



Định lượng rủi ro và mức độ nghiêm trọng tai nạn trên đường bộ cao tốc phân kỳ: Tiếp cận bằng mô hình Logit thứ bậc

Quantifying risk and accident severity on staged-development expressways: An Ordered Logit approach

Lê Xuân Quý^{1,*}, Chu Tiến Dũng¹, Phan Văn Mên¹

¹Trường Đại học Giao thông vận tải

Từ khóa:

Cao tốc phân kỳ;
Tai nạn giao thông;
Hồi quy Logit thứ bậc.

TÓM TẮT

Trong bối cảnh nguồn vốn hạn hẹp, giải pháp đầu tư phân kỳ (2 làn xe hoặc 4 làn xe hạn chế) đã được áp dụng rộng rãi tại Việt Nam. Tuy nhiên, mô hình này đang bộc lộ nhiều thách thức về an toàn do thiếu hụt dải phân cách cứng và làn dừng khẩn cấp liên tục. Nghiên cứu này phân tích đặc điểm và lượng hóa các nhân tố ảnh hưởng đến mức độ nghiêm trọng của tai nạn giao thông dựa trên dữ liệu 170 vụ tai nạn thực tế giai đoạn 2018-2025. Bằng phương pháp thống kê mô tả kết hợp mô hình hồi quy Logit thứ bậc, nghiên cứu xác định xe tải là phương tiện chủ đạo trong hơn 71% các vụ va chạm. Kết quả mô hình chỉ ra rằng các vụ tai nạn vào ngày thường có xác suất nghiêm trọng cao gấp 2,445 lần so với cuối tuần. Đặc biệt, loại hình tai nạn lật xe hoặc lao ra khỏi đường - hệ quả từ hạ tầng thiếu làn dừng khẩn cấp - làm tăng nguy cơ nghiêm trọng lên gấp 4,178 lần. Những bằng chứng khoa học này cung cấp cơ sở để các cơ quan quản lý ưu tiên lộ trình nâng cấp hạ tầng và điều chỉnh phương án tổ chức giao thông trên các tuyến cao tốc phân kỳ.

Keywords:

Staged-development expressway;
Traffic accident;
Ordered Logit regression.

ABSTRACT

With limited public investment, phased investment strategies (2-lane or limited 4-lane configurations) are widely adopted in Vietnam. However, this model poses safety risks due to the lack of median barriers and continuous emergency lanes. This study analyzes and quantifies factors influencing traffic accident severity using 170 accident records from 2018-2025. Integrating descriptive statistics and hierarchical logistic regression, the research identifies heavy trucks as the primary vehicles in over 71% of collisions. The model reveals that accidents on weekdays pose a severity risk 2.445 times higher than on weekends. Notably, "vehicle overturn or run-off-road" accidents - a direct result of inadequate emergency lanes - increase the probability of severe outcomes by 4.178 times. These findings provide a scientific basis for prioritizing infrastructure upgrades and optimizing traffic management on phased expressways.

* Lê Xuân Quý. Trường Đại học Giao thông vận tải
Email: lexuanquy@utc.edu.vn

<https://www.doi.org/10.55228/JTST150404>

Ngày nhận bài: 10/01/2026; Ngày nhận bài sửa: 07/05/2026; Ngày chấp nhận đăng: 21/05/2026

pISSN: 1859-4263.

1. Giới thiệu

Phát triển mạng lưới đường bộ cao tốc là một trụ cột chiến lược nhằm thúc đẩy phát triển kinh tế - xã hội và đảm bảo an ninh quốc phòng tại Việt Nam. Trong những năm gần đây, hàng loạt dự án, đặc biệt là các hợp phần của tuyến cao tốc Bắc - Nam phía Đông, đã được đẩy nhanh tiến độ và đưa vào khai thác, từng bước hình thành một hệ thống giao thông huyết mạch hiện đại. Tuy nhiên, trong bối cảnh nguồn vốn đầu tư công còn hạn hẹp và áp lực phải nhanh chóng hoàn thành mục tiêu, giải pháp "đầu tư phân kỳ" đã được áp dụng rộng rãi [1], [2]. Thực tiễn này đã tạo ra hai nhóm tuyến đường cao tốc có đặc thù vận hành rủi ro cao: các tuyến 2 làn xe và các tuyến 4 làn xe hạn chế (nền đường 17m).

Một đặc điểm kỹ thuật chung và cũng là nguy cơ lớn nhất của các tuyến phân kỳ 2 làn là việc thiếu dải phân cách cứng ở giữa, chỉ sử dụng vạch sơn. Điều này tạo ra rủi ro thường trực của các vụ va chạm đối đầu trực diện ở tốc độ cao - loại tai nạn thường dẫn đến hậu quả thảm khốc nhất. Song song đó, đặc điểm thiết kế của các tuyến phân kỳ 4 làn xe là thiếu làn dừng xe khẩn cấp liên tục (chỉ thiết kế các dải an toàn, bố trí gián đoạn cả hai bên theo chiều xe chạy, cách nhau từ 4,0 – 5,0 km). Các dải an toàn (hay dải dừng khẩn cấp) được thiết kế với chiều dài tối thiểu 170 m (bao gồm cả chiều dài đoạn tăng tốc và đoạn giảm tốc), có bề rộng tối thiểu 2,5m (hoặc 2,0 m trong điều kiện khó khăn hoặc đặc biệt). Tuy nhiên, việc bố trí các dải dừng khẩn cấp mà không bố trí làn dừng khẩn cấp liên tục có thể khiến các phương tiện gặp sự cố không có nơi trú ẩn an toàn, buộc phải dừng ngay trên làn xe chạy. Các phương tiện này lập tức trở thành những "chướng ngại vật tĩnh", gợi ý rằng các vụ va chạm đâm từ phía sau với động năng lớn, dẫn đến tai nạn giao thông đặc biệt nghiêm trọng.

Tại Việt Nam, đã có nhiều nghiên cứu về an toàn giao thông trên các tuyến đường bộ cao tốc. Các nghiên cứu tập trung về vận tốc khai thác của các loại phương tiện khác nhau trên một số tuyến đường bộ cao tốc [3], [4] chỉ ra rằng các loại xe khác nhau (xe tải, xe khách, xe con...) có tốc độ khai thác khác nhau do đó cần thiết phải có các giải pháp để tách các dòng xe có tốc độ khác nhau nhằm giảm xung đột trong quá trình khai thác, ví dụ như sử dụng biển báo hoặc phân làn cho các phương tiện chạy chậm đi về bên phải. Nghiên cứu [5] tiếp cận từ phương pháp thiết kế kết

cấu mặt đường cao tốc thích hợp để phù hợp với tải trọng trục xe khai thác thực tế nhằm xây dựng được kết cấu bền vững và đảm bảo an toàn cho phương tiện khai thác. Nghiên cứu của Chu Tiến Dũng và cộng sự [6] dựa trên số liệu thống kê 108 vụ tai nạn trên các tuyến đường bộ cao tốc ở Việt Nam từ tháng 1 năm 2023 đến tháng 2 năm 2025 đã chỉ ra rằng các điều kiện về hạ tầng, loại xe, và hình thức xảy ra tai nạn có ảnh hưởng đáng kể đến mức độ nghiêm trọng của tai nạn giao thông trên đường cao tốc, trong đó, yếu tố dải phân cách và làn dừng khẩn cấp giúp làm giảm đáng kể chấn thương và nguy cơ xảy ra chết người khi xảy ra tai nạn. Tuy nhiên, phần lớn các công trình phân tích chuyên sâu, định lượng về đặc thù tai nạn trên chính các tuyến cao tốc phân kỳ còn rất hạn chế. Các cơ quan quản lý và đơn vị vận hành cần một bằng chứng khoa học rõ ràng để trả lời câu hỏi: Mức độ ảnh hưởng của các yếu tố trên các tuyến đường bộ cao tốc phân kỳ đầu tư đến xác suất xảy ra tai nạn nghiêm trọng là bao nhiêu? Bài báo này giải quyết khoảng trống nghiên cứu đó bằng cách xây dựng mô hình thống kê dự báo mức độ nghiêm trọng của tai nạn giao thông dựa trên bộ dữ liệu tai nạn thực tế. Mục tiêu chính là: (1) phân tích các yếu tố ảnh hưởng đến tai nạn và (2) lượng hóa một cách khoa học tác động của việc thiếu các hạng mục an toàn then chốt. Kết quả của nghiên cứu sẽ là cơ sở quan trọng để đề xuất các giải pháp kỹ thuật, giải pháp tổ chức giao thông và hoạch định chiến lược nâng cấp, mở rộng các tuyến cao tốc này trong tương lai.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Dữ liệu

Cơ sở dữ liệu được sử dụng trong nghiên cứu này là tập hợp hồ sơ các vụ tai nạn giao thông (TNGT) xảy ra trên các tuyến cao tốc phân kỳ (cả loại hình 2 làn xe và 4 làn xe hạn chế) tại Việt Nam. Dữ liệu được thu thập và tổng hợp thủ công từ các nguồn công bố chính thức, bao gồm các cổng thông tin điện tử của cơ quan quản lý, các báo cáo nhanh vụ việc và thông tin đăng tải trên các cơ quan truyền thông, báo chí uy tín trong khoảng thời gian từ tháng 01/2018 đến hết tháng 12/2025.

Cần lưu ý rằng do nguồn dữ liệu chủ yếu từ các báo cáo và truyền thông, mẫu nghiên cứu có sự tập trung cao vào các vụ tai nạn có thương vong. Điều này tạo ra một sự sai lệch lựa chọn (Selection Bias) nhất định

khi so sánh với phân phối thực tế của toàn bộ các vụ va chạm (bao gồm cả va chạm nhẹ). Tuy nhiên, cách tiếp cận này cho phép nghiên cứu tập trung sâu vào việc phân tích các yếu tố dẫn đến hậu quả khốc liệt nhất trên các tuyến phân kỳ – vốn là đối tượng ưu tiên trong công tác quản lý an toàn giao thông.

Sau quá trình làm sạch, chuẩn hóa và loại bỏ các bản ghi trùng lặp hoặc thiếu thông tin cốt lõi, bộ dữ liệu cuối cùng được sử dụng để phân tích bao gồm 170 vụ TNGT. Mỗi hồ sơ vụ tai nạn (tương ứng với một hàng trong cơ sở dữ liệu) được mô tả bằng 29 biến số khác nhau. Các biến số này được mã hóa và phân thành các nhóm chính để phục vụ cho việc xây dựng mô hình, trong đó quan trọng nhất là:

- **Biến phụ thuộc (Y):** Mục tiêu của mô hình là dự báo Mức độ nghiêm trọng của tai nạn. Biến này được xây dựng dưới dạng biến thứ bậc dựa trên dữ liệu thu thập được. Mức độ nghiêm trọng của tai nạn được xây dựng từ công thức đề xuất của Geurts và cộng sự [7] (Công thức (1)), theo đó, mức độ nghiêm trọng của tai nạn (SI) được đánh trọng số 1, 3, và 5 cho số lượng người bị thương nhẹ ($N_{thương\ nhẹ}$), thương nặng ($N_{thương\ nặng}$) hoặc tử vong ($N_{tử\ vong}$). Giá trị SI này sau đó được mã hóa thành 4 cấp độ tăng dần theo mức độ nghiêm trọng của hậu quả: Cấp 1 ($Y=0, SI=0$): Tai nạn không nghiêm trọng - Tai nạn chỉ gây thiệt hại về phương tiện hoặc hạ tầng, không có bất kỳ thương vong nào về người (Số người chết = 0, Số người bị thương = 0); Cấp 2 ($Y=1, SI \leq 5$): Tai nạn nhẹ; Cấp 3 ($Y=2, 5 < SI \leq 10$): Tai nạn nặng; Cấp 4 ($Y=3, SI > 10$) Tai nạn rất nặng.

$$SI = 1 \times N_{thương\ nhẹ} + 3 \times N_{thương\ nặng} + 5 \times N_{tử\ vong} \quad (1)$$

- **Biến độc lập (X):** Đây là các yếu tố dự báo, được phân thành các nhóm: (1) Đặc điểm hình học của đường; (2) Đặc điểm thời gian và môi trường; (3) Đặc điểm va chạm; (4) Loại phương tiện liên quan.

Cấu trúc chi tiết của các biến được trình bày trong Bảng 1. Bộ dữ liệu này là cơ sở đầu vào để đưa vào mô hình thống kê trong phần phương pháp nghiên cứu tiếp theo.

Bảng 1. Nhóm biến mô tả cấu trúc cơ sở dữ liệu tai nạn giao thông trên cao tốc phân kỳ.

Nhóm	Mã biến	Tên biến	Mô tả/Mã hóa
Đặc điểm hình học của đường	X1	Số làn xe	Biến phân loại (2 và 4 làn xe)
	X2	Có dải phân cách cứng	Biến nhị phân (1=Có, 0=Không)
	X3	Có làn dừng khẩn cấp	Biến nhị phân (1=Có, 0=Không)
	X4	Tốc độ cho phép (Km/h)	Biến phân loại (80, 90, 100Km/h)
Đặc điểm thời gian và môi trường	X5	Tai nạn khung giờ 0h-6h	Biến nhị phân (1=Có, 0=Không)
	X6	Tai nạn khung giờ 6h-12h	Biến nhị phân (1=Có, 0=Không)
	X7	Tai nạn khung giờ 12h-18h	Biến nhị phân (1=Có, 0=Không)
	X8	Tai nạn khung giờ 18h-24h	Biến nhị phân (1=Có, 0=Không)
	X9	Xảy ra vào ngày thường (T2-T6)	Biến nhị phân (1=Có, 0=Không)
	X10	Xảy ra vào ngày cuối tuần (T7, CN)	Biến nhị phân (1=Có, 0=Không)
	X11	Điều kiện thời tiết (Mưa)	Biến nhị phân (1=Có, 0=Không)
Đặc điểm va chạm	X12	Không giữ khoảng cách an toàn	Biến nhị phân (1=Có, 0=Không)
	X13	Đổi đầu	Biến nhị phân (1=Có, 0=Không)
	X14	Đâm vào xe đang dừng đỗ	Biến nhị phân (1=Có, 0=Không)
	X15	Lật xe/lao ra khỏi đường	Biến nhị phân (1=Có, 0=Không)
Loại phương tiện	X16	Xe tải và xe khách	Biến nhị phân (1=Có, 0=Không)

Nhóm	Mã biến	Tên biến	Mô tả/Mã hóa
liên quan	X17	Xe tải và xe tải	Biến nhị phân (1=Có, 0=Không)
	X18	Xe khách và xe khách	Biến nhị phân (1=Có, 0=Không)
	X19	Xe khách và xe con	Biến nhị phân (1=Có, 0=Không)
	X20	Xe tải và xe con	Biến nhị phân (1=Có, 0=Không)
	X21	Tai nạn liên hoàn (từ 3 xe trở lên)	Biến nhị phân (1=Có, 0=Không)
	X22	Xe con và xe con	Biến nhị phân (1=Có, 0=Không)
	X23	Va chạm với xe máy	Biến nhị phân (1=Có, 0=Không)
	X24	Va chạm với người đi bộ	Biến nhị phân (1=Có, 0=Không)
	X25	Tự gây tai nạn	Biến nhị phân (1=Có, 0=Không)
Mức độ nghiêm trọng của tai nạn	Y	Va chạm chỉ hư hỏng tài sản	Biến phân loại (Y=0)
		Tai nạn nhẹ	Biến phân loại (Y=1)
		Tai nạn nặng	Biến phân loại (Y=2)
		Tai nạn rất nặng	Biến phân loại (Y=3)

2.2. Phương pháp nghiên cứu

Quy trình nghiên cứu được thực hiện thông qua hai bước chính: (1) Phân tích thống kê mô tả và (2) Phân tích mô hình hồi quy.

2.2.1. Phân tích thống kê mô tả

Ở bước đầu tiên, phân tích thống kê mô tả [8], [9], [10], [11] được thực hiện trên toàn bộ cơ sở dữ liệu. Mục tiêu của bước này là cung cấp một bức tranh tổng thể về đặc điểm của các vụ tai nạn trên cao tốc phân kỳ.

Nghiên cứu thực hiện phân tích tần suất và tính toán tỷ lệ phần trăm cho tất cả các biến quan tâm để giúp làm rõ các câu hỏi cơ bản như: Tai nạn thường xảy ra vào thời điểm nào (khung giờ, ngày trong tuần)? Loại hình va chạm nào là phổ biến nhất (ví dụ: đối đầu, lật xe, đâm vào xe đang dừng)? Tỷ lệ các vụ tai nạn xảy ra trên các đoạn có và không có làn dừng khẩn cấp/dải phân cách cứng là bao nhiêu?... Đây là cơ sở quan trọng để xác định các xu hướng ban đầu trước khi đi vào mô hình hóa.

2.2.2. Mô hình hồi quy Logit thứ bậc (Ordered Logit regression)

Nghiên cứu sử dụng Mô hình hồi quy Logit thứ bậc [12], [13], [14], [15], [16] để phân tích các yếu tố ảnh hưởng. Mô hình này được lựa chọn vì biến phụ thuộc (Y) - Mức độ nghiêm trọng của tai nạn là một biến phân loại có tính thứ bậc tự nhiên.

Việc sử dụng mô hình Logit thứ bậc cho phép tận dụng được thông tin về thứ bậc của dữ liệu (mức độ nghiêm trọng tăng dần từ Y=0 đến Y=3). Điều này giúp mô hình vượt trội hơn so với Hồi quy Logistic nhị phân [9] (chỉ phân biệt được Tai nạn/Không tai nạn) hoặc Hồi quy Logistic đa thức [10] (không phân biệt được thứ tự nghiêm trọng giữa các nhóm), từ đó cung cấp các ước lượng chính xác hơn về tác động của từng biến số độc lập lên xác suất xảy ra các mức độ tai nạn nghiêm trọng hơn.

Mô hình sẽ ước tính xác suất một vụ tai nạn rơi vào một mức độ nghiêm trọng nhất định (hoặc cao hơn), dựa trên một tập hợp các biến độc lập (X). Phương trình tổng quát của mô hình có dạng:

$$\ln \left[\frac{P(Y \leq j)}{1 - P(Y \leq j)} \right] = \alpha_j - \sum_{i=1}^k \beta_i X_i \quad (2)$$

Trong đó: $P(Y \leq j)$ là xác suất tích lũy của vụ tai nạn có mức độ nghiêm trọng là j hoặc thấp hơn; α_j là các điểm cắt (cut-off points) giữa các cấp độ nghiêm trọng; X_i là các biến độc lập (như thời tiết, loại va chạm, loại xe...); β_i là các hệ số hồi quy, cho biết tác động của từng biến X_i .

Kết quả (dưới dạng Tỷ lệ Chênh - Odds Ratios) sẽ cho phép chúng ta kết luận một cách khoa học rằng việc thiếu hụt các yếu tố an toàn này làm tăng (hay giảm) khả năng xảy ra một vụ tai nạn thảm khốc lên bao nhiêu lần, sau khi đã kiểm soát ảnh hưởng của các yếu tố khác.

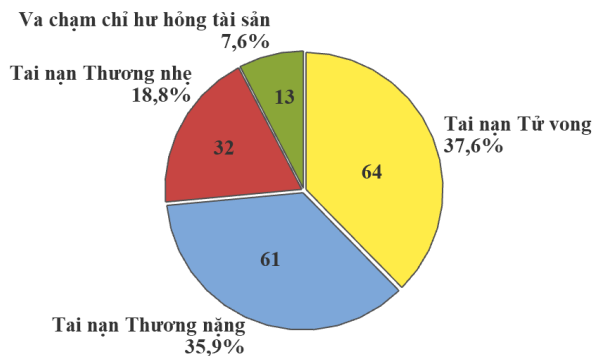
3. Kết quả nghiên cứu

3.1. Phân tích thống kê mô tả

Trong phần phân tích này, số liệu được trình bày dưới dạng tần suất xuất hiện và tỷ trọng phần trăm của các vụ tai nạn. Nhóm tác giả nhận thấy rằng do các tuyến cao tốc phân kỳ được đưa vào khai thác tại các thời điểm khác nhau trong giai đoạn nghiên cứu, việc chuẩn hóa số vụ theo chiều dài đường và lưu lượng xe gặp nhiều hạn chế về tính đồng bộ. Do đó, kết quả thống kê mô tả dưới đây tập trung vào việc nhận diện các 'điểm nóng' và đặc điểm phân bố của tai nạn trên các loại hình hạ tầng cụ thể, làm tiền đề cho các phân tích định lượng chuyên sâu ở các nội dung tiếp theo.

Bộ dữ liệu nghiên cứu bao gồm 170 vụ TNGT được thu thập trên các tuyến cao tốc phân kỳ tại Việt Nam (trình bày ở Mục 2.1). Hình 1 cung cấp một cái nhìn tổng quan về hậu quả của 170 vụ tai nạn được phân tích. Các vụ tai nạn đã được phân loại theo mức độ nghiêm trọng cao nhất được ghi nhận. Kết quả cho thấy một bức tranh khá rõ ràng về sự phân bố: Nhóm "Tai nạn Tử vong" chiếm tỷ lệ 37,6% (tương ứng 64 vụ). Nhóm "Tai nạn Thương nặng" chiếm tỷ lệ 35,9% (tương ứng 61 vụ). Khi cộng gộp lại, hai nhóm có hậu quả nghiêm trọng nhất này chiếm 73,5%, tương đương gần ba phần tư tổng số vụ tai nạn trong bộ dữ liệu. Các nhóm còn lại, bao gồm "Tai nạn Thương nhẹ" (18,8% ~ 32 vụ) và "Va chạm chỉ hư hỏng tài sản" (7,6% ~ 13 vụ), chiếm tỷ lệ nhỏ hơn đáng kể.

Việc các vụ tai nạn có hậu quả nghiêm trọng chiếm tỷ lệ lớn trong dữ liệu cho thấy tầm quan trọng của việc tìm hiểu các yếu tố góp phần. Điều này tạo cơ sở để chúng ta xem xét kỹ hơn vai trò của các yếu tố ảnh hưởng trong các phần phân tích tiếp theo.



Hình 1. Phân bố các vụ TNGT theo mức độ nghiêm trọng (N=170) trên các tuyến cao tốc phân kỳ ở Việt Nam từ 01/2018 đến 12/2025.

Bảng 2 cung cấp thông tin chi tiết về bối cảnh thời gian và môi trường khi 170 vụ tai nạn xảy ra cùng với các thông tin về đặc điểm va chạm và loại phương tiện xuất hiện trong các vụ tai nạn.

Bảng 2. Thống kê mô tả các đặc điểm chính của tai nạn giao thông trên cao tốc phân kỳ tại Việt Nam (N=170).

Hạng mục	Biến số	Số vụ	Tỷ lệ (%)
Đặc điểm đường			
(X1) Số làn xe	2 làn	90	52,94
	4 làn	80	47,06
(X2) Dải phân cách cứng	Không (0)	96	56,47
	Có (1)	74	43,53
(X3) Làn dừng khẩn cấp	Không (0)	152	89,41
	Có (1)	18	10,59
(X4) Tốc độ cho phép	80 km/h	104	61,18
	90 km/h	66	38,82
Thời gian & Môi trường			
(X5-X8) Khung giờ	0h-6h	57	33,53
	6h-12h	39	22,94
	12h-18h	54	31,76
	18h-24h	20	11,76
(X9/X10) Ngày trong tuần	Ngày thường	114	67,06
	Cuối tuần	56	32,94
(X11) Điều kiện thời tiết	Tốt (Không mưa)	141	81,98
	Mưa (1)	31	18,02
Đặc điểm va chạm			
(X12) Không giữ khoảng cách an toàn	Có (1)	73	41,95
(X13) Đối đầu	Có (1)	42	24,14
(X14) Đâm vào xe dừng đỗ	Có (1)	29	16,67
(X15) Lật xe/lao ra khỏi đường	Có (1)	30	17,24
Loại phương tiện			
(X16) Xe tải và Xe khách	Có (1)	34	19,65
(X17) Xe tải và Xe tải	Có (1)	58	33,53
(X18) Xe khách và Xe khách	Có (1)	7	4,05
(X19) Xe khách và Xe con	Có (1)	6	3,47

Hạng mục	Biến số	Số vụ	Tỷ lệ (%)
(X20) Xe tải và Xe con	Có (1)	20	11,56
(X21) Tai nạn liên hoàn (>= 3 xe)	Có (1)	12	6,94
(X22) Xe con và Xe con	Có (1)	4	2,31
(X23) Va chạm với Xe máy	Có (1)	0	0,00
(X24) Va chạm với Người đi bộ	Có (1)	3	1,73
(X25) Tự gây tai nạn	Có (1)	29	16,76

Phân tích về Đặc điểm hạ tầng của đường cao tốc (ứng với các biến X1, X2, X3, X4):

- Số liệu tai nạn ở các đoạn cao tốc phân kỳ 2 làn xe và 4 làn xe hạn chế (chỉ có dải an toàn) (X1) cho thấy tai nạn xảy ra trên các tuyến 2 làn (52,94%) nhiều hơn một chút so với các tuyến 4 làn hạn chế (47,06%), tuy nhiên không có quá nhiều sự khác biệt;

- Việc thiếu hụt dải phân cách cứng (X2) làm tăng nguy cơ tai nạn khoảng 29,7% (96 vụ trên các đoạn không bố trí dải phân cách cứng so với 74 vụ trên các đoạn có bố trí). Việc thiếu sự ngăn cách vật lý giữa hai chiều xe chạy có mối liên hệ thống kê với các vụ va chạm đối đầu hoặc va chạm bên khi xe lấn làn;

- Việc bố trí làn dừng khẩn cấp (X3): Trong mẫu khảo sát, tỷ trọng tai nạn xảy ra trên các đoạn không có làn dừng khẩn cấp chiếm ưu thế tuyệt đối (152 vụ) so với các đoạn có bố trí (18 vụ). Mặc dù số liệu này chưa được chuẩn hóa theo chiều dài tuyến, nhưng sự chênh lệch lớn về số vụ tuyệt đối cho thấy các điểm không có nơi trú ẩn an toàn là những vị trí ghi nhận nhiều va chạm dừng đỗ nhất;

- Dữ liệu về tốc độ cho phép (X4) cho thấy số vụ tai nạn tại các đoạn giới hạn tốc độ 80 km/h (104 vụ - 61,18%) cao hơn đáng kể so với các đoạn 90 km/h (66 vụ - 38,82%). Điều này cho thấy tốc độ cao hơn (90 km/h) không hẳn là nguyên nhân gây ra nhiều tai nạn hơn. Ngược lại, các đoạn 80 km/h thường gắn liền với hạ tầng 2 làn xe, không dải phân cách. Sự kết hợp giữa hạ tầng hạn chế và tốc độ thấp thường gây

ra sự ức chế tâm lý và các hành vi vượt ẩu rủi ro, trong khi các đoạn 90 km/h (thường là 4 làn, có dải phân cách) lại giúp dòng giao thông vận hành ổn định hơn.

Phân tích về Đặc điểm thời gian và môi trường (X5-X11):

- Số liệu phân tích cho thấy rằng yếu tố thời tiết (X11) không phải là nguyên nhân chính. Đại đa số các vụ tai nạn (81,98%) xảy ra trong điều kiện thời tiết Tốt (Không mưa). Kết quả cho thấy rằng tai nạn không phải do "trời mưa đường trơn" là nguyên nhân chính gây tai nạn;

- Về khung giờ xảy ra tai nạn (X5-X8): Dữ liệu cho thấy hai khung giờ có tỷ lệ tai nạn cao nhất là 0h-6h (33,53%) và 12h-18h (31,76%), cho thấy mối liên hệ chặt chẽ với trạng thái sinh lý của người điều khiển phương tiện. Trong đó, khung giờ từ 0h-6h sáng là thời điểm nguy hiểm nhất do nhịp sinh học tự nhiên khiến cơ thể rơi vào trạng thái buồn ngủ sâu, dễ gây ra hiện tượng "ngủ gật trắng" khiến lái xe mất kiểm soát hoàn toàn. Đối với khung giờ 12h-18h, tai nạn thường xuất phát từ sự mệt mỏi tích tụ sau nhiều giờ làm việc và trạng thái giảm tập trung sau bữa trưa (post-lunch dip). Việc chiếm tổng cộng hơn 65% số vụ tai nạn tại hai thời điểm này khẳng định rằng buồn ngủ và mệt mỏi là những nguyên nhân hàng đầu dẫn đến rủi ro trên đường cao tốc, đòi hỏi lái xe phải có chế độ nghỉ ngơi nghiêm ngặt để đảm bảo an toàn;

- Về yếu tố thời gian theo ngày (X9, X10), số liệu ghi nhận 114 vụ xảy ra vào ngày thường (chiếm 67,06%) và 56 vụ vào cuối tuần (chiếm 32,94%). Dù con số tuyệt đối vào ngày thường cao gấp đôi, nhưng nếu phân tích dựa trên hiệu suất bình quân theo ngày, có thể thấy trung bình mỗi ngày thường xảy ra khoảng 22,8 vụ, trong khi mỗi ngày cuối tuần (Thứ 7 và Chủ nhật) con số này lên tới 28 vụ. Điều này cho thấy rủi ro tai nạn vào cuối tuần thực tế cao hơn khoảng 23% so với ngày thường. Nguyên nhân có thể do lưu lượng xe cá nhân và xe du lịch tăng đột biến, kết hợp với tâm lý chủ quan hoặc lái xe trong trạng thái hưng phấn, mệt mỏi sau các chuyến đi chơi xa, dẫn đến mật độ va chạm cao hơn dù tổng thời gian ngắn hơn.

Phân tích về đặc điểm va chạm và loại phương tiện (X12-X15): Phân tích này cho thấy rõ mối liên hệ giữa Thiết kế hạ tầng và Loại hình tai nạn. Về hành vi: Loại hình phổ biến nhất là "Không giữ khoảng cách an

toàn" (X12 - chiếm 41,95%). Đây là lỗi hành vi trực tiếp, nhưng nó bị khuếch đại bởi một hạ tầng không cho phép sai sót. Số liệu phân tích cũng cho thấy 89,41% số vụ xảy ra ở nơi thiếu làn dừng khẩn cấp, điều này dẫn đến 16,67% số vụ tai nạn là "Đâm vào xe đang dừng đỗ" (X14). Đây là bằng chứng rõ ràng cho thấy sự thiếu hụt thiết kế (X3) đã trực tiếp dẫn đến một loại hình tai nạn cụ thể. Tương tự, 43,53% số vụ xảy ra ở nơi thiếu dải phân cách cứng (X2). Điều này tạo điều kiện cho 24,14% số vụ tai nạn là "Va chạm đối đầu" (X13), loại hình va chạm có mức độ nghiêm trọng cao nhất.

Về loại phương tiện (X17-X25): Dựa trên số liệu, xe tải là đối tượng trọng tâm trong các vụ tai nạn cao tốc khi góp mặt trong trên 71% tổng số vụ va chạm. Đáng chú ý nhất là các vụ va chạm giữa xe tải với xe tải chiếm tỷ lệ lớn nhất với 33,53%, tiếp theo là va chạm với xe khách (19,65%) và xe con (11,56%). Sự hiện diện của xe tải không chỉ phản ánh áp lực vận tải hàng hóa mà còn giải thích cho tỷ lệ tai nạn cao vào khung giờ 0h-6h, khi các lái xe đường dài dễ rơi vào trạng thái buồn ngủ và mệt mỏi. Với khối lượng lớn và điểm mù rộng, các vụ tai nạn liên quan đến xe tải thường gây ra hậu quả nghiêm trọng, biến đây trở thành nhóm phương tiện cần được kiểm soát chặt chẽ nhất về tốc độ và thời gian lái xe để đảm bảo an toàn.

3.2. Kết quả Mô hình hồi quy Logit thứ bậc

Để lượng hóa tác động của các yếu tố hạ tầng và các yếu tố khác lên mức độ nghiêm trọng của TNGT, mô hình hồi quy Logit thứ bậc đã được thực hiện. Biến phụ thuộc (Y) là Mức độ nghiêm trọng được phân loại theo 4 cấp độ thứ bậc (từ Va chạm chỉ hư hỏng tài sản đến Tai nạn rất nặng theo giá trị SI – Mức độ nghiêm trọng của tai nạn). Các biến độc lập được đưa vào mô hình từ Bảng 1.

Trong mô hình hồi quy, để đánh giá tác động của các yếu tố định tính như khung giờ, ngày trong tuần hay loại phương tiện, các biến này được chuyển đổi sang dạng biến giả (dummy variables). Nhằm tránh hiện tượng đa cộng tuyến hoàn hảo hay còn gọi là 'bẫy biến giả' (dummy variable trap) Nghiên cứu thực hiện quy trình phân tích lặp và lựa chọn một nhóm đặc tính làm biến tham chiếu (reference category) để loại bỏ khỏi phương trình hồi quy. Cụ thể, số lượng biến giả được đưa vào mô hình luôn là

$n-1$ (với n là số trạng thái của biến). Việc lựa chọn biến tham chiếu dựa trên nguyên tắc nhóm có tần suất xuất hiện cao nhất hoặc nhóm có tính chất làm nền tảng so sánh (ví dụ: điều kiện thời tiết 'Bình thường' hoặc khung giờ '18h-24h' có tỷ lệ tai nạn thấp nhất). Kết quả các hệ số hồi quy của các biến giả còn lại sẽ được diễn giải là sự thay đổi trong biến phụ thuộc so với nhóm tham chiếu này, đảm bảo ma trận hiệp phương sai không bị suy biến và các ước lượng thu được là không chệch. Nghiên cứu đã thực hiện loại bỏ các biến tham chiếu (reference variables) bao gồm X8 (Khung giờ 18h-24h) và X10 (Ngày cuối tuần). Việc lựa chọn X8 làm biến nền là do đây là khung giờ có tỷ trọng tai nạn thấp nhất (11,76%), tạo cơ sở so sánh tương quan để làm nổi bật sự gia tăng rủi ro của các khung giờ khác như 0h-6h hay 12h-18h.

Bên cạnh đó, mô hình thực hiện quy trình loại bỏ lặp đối với các biến có số lượng quan sát thấp. Cụ thể, các biến như va chạm giữa xe khách và xe con (X18), xe khách và xe con (X19), xe con và xe con (X22), va chạm với xe máy (X23) hay va chạm với người đi bộ (X24) (lần lượt chiếm 4,05%, 3,47%, 2,31%, 0% và 1,73% mẫu) nên không đủ bằng chứng thống kê để khẳng định tác động có ý nghĩa đến mô hình tổng thể. Việc loại bỏ các biến này giúp giảm nhiễu, tối ưu hóa bậc tự do và tăng cường độ tin cậy cho các hệ số hồi quy của nhóm phương tiện chủ chốt là xe tải (chiếm trên 71% tổng số vụ). Kết quả cuối cùng là một mô hình tinh gọn, tập trung vào các biến số thực sự có sức mạnh giải thích, đảm bảo tính hội tụ của thuật toán ước lượng và tránh được hiện tượng quá khớp (overfitting).

Để đảm bảo tính hợp lệ của mô hình Logit thứ bậc, nghiên cứu đã tiến hành kiểm định giả định tỷ lệ chênh đồng đều (Proportional Odds Assumption – POA) bằng kiểm định Brant. Kết quả cho thấy không có đủ bằng chứng để bác bỏ giả định này (P-value của các kiểm định đều lớn hơn 0,05), do đó việc sử dụng mô hình Ordered Logit là phù hợp với cấu trúc dữ liệu. Bên cạnh đó, hiện tượng đa cộng tuyến giữa các biến độc lập được kiểm tra thông qua hệ số phóng đại phương sai (Variance Inflation Factor – VIF). Kết quả kiểm tra VIF (Bảng 3) cho thấy một số biến có mức tương quan trung bình đến tương đối cao, bao gồm X25 (6,76), X12 (6,70), X2 (6,46), X4 (5,93) và X13 (5,25). Điều này phản ánh mối liên hệ cấu trúc giữa

các đặc điểm hạ tầng, điều kiện vận hành và loại hình va chạm trong dữ liệu tai nạn giao thông. Tuy nhiên, các giá trị này vẫn chưa vượt ngưỡng đa cộng tuyến nghiêm trọng thường được chấp nhận trong hồi quy logistic ($VIF > 10$). Ngoài ra, mô hình không ghi nhận hiện tượng bất ổn định đáng kể về hệ số hồi quy hoặc sai số chuẩn. Do đó, các biến vẫn được giữ lại trong mô hình nhằm đảm bảo tính đầy đủ về mặt lý thuyết và kỹ thuật.

Kết quả ước tính của các hệ số (α/β), giá trị P-value, và tỷ lệ Chênh (Odds Ratio - OR) được trình bày trong Bảng 3. Các nhận xét có thể được rút ra từ kết quả phân tích như sau:

- Sự tương tác giữa hạ tầng và loại hình va chạm: Mặc dù các biến hạ tầng như số làn xe (X1), dải phân cách (X2) và làn dừng khẩn cấp (X3) không đạt ý nghĩa thống kê trong mô hình hồi quy ($P > 0,05$), nhưng thống kê mô tả cho thấy chúng là "nền tảng" cho các vụ tai nạn. Mô hình hồi quy xác nhận biến X15 (Lật xe/lao ra khỏi đường) có ý nghĩa thống kê mạnh ($P\text{-value} = 0,024$) với tỷ số chênh OR = 4,178. Điều này gợi ý rằng khi hạ tầng không có làn dừng khẩn cấp (X3 chiếm 89,41% số vụ), lái xe khi gặp sự cố dễ dẫn đến hành vi mất lái và lật xe, khiến mức độ nghiêm trọng của tai nạn tăng vọt gấp hơn 4 lần so với các loại hình khác. Kết quả kiểm định đa cộng tuyến cho thấy mặc dù các biến hạ tầng (số làn xe, dải phân cách cứng, làn dừng khẩn cấp) có mối quan hệ cấu trúc do đặc điểm thiết kế tuyến, nhưng không tồn tại đa cộng tuyến nghiêm trọng về mặt thống kê. Điều này cho thấy việc các biến này không đạt ý nghĩa thống kê riêng lẻ trong mô hình không phải do vấn đề kỹ thuật, mà phản ánh bản chất tương tác của hệ thống hạ tầng. Các yếu tố này đóng vai trò như các biến nền, ảnh hưởng gián tiếp đến mức độ nghiêm trọng của tai nạn thông qua các loại hình va chạm cụ thể, điển hình là tai nạn lật xe hoặc lao ra khỏi đường (X15), vốn có ý nghĩa thống kê cao trong mô hình.

- Về thời gian và đặc điểm giao thông các ngày trong tuần: Một kết quả đáng chú ý từ mô hình hồi

quy là biến X9 (Ngày thường) có ý nghĩa thống kê cao ($P\text{-value} = 0,013$) với OR = 2,445. Kết quả này củng cố cho phân tích mô tả về sự hiện diện của xe tải (chiếm trên 71% số vụ) thường tập trung vào các ngày làm việc trong tuần. Mặc dù thống kê mô tả cho thấy tỷ lệ tai nạn bình quân mỗi ngày cuối tuần cao hơn, nhưng mô hình hồi quy lại khẳng định các vụ tai nạn xảy ra vào ngày thường có rủi ro nghiêm trọng cao gấp 2,445 lần. Điều này phản ánh hệ quả của việc pha trộn dòng giao thông giữa xe tải nặng và xe khách vào ngày thường, tạo ra các xung đột giao thông khốc liệt hơn so với sự gia tăng thuần túy của xe cá nhân vào cuối tuần.

- Tác động của khung giờ xảy ra tai nạn và hình thức xảy ra tai nạn: Thống kê mô tả chỉ ra hai "điểm nóng" về thời gian là 0h-6h (33,53%) và 12h-18h (31,76%), liên quan trực tiếp đến tình trạng buồn ngủ và mệt mỏi của tài xế. Tuy nhiên, trong mô hình hồi quy, biến X25 (Tự gây tai nạn) lại cho thấy hệ số β âm (-3,507) và OR thấp (0,03) với $P\text{-value} = 0,001$. Điều này chỉ ra một quy luật quan trọng: các vụ tai nạn do buồn ngủ dẫn đến tự mất lái, đâm vào hộ lan hoặc tự lật (dù chiếm tỷ lệ lớn về số lượng) thường có mức độ thương vong thấp hơn so với các vụ va chạm liên phương tiện. Ngược lại, lỗi hành vi "Không giữ khoảng cách an toàn" (X12 - 41,95%) khi kết hợp với tốc độ cao thường dẫn đến các vụ va chạm trực diện, vốn là tác nhân chính gây ra hậu quả thảm khốc trên các tuyến cao tốc hạ tầng hạn chế.

- Về các biến bị loại bỏ và tính tối ưu của mô hình: Việc loại bỏ các biến tham chiếu X8 (18h-24h) và X10 (Cuối tuần) cùng các biến có dữ liệu thống kê thấp (X18, X19, X22, X23, X24) đã giúp mô hình tập trung vào các nhân tố thực sự có sức mạnh giải thích. Các biến như tốc độ giới hạn (X4) không có ý nghĩa trong hồi quy ($P\text{-value} = 0,384$), có thể phản ánh tốc độ không phải là biến độc lập gây tai nạn, mà rủi ro thực sự nằm ở sự kết hợp giữa hạ tầng thiếu làn dừng và trạng thái tâm lý ức chế của lái xe khi vận hành trên các đoạn tuyến hẹp.

Bảng 3. Kết quả ước tính mô hình hồi quy Logit thứ bậc cho các yếu tố ảnh hưởng đến Mức độ nghiêm trọng TNGT.

Hạng mục	Biến số	Tên biến	Hệ số (β)	VIF	P-value	Tỷ lệ chênh (OR)	95% CI
Hạ tầng	X1	Số làn xe	-0,303	4,12	0,641	0,738	[0,206 - 2,647]

Hạng mục	Biến số	Tên biến	Hệ số (β)	VIF	P-value	Tỷ lệ chênh (OR)	95% CI
Thời gian & Môi trường	X2	Dải phân cách cứng	-0,133	6,46	0,866	0,875	[0,186 - 4,109]
	X3	Làn dừng khẩn cấp	0,075	1,22	0,884	1,077	[0,396 - 2,933]
	X4	Tốc độ cho phép	-0,66	5,93	0,384	0,517	[0,117 - 2,281]
	X5	Khung giờ 0h-6h	-0,035	3,05	0,951	0,966	[0,323 - 2,886]
	X6	Khung giờ 6h-12h	-0,202	2,69	0,729	0,817	[0,261 - 2,556]
	X7	Khung giờ 12h-18h	-0,512	3,06	0,357	0,6	[0,202 - 1,78]
	X9	Ngày thường (T2-T6)	1	1,16	0,013	2,445	[1,206 - 4,956]
	X11	Điều kiện Mưa	0,119	1,27	0,78	1,127	[0,488 - 2,6]
	X12	Không giữ K/C an toàn	-1	6,70	0,471	0,579	[0,131 - 2,56]
	X13	Đối đầu	0,625	5,25	0,415	1,868	[0,415 - 8,402]
Đặc điểm va chạm	X14	Đâm vào xe dừng đỗ	1	4,35	0,248	2,566	[0,519 - 12,694]
	X15	Lật xe/lao ra khỏi đường	1,430	2,38	0,024	4,178	[1,21 - 14,427]
	X16	Xe tải và Xe khách	0	2,69	0,536	1,488	[0,422 - 5,251]
	X17	Xe tải và Xe tải	-1	3,05	0,136	0,422	[0,136 - 1,312]
	X20	Xe tải và Xe con	-1,273	2,12	0,075	0,28	[0,069 - 1,138]
Loại phương tiện	X21	Tai nạn liên hoàn	-1	1,88	0,305	0,44	[0,091 - 2,115]
	X25	Tự gây tai nạn	-3,507	6,76	0,001	0,03	[0,004 - 0,241]
	α_1	Ngưỡng không nghiêm trọng sang nhẹ	-3,981	-	-	-	-
Ngưỡng giữa 4 mức độ tai nạn (Cut-off points)	α_2	Ngưỡng tai nạn nhẹ sang nặng	-1	-	-	-	-
	α_3	Ngưỡng tai nạn nặng sang rất nặng	0,738	-	-	-	-

Model fit statistics: Log-likelihood = -176,27; AIC = 394,55; McFadden's R^2 = 0,165.

4. Kết luận và kiến nghị

Nghiên cứu đã phân tích 170 vụ tai nạn trên các tuyến cao tốc phân kỳ tại Việt Nam, sử dụng thống kê mô tả và mô hình hồi quy Logit thứ bậc để lượng

hóa tác động của hạ tầng đến mức độ nghiêm trọng của TNGT. Nghiên cứu đã chỉ ra rằng rủi ro tai nạn trên đường cao tốc không đến từ một yếu tố đơn lẻ mà là sự tương tác phức tạp giữa hạ tầng, hành vi và thời gian.

Các kết luận chính có thể rút ra từ kết quả nghiên cứu:

- Về hạ tầng: Mặc dù mô hình hồi quy chưa cho thấy tác động trực tiếp có ý nghĩa thống kê của các biến hạ tầng đơn lẻ (X1, X2, X3) đến mức độ nghiêm trọng của tai nạn, nghiên cứu gợi ý rằng đặc điểm hạ tầng cao tốc phân kỳ đóng vai trò là yếu tố nền tảng. Sự thiếu vắng dải phân cách cứng hay làn dừng khẩn cấp liên tục tạo ra một môi trường vận hành nhạy cảm, nơi các sai sót hành vi (như lấn làn, lật xe) dễ dàng chuyển hóa thành các vụ tai nạn thảm khốc. Do đó, các kiến nghị về việc nâng cấp hạ tầng cần được xem xét như một giải pháp phòng ngừa rủi ro thụ động, giúp giảm thiểu hậu quả khi tai nạn xảy ra;

- Về loại phương tiện và khung giờ: Xe tải là đối tượng trung tâm khi tham gia vào hơn 71% tổng số vụ tai nạn. Đặc biệt, hai khung giờ "điểm nóng" là 0h-6h và 12h-18h (chiếm tổng cộng hơn 65%) cho thấy sự suy giảm về trạng thái sinh lý (buồn ngủ, mệt mỏi) của lái xe có thể dẫn đến tai nạn;

- Về yếu tố thời gian: Mô hình hồi quy khẳng định rủi ro nghiêm trọng vào ngày thường cao gấp 2,445 lần so với cuối tuần. Điều này phản ánh áp lực từ dòng xe tải và xe khách lưu thông với mật độ cao vào những ngày làm việc, gây ra các xung đột giao thông có mức độ khốc liệt lớn hơn;

- Về yếu tố môi trường: Thời tiết không phải là nguyên nhân chính khi có tới 81,98% vụ tai nạn xảy ra trong điều kiện bình thường, cho thấy tâm lý chủ quan của lái xe khi đường vắng và thời tiết tốt.

Từ kết quả nghiên cứu, các giải pháp cần được thực hiện: (1) Về đầu tư và quy hoạch: Các cơ quan quản lý cần coi việc bổ sung dải phân cách cứng và làn dừng khẩn cấp liên tục là yêu cầu bắt buộc đối với mọi dự án cao tốc mới và nâng cấp. Mặc dù mô hình không tìm thấy ảnh hưởng lên mức độ nghiêm trọng, việc hoàn thiện hạ tầng là cần thiết để ngăn ngừa các vụ va chạm năng lượng cao (như đối đầu và đâm xe dừng) xảy ra, qua đó giảm tần suất TNGT. Ưu tiên nguồn vốn khẩn cấp để nâng cấp ngay lập tức các đoạn tuyến 4 làn hạn chế (thiếu làn dừng khẩn cấp - chiếm 89,41% số vụ) và 2 làn xe hiện hữu. (2) Về giải pháp ngắn hạn: Tăng cường giám sát và xử phạt nghiêm khắc các hành vi dừng đỗ trái phép trên làn đường và không giữ khoảng cách an toàn. Ngoài ra, cũng cần khuyến

cáo tài xế không lái xe liên tục trong thời gian dài và chủ động nghỉ ngơi trong các khung giờ dễ xảy ra tai nạn (sau ăn trưa và rạng sáng).

Đóng góp của các tác giả trong bài báo

Lê Xuân Quý: Điều tra, Phân tích chính thức, Phản hồi ý kiến phản biện, Viết – bản thảo gốc, Chỉnh sửa bản thảo. **Chu Tiến Dũng:** Phương pháp, Điều tra, Tiếp nhận tài trợ, Giám sát. **Phan Văn Mèn:** Điều tra, Biên soạn dữ liệu, Phân tích dữ liệu, Viết – bản thảo gốc.

Tuyên bố không xung đột lợi ích và cam kết bản quyền

Các tác giả tuyên bố về sự không xuất hiện những xung đột tiềm ẩn từ nghiên cứu này, và cam kết bài báo chưa từng được công bố trước đây.

Chia sẻ dữ liệu theo yêu cầu

Dữ liệu sẽ được cung cấp theo yêu cầu

Lời cảm ơn.

Nghiên cứu này được tài trợ bởi Bộ Giáo dục và Đào tạo (MoET) trong Đề tài mã số B2025-GHA-01.

1st Le Xuan Quy*. *University of Transport and Communications.*

2nd Chu Tien Dung. *University of Transport and Communications.*

3rd Phan Van Men. *University of Transport and Communications.*

*Corresponding author: lexuanquy@utc.edu.vn

Tài liệu tham khảo

- [1] TCCS 42:2022/TCĐBVN, *Đường ô tô cao tốc - Thiết kế và tổ chức giao thông trong giai đoạn phân kỳ đầu tư xây dựng*, Hà Nội, 2022.
- [2] TCVN 5729:2012, *Đường ô tô cao tốc - Yêu cầu thiết kế*, Hà Nội, 2012.
- [3] Đặng Minh Tân and Vũ Quang Huy, "Vấn đề sử dụng làn ở một số tuyến đường cao tốc khu vực phía Bắc và đề xuất nâng cao hiệu quả tổ chức an toàn giao thông," *Hội nghị An toàn giao thông Việt Nam*, Hà Nội, 2023.
- [4] Vũ Minh Tâm, "Nghiên cứu quy trình đề xuất phương án tổ chức giao thông phân làn trên đường cao tốc," *Tạp chí Khoa học Giao thông vận tải*, vol. 76, no. 5, pp. 779–792, Jun. 2025, doi: 10.47869/tcsj.76.5.11.
- [5] X. Q. Le, T. D. Chu, and T. A. Vu, "A case study on scenario-based optimization for axle load survey

- data selection for flexible pavement design in Vietnam," *Transport and Communications Science Journal*, vol. 75, no. 4, pp. 1555–1566, May 2024, doi: 10.47869/tcsj.75.4.7.
- [6] Chu Tiến Dũng and Lê Xuân Quý, "Ứng dụng mô hình poisson và mô hình nhị thức âm trong dự báo tai nạn giao thông trên đường cao tốc," *Tạp chí Khoa học Giao thông vận tải*, vol. 76, no. 6, pp. 821–835, Aug. 2025, doi: 10.47869/tcsj.76.6.1.
- [7] K. Geurts, G. Wets, T. Brijs, and K. Vanhoof, "Identification and Ranking of Black Spots: Sensitivity Analysis," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1897, no. 1, pp. 34–42, Jan. 2004, doi: 10.3141/1897-05.
- [8] AASHTO, *Highway Safety Manual*, 1st Edition. American Association of State Highway and Transportation Officials, 2010.
- [9] W. Greene, *Econometric analysis*, Eighth edition. New York, NY: Pearson, 2018.
- [10] S. Washington, M. G. Karlaftis, F. Mannering, and P. Anastasopoulos, *Statistical and Econometric Methods for Transportation Data Analysis*, 3rd ed. New York: Chapman and Hall/CRC, 2020. doi: 10.1201/9780429244018.
- [11] T. Sayed and E. Sacchi, "Evaluating the safety impact of increased speed limits on rural highways in British Columbia," *Accident Analysis & Prevention*, vol. 95, pp. 172–177, Oct. 2016, doi: 10.1016/j.aap.2016.07.012.
- [12] A. S. Al-Ghamdi, "Using logistic regression to estimate the influence of accident factors on accident severity," *Accident Analysis & Prevention*, vol. 34, no. 6, pp. 729–741, Nov. 2002, doi: 10.1016/S0001-4575(01)00073-2.
- [13] D. W. H. Jr, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*. John Wiley & Sons, 2013.
- [14] K. M. Kockelman and Y.-J. Kweon, "Driver injury severity: an application of ordered probit models," *Accident Analysis & Prevention*, vol. 34, no. 3, pp. 313–321, May 2002, doi: 10.1016/S0001-4575(01)00028-8.
- [15] P. McCullagh, "Regression Models for Ordinal Data," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 42, no. 2, pp. 109–127, Jan. 1980, doi: 10.1111/j.2517-6161.1980.tb01109.x.
- [16] C. J. O'Donnell and D. H. Connor, "Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice," *Accident Analysis & Prevention*, vol. 28, no. 6, pp. 739–753, Nov. 1996, doi: 10.1016/S0001-4575(96)00050-4.