssr.2023.09.003
AASHTO (2010). Highway safety manual. American Association of State Highway and Transportation Officials: Washington, DC, USA, 19192.
K. Geurts, G. Wets, T. Brijs, and K. Vanhoof, Identification and ranking of black spots: Sensitivity analysis, Transportation Research Record, 1897(2004) 34-42. https://doi.org/10.3141/1897-05