Evaluating Safety Performance of Automated Shuttles Through Hard Braking Analysis in European Urban Pilots

Marios Sekadakis[1*][0000-0002-7586-9283], Apostolos Ziakopoulos[1][0000-0001-6252-6743], and George Yannis[1][0000-0002-2196-2335]

[1] National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Iroon Polytechniou St, Athens GR-15773, Greece
*msekadakis@mail.ntua.gr

Abstract. The emergence of Cooperative, Connected, and Automated Mobility (CCAM) is progressing steadily, yet its safety implications remain only partially understood. This study investigates the safety performance of automated shuttles by analyzing Hard Braking (HB) incidents collected from real-world deployments across multiple European urban pilot sites. The analysis incorporated data from ten distinct locations involved in the SHOW project: Brno, Carabanchel, Graz, Karlsruhe, Klagenfurt, Les Mureaux, Linköping, Pörtschach, Tampere, and Trikala. The dataset provides broad geographical and contextual representation encompassing a total of 1,796 daily records and 4,820 HB events. A Negative Binomial regression model was used to correlate key explanatory variables including average speed, acceleration variance, as well as the pilot site itself. Multicollinearity was found to be negligible, as confirmed through variance inflation factors, while Marginal Effects to the Mean (MEM) were estimated to assess the influence of each variable. Results indicated that both average speed and acceleration variance significantly contributed to increased HB occurrences. Linköping exhibited the fewest HB events, whereas Klagenfurt had the highest, with Graz, Pörtschach, and Trikala also showing elevated rates. These findings emphasize the impact of operational conditions on shuttle safety and support the development of site-specific safety enhancements.

Keywords: Naturalistic Pilots, Negative Binomial Model, Automated Shuttles, Harsh Events, Traffic Safety.

1 Introduction

Cooperative, Connected and Automated Mobility (CCAM) is moving from controlled pilots to everyday operations. As automated shuttles begin serving public transport roles, safety performance in mixed-traffic conditions becomes a primary concern. Among the available operational indicators, hard braking events offer a direct lens on how often shuttles are forced into abrupt deceleration, signalling tight interactions with other road users, late detections, or conservative control strategies. While crash data remain sparse for automated fleets, hard braking counts provide a surrogate safety measure that can be monitored continuously and compared across sites [1–3].

Most evidence to date on automated vehicle safety stems from closed-course trials, short demonstrations, or single-site studies. These settings rarely capture the variability of European urban traffic: differing street layouts, pedestrian activity, cycling volumes, public transport priority, speed management, and weather patterns. As a result, conclusions drawn in one context may not transfer to another. Multi-site analyses can close this gap by linking vehicle dynamics to local operating environments [4].

This study was conducted within the framework of the SHOW project (SHared automation Operating models for Worldwide adoption), a Horizon 2020 initiative deploying automated shuttles across European cities to assess their impacts on safety, efficiency, and user acceptance [5]. This study investigates hard braking in automated shuttles deployed across ten SHOW pilot locations: Brno, Carabanchel, Graz, Karlsruhe, Klagenfurt, Les Mureaux, Linköping, Pörtschach, Tampere, and Trikala. The dataset spans 1,796 site-days and 4,820 hard braking events, offering sufficient statistical power for comparative analysis. The study focuses on two interpretable covariates that are routinely available in operations dashboards, average speed and acceleration variance, while explicitly modelling site effects to reflect local context. Negative Binomial regression is adopted to address overdispersion in daily event counts and to yield incidence rate ratios and marginal effects that are meaningful to operators and policymakers [6].

The work makes three contributions. First, it delivers one of the earliest cross-European, real-world assessments of automated shuttle safety using a common telemetry-based definition of hard braking. Second, it quantifies the association between speed, acceleration variability and hard braking frequency, translating coefficients into marginal effects to support operational decision-making (e.g., expected change in events for a 1 km/h speed increase). Third, it reveals substantial site-to-site differences after controlling for vehicle dynamics, highlighting the role of local traffic conditions, route design, and interaction density with vulnerable road users.

The findings speak directly to safety management in automated public transport. If small increases in speed or instability in longitudinal control materially elevate hard braking counts, operators may prioritize smoother profiles, revised speed targets by segment, or adaptive behaviours that anticipate conflicts rather than relying on conservative low speeds alone. Likewise, systematic site differences point to where targeted interventions, signal timing, curb management, crossing protection, or perception tuning, are likely to yield measurable gains.

2 Methodology

2.1 Data Collection

The analysis draws on operational data from the Horizon 2020 SHOW project, which deployed automated shuttles across multiple European cities to evaluate their safety, efficiency, and integration into public transport. Ten pilot locations are included in this study: Brno, Carabanchel (Madrid), Graz, Karlsruhe, Klagenfurt, Les Mureaux, Linköping, Pörtschach, Tampere, and Trikala.

Across these sites, automated shuttles recorded 1,796 site-days of operation and 4,820 Hard Braking (HB) events. Hard braking was defined consistently across pilots as longitudinal decelerations exceeding 3 m/s^2, a threshold set by operators in line with earlier traffic safety studies where abrupt deceleration is used as a surrogate safety measure for near-crash conditions. Unlike self-reported events, all HB counts were extracted directly from vehicle telemetry, ensuring objective and comparable measurements across pilots. The SHOW Dashboard (https://show-project.eu/show-dashboard/) offers a centralized platform for accessing real-time and historical performance data from all pilot sites, including vehicle telemetry, route metrics, and safety indicators, facilitating standardised monitoring and cross-site comparison.

2.2 Variables

Three categories of variables are incorporated:

- Average speed (km/h): mean daily operating speed of the shuttles on their designated routes.
- Acceleration variance $(m/s^2)^2$: daily variation in longitudinal acceleration, reflecting stability of vehicle control.
- Pilot site (categorical): ten sites treated as factor levels, with Linköping selected as the reference category due to its large number of observations and lowest raw HB counts.

These variables were selected to balance parsimony with operational interpretability. Both speed and acceleration variance are routinely monitored in fleet dashboards, while site dummies capture context-specific heterogeneity such as route design, traffic density, or interaction with Vulnerable Road Users (VRUs).

2.3 Statistical Modelling

Hard braking events were modelled using a Negative Binomial (NB) regression framework, appropriate for count data exhibiting overdispersion. Overdispersion was confirmed in the dataset, with the variance of daily HB counts exceeding the mean by a factor of 20.15. The NB specification introduces a dispersion parameter to correct for this, offering more reliable coefficient estimates than the Poisson model [7].

Site effects were first tested as random intercepts but yielded no additional explanatory power relative to fixed effects. Therefore, the final specification retained fixed site dummies with Linköping as the reference.

2.4 Diagnostics and Marginal Effects

To ensure model robustness, Variance Inflation Factors (VIFs) were computed. The highest observed VIF was 2.333, well below the conventional cut-off of 5, indicating negligible multicollinearity [8].

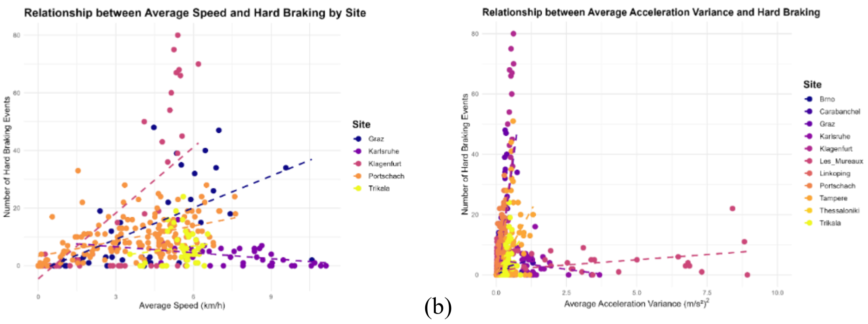
To aid interpretation, Marginal Effects at the Mean (MEMs) were derived for continuous predictors (speed and acceleration variance) as well as site dummies. MEMs quantify the change in expected HB counts associated with a one-unit increase in a predictor, holding other variables at their mean values. This approach allows results to be translated into operationally relevant terms (e.g., additional HB events per 1 km/h increase in mean speed).

3 Results and Discussion

3.1 Descriptive Patterns

Across the ten SHOW pilot sites, a total of 4,820 HB events were recorded over 1,796 site-days. Daily averages ranged from fewer than 0.1 events in Linköping to more than 15 events in Klagenfurt. Pörtschach, Graz, and Trikala also exhibited elevated frequencies, while Brno, Les Mureaux, and Tampere reported relatively lower values. These initial differences suggest that site-specific operational and traffic contexts strongly shape shuttle safety performance.

Scatterplots were generated to explore bivariate relationships. Fig. 1 (a) indicates only a weak linear relationship between average speed and HB counts, with site-specific R^2 values remaining low. Nonetheless, some locations (e.g., Graz) displayed greater variability at higher speeds, likely reflecting denser urban conditions and interactions with Vulnerable Road Users (VRUs).



(a) (b)

Fig. 1. (a) Scatterplot of SHOW shuttle hard brakings along with its speed
(b) Scatterplot of SHOW shuttle hard brakings along with its acceleration variance

Fig. 1 (b) examined acceleration variance, where upward-sloping trends suggested that unstable speed control may contribute to HB frequency. Yet, as with speed, relationships differed substantially across sites, indicating that acceleration variance alone cannot fully explain observed differences. Together, these descriptive patterns illustrate that although speed, acceleration variance play a role, HB dynamics arise from a multifactor interaction shaped by both vehicle behaviour and site environment.

3.2 Negative Binomial Regression

The regression results are presented in Table 1. Both continuous predictors were statistically significant at the 0.001 level. A 1 km/h increase in average speed was associated with a 34% higher expected HB rate, while a one-unit increase in acceleration variance increased the rate by approximately 5%. These findings confirm that both faster operation and greater instability in longitudinal control elevate the likelihood of abrupt braking.

Table 1. Negative Binomial Regression Results For Hard Braking Events

Variable	Estimate	Std. Error	z-value	p-value	
Intercept	-4.938	0.232	-21.267	<0.0001	***
Average Speed	0.292	0.024	12.256	<0.0001	***
Average Acceleration Variance	0.052	0.010	5.133	<0.0001	***
Site: Brno [Ref. Cat. Linköping]	1.579	0.302	5.225	<0.0001	***
Site: Carabanchel [Ref. Cat. Linköping]	3.262	0.210	15.500	<0.0001	***
Site: Graz [Ref. Cat. Linköping]	5.962	0.287	20.790	<0.0001	***
Site: Karlsruhe [Ref. Cat. Linköping]	4.098	0.289	14.164	<0.0001	***
Site: Klagenfurt [Ref. Cat. Linköping]	6.179	0.269	22.963	<0.0001	***
Site: Les Mureaux [Ref. Cat. Linköping]	2.738	0.183	14.933	<0.0001	***
Site: Pörtschach [Ref. Cat. Linköping]	6.057	0.210	28.894	<0.0001	***
Site: Tampere [Ref. Cat. Linköping]	1.455	0.330	4.407	<0.0001	***
Site: Trikala [Ref. Cat. Linköping]	5.127	0.268	19.101	<0.0001	***
Dependent variable: Hard Braking Counts per day Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Null deviance: 3109.5 on 1795 df. Residual deviance: 1218.4 on 1784 df. AIC: 4843.4					

Individual effects that were native to each site were also pronounced. Compared to Linköping (reference), all other sites showed significantly higher HB rates. Klagenfurt, Pörtschach, Graz, and Trikala emerged as outliers, with Incidence Rate Ratios (IRRs) several orders of magnitude above the baseline. In contrast, Brno, Les Mureaux, and Tampere displayed more modest increases. These patterns reinforce that local operational context, including traffic density, route design, and pedestrian activity, plays a decisive role in shaping shuttle safety performance.

Model diagnostics confirmed robustness. The Variance Inflation Factor (VIF) values did not exceed 2.333, ruling out multicollinearity concerns. Residual deviance (1218.4 on 1784 df) indicated good model fit relative to the null model (deviance = 3109.5).

3.3 Marginal Effects

Table 2 presents Marginal Effects at the Mean (MEMs), translating coefficients into operationally interpretable impacts. On average, a 1 km/h speed increase resulted in 0.155 additional daily HB events, while a one-unit rise in acceleration variance yielded 0.863 additional daily events.

Table 2. Marginal Effects to the Mean (MEM)

Marginal Effects	Estimate	SE	z	p	lower	upper
Average Speed	0.155	0.036	4.304	0.000	0.084	0.225
Average Acceleration Variance	0.863	0.121	7.133	0.000	0.626	1.101
Site: Brno [Ref. Cat. Linköping]	0.646	0.204	3.170	0.002	0.246	1.045
Site: Carabanchel [Ref. Cat. Linköping]	4.205	0.884	4.755	0.000	2.472	5.938
Site: Graz [Ref. Cat. Linköping]	64.933	22.191	2.926	0.003	21.440	108.426
Site: Karlsruhe [Ref. Cat. Linköping]	9.920	2.931	3.385	0.001	4.176	15.664
Site: Klagenfurt [Ref. Cat. Linköping]	80.705	27.912	2.891	0.004	25.999	135.412
Site: Les Mureaux [Ref. Cat. Linköping]	2.422	0.377	6.419	0.000	1.682	3.161
Site: Pörtschach [Ref. Cat. Linköping]	71.382	20.510	3.480	0.001	31.183	111.581
Site: Tampere [Ref. Cat. Linköping]	0.551	0.137	4.021	0.000	0.282	0.819
Site: Trikala [Ref. Cat. Linköping]	28.065	8.656	2.242	0.001	11.100	45.030

Site-specific MEMs highlighted considerable heterogeneity. For example, Klagenfurt and Pörtschach recorded over 70 additional daily HB events relative to Linköping, while Carabanchel and Les Mureaux showed moderate increases (4-5 events per day). Lower MEM values for Brno and Tampere suggest comparatively stable operations, possibly due to more predictable traffic conditions or less complex interactions.

3.4 Interpretation and Implications

The regression analysis supports three main insights. First, both speed and stability of acceleration control are significant determinants of hard braking. Even small increases in these variables substantially elevate HB risk, implying that smoother longitudinal control should be prioritized in shuttle programming. Second, contextual heterogeneity is large: site differences persist even after controlling for vehicle dynamics, indicating that safety outcomes are strongly mediated by local environments. Third, automation share alone does not explain safety performance, suggesting that adaptive control strategies tailored to urban density and interaction patterns are more relevant than automation exposure per se. These results are consistent with earlier research linking speed management and acceleration variability to safety-critical events in mixed traffic [9]. They also repeat recent findings that local context remains a dominant determinant of AV safety outcomes.

In addition to confirming the importance of vehicle dynamic parameters, model outputs also showcase the potential of harsh braking events as Surrogate Safety Measures (SSMs) and as an operational safety Key Performance Indicator (KPI) for automated shuttles, which are analysed proactively to anticipate the advent of automation [10]. Due to the fact that HB events are relatively frequent compared to crashes, they provide a practical, real-time signal that operators can monitor through fleet dashboards. Integrating HB thresholds into automated shuttle control logic or adaptive scheduling could enable proactive interventions, such as adjusting maximum speeds on specific route segments or reprogramming acceleration profiles in areas prone to conflicts with Vulnerable Road Users (VRUs). This new technical venue opens a pathway for using SSMs not just for retrospective evaluation, but also for continuous operational management of automated fleets, also with unique possibilities of new analytic metrics [11]. This can also be a promising validation mechanism for the simulated transport networks which are created in anticipation of the advent of automation [12].

Another key implication concerns transferability across pilot sites. While the negative binomial model controlled for vehicle dynamics, the strong site effects suggest that safety outcomes are not easily generalizable across contexts. A model calibrated in Klagenfurt, for example, would likely overestimate risks in Tampere or Brno, leading to overly conservative operations. Conversely, ignoring local heterogeneity could underestimate safety risks in dense or complex traffic environments. This underlines the need for site-sensitive calibration of automated vehicle algorithms, possibly through dynamic geofencing or context-aware AI modules that adapt control strategies based on local operating conditions. Such adaptive frameworks would ensure that lessons from individual pilots can inform broader deployments without assuming that one model fits all.

4 Conclusions and Future Work

This paper evaluated the safety performance of automated shuttles by analysing HB events across ten European pilot sites within the Horizon 2020 SHOW project. Using 1,796 site-days of data and 4,820 objectively recorded HB events, a Negative Binomial regression framework was applied to address overdispersion in daily event counts. The analysis produced several key findings. First, vehicle dynamics strongly influenced HB frequency. Higher average speeds and greater acceleration variance significantly increased the number of abrupt braking events. Even marginal changes in these variables translated into measurable increases in daily HB counts, highlighting the operational importance of speed management and stable longitudinal control. Second, substantial site-to-site heterogeneity was observed. Locations such as Klagenfurt, Pörtschach, Graz, and Trikala exhibited much higher HB rates than Linköping, the reference site, even after accounting for vehicle dynamics. These differences likely reflect contextual influences such as road design, traffic density, pedestrian volumes, and local operational practices. By contrast, Brno, Tampere, and Les Mureaux demonstrated comparatively stable operations with lower HB counts. Together, these findings underscore that context-aware automation strategies are essential for safe integration of shuttles into mixed urban traffic. Adaptive speed regulation, smoother acceleration control, and location-specific operational tuning may help reduce abrupt braking events and enhance passenger comfort and safety.

Future research should expand on this foundation in four directions. First, broader safety indicators such as near-crash metrics and conflict severity measures would provide a richer perspective on automated shuttle safety. Second, contextual integration is needed by linking hard braking data with traffic volumes, vulnerable road user densities, and road geometry to allow more precise attribution of site differences. Third, human factors should be examined, including passenger comfort, public perception, and interactions with VRUs, which are critical for societal acceptance. Finally, technological evolution must be considered, as advances in sensor fusion, vehicle-to-infrastructure communication, and AI-driven control are expected to influence braking dynamics and overall safety outcomes.

Acknowledgments
The SHOW project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 875530.

5 References

1. Nikolaou, D., Ziakopoulos, A., Yannis, G.: A Review of Surrogate Safety Measures Uses in Historical Crash Investigations. Sustain. 15, (2023). https://doi.org/10.3390/su15097580.
2. Noy, I.Y., Shinar, D., Horrey, W.J.: Automated driving: Safety blind spots. Saf. Sci. 102, 68–78 (2018). https://doi.org/10.1016/j.ssci.2017.07.018.
3. Papadoulis, A., Quddus, M., Imprialou, M.: Evaluating the safety impact of connected and autonomous vehicles on motorways. Accid. Anal. Prev. 124, 12–22 (2019). https://doi.org/10.1016/j.aap.2018.12.019.
4. Kim, W., Romo, A., McDonough, J.: A Multi-site Examination for the Impact of Changes in Posted Speed Limit on Traffic Safety. AAA Found. Traffic Saf. (2023).
5. EU H2020 SHOW project: SHared automation Operating models for Worldwide adoption, https://show-project.eu/.
6. Wang, C., Chen, F., Cheng, J., Bo, W., Zhang, P., Hou, M., Xiao, F.: Random-parameter multivariate negative binomial regression for modeling impacts of contributing factors on the crash frequency by crash types. Discret. Dyn. Nat. Soc. 2020, (2020). https://doi.org/10.1155/2020/6621752.
7. Lord, D., Washington, S.P., Ivan, J.N.: Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accid. Anal. Prev. 37, 35–46 (2005). https://doi.org/10.1016/j.aap.2004.02.004.
8. O’Brien, R.M.: A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41, 673–690 (2007). https://doi.org/10.1007/s11135-006-9018-6.
9. Torrão, G.: Speed related variables for crash injury risk analysis : what has been used ? Int. J. Crashworthiness. 27, 1560–1567 (2022). https://doi.org/10.1080/13588265.2021.1959152.
10. Ziakopoulos, A., Oikonomou, M.G., Vlahogianni, E.I., Yannis, G.: Quantifying the implementation impacts of a point to point automated urban shuttle service in a large-scale network. Transp. Policy. 114, 233–244 (2021). https://doi.org/10.1016/j.tranpol.2021.10.006.
11. Ziakopoulos, A., Oikonomou, M.G., Sekadakis, M., Yannis, G.: Safety evaluation via conflict classification during automated shuttle bus service operations. Eur. Transp. Res. Rev. 16, 38 (2024). https://doi.org/10.1186/s12544-024-00653-1.
12. Oikonomou, M.G., Yannis, G.: Validating traffic simulation for crash risk assessment using field crash data. J. Safety Res. 94, 24–34 (2025). https://doi.org/10.1016/j.jsr.2025.06.002.