A Hybrid Machine Learning and Statistical Modeling Approach for Analyzing the Crash Severity of Mobility Scooter Users Considering Temporal Instability

Matin Sadeghi a, Kayvan Aghabayk a, *\**, Mohammed Quddusb

*a School of Civil Engineering, College of Engineering, University of Tehran, Iran*

*b Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, UK*

*\* Corresponding author's E-mail address: kayvan.aghabayk@ut.ac.ir*

**Abstract**

One of the main objectives in improving the quality of life for individuals with disabilities, especially those experiencing mobility issues such as the elderly, is to enhance their day-to-day mobility. Enabling easy mobility contributes to their independence and access to better healthcare, leading to improvements in both physical and mental well-being. Mobility Scooters have become increasingly popular in recent years as a means of facilitating mobility, yet traffic safety issues such as crash severity have not been adequately investigated in the literature. This study addresses this knowledge gap by employing a hybrid method that combines a machine learning approach using the eXtreme Gradient Boosting (XGBoost) algorithm with Shapley Additive exPlanations (SHAP) and an advanced statistical model called Random Parameters Binary Logit accounting for heterogeneity in means and variances. Analyzing the United Kingdom mobility scooter crash data from 2018 to 2022, the study examined temporal instability using a likelihood ratio test. The results revealed that there was instability over the three distinct periods of time based on the coronavirus (COVID) pandemic, namely, Pre-COVID, during COVID, and post-COVID. Moreover, the results revealed that mobility scooter crashes occurring at a give-way or uncontrolled junctions has a random effect on the severity, while factors such as mobility scooter riders aged over 80, rear-end and sideswipe crashes, and crashes during winter months increase the risk of severe injuries. Conversely, mobility scooter riders involved in crashes while riding on the footway are less likely to experience severe injuries. These findings offer valuable insights for enhancing road safety measures that can be utilized to effectively reduce the crash severity of mobility scooter riders.

**Keywords:**

Random parameters binary logit, Unobserved heterogeneity, XGBoost, SHAP, COVID impact, Mobility Scooter Users

1. **Introduction**

The World Health Organization (WHO) reports that 16% of the global population confronts diverse disabilities, leading to distinct challenges and health disparities (WHO 2023). People with disabilities, making up 1 in 6 individuals, often experience a higher risk of depression, stroke, and other chronic diseases due to factors such as immobility and limited access to healthcare services (WHO 2013). Therefore, this segment of the population, comprising individuals with disabilities and vulnerabilities, necessitates and requests increased attention and support.

Various types of disabilities affect diverse population groups, with one of the fastest-growing demographics in the past decade being the elderly. Notably, the elderly population has witnessed significant growth in the past decade, driven by improvements in life expectancy. In 2020, the number of individuals aged 60 and older exceeded 1 billion, surpassing the population of children under 5 years. Projections indicate further increases to 1.4 billion by 2030 and 2.1 billion by 2050 (WHO 2022).

According to the studies by Delbosc and Currie (2011) and Nordbakke and Schwanen (2014) , the inclusion of disabled people in society and assistance in achieving a more normal daily life contribute to increased independence, enabling them to respond to their own needs. This, in turn, leads to improved mental health and an enhanced overall quality of life. One of the significant barriers to active participation in society for individuals with disabilities is their daily transportation and mobility needs (Luiu and Tight 2021). By addressing these challenges, individuals with disabilities can experience increased independence and confidence, reducing their reliance on family members or external assistance (Paparone 2013). This not only benefits the individuals themselves but also makes a positive impact on society as a whole. This facilitation of daily trips and improved mobility can take various forms, depending on individuals' abilities and the infrastructure available in different regions. Solutions range from fully automated vehicles to motorized bikes or scooters, which are widely utilized in high-income countries such as the United Kingdom (UK). In the UK, with a disabled population of 16 million in 2022 and an ageing population exceeding 11 million individuals aged over 65 in 2023 (projected to increase by 32% by 2043), mobility facilitators, including motorized mobility scooters, have gained popularity among citizens and policymakers (Kirk-Wade 2023).

Motorized mobility scooter, commonly referred to as mobility scooter, is one of the most contemporary vehicles. These three or four-wheeled electric vehicles are designed to facilitate the daily trips of individuals with disabilities (see Figure 1). First introduced in 1950 to address the mobility needs of disabled and older consumers, as documented by Krahelski *et al.* (2022), mobility scooters in the UK are categorized into two classes: class 2 and class 3. Class 2 mobility scooters are intended solely for use on pavements, with a maximum speed limit of 4 miles per hour (6.4 kilometers per hour), adhering to legal pavement usage limits. In contrast, class 3 mobility scooters have the flexibility to operate on both roads and pavements. They are governed by a speed limit of 8 miles per hour (12.9 kilometers per hour) on roads and 4 miles per hour on pavements. Additionally, class 3 scooters are required to be equipped with lights and indicators, ensuring enhanced safety during operation (GOV.UK 2023).



Figure 1: Three and four wheeled mobility scooters (Peters et al. 2022)

According to Bækgaard (2017), there were between 300,000 and 350,000 mobility scooters in the UK in 2017, representing the latest available statistics for these vehicles. This marked a significant increase from 90,000 in 2005, and approximately 80,000 new users were expected to join annually in the future. This substantial growth demonstrates the increasing popularity of these vehicles among the UK population. However, according to RICA (2014) safety challenges are also highlighted, with 21% of mobility scooter users expressing a feeling of being unsafe due to experienced crashes, a notably high percentage for a specific type of vehicle.

Experiencing a crash, whether with or without serious injuries, poses a significant issue for every type of vehicle users and experts and policymakers strive to prevent crashes or reduce their severity. Mobility scooters, designed to enhance users' mobility and improve their quality of life (Johnson *et al.* 2013), become crucial in this context. Nevertheless, crashes and additional injuries not only create immediate problems but also make daily life more challenging, lowering the overall quality of life. Therefore, investigating contributing factors to the severity of crashes involving mobility scooter users becomes crucial in ensuring their safety and well-being, highlighting the importance of understanding these factors in the context of modern transportation mode and a supportive service for disabled individuals.

Despite the mentioned yearly growth in the number of users and the importance of reducing the frequency and severity of mobility scooter crashes, there have been very few studies in this field. To the best of the author's knowledge, no study has explored the crash severity of mobility scooters from a transportation safety perspective with the aim of revealing effective factors and devising relevant countermeasures.

To address existing research gaps, this study assesses the factors influencing crash severity among mobility scooter users using five years of data (2018-2022) from the UK roads. A novel hybrid method is employed, combining a machine learning-based technique—the eXtreme Gradient Boosting (XGBoost) algorithm—with the random parameter binary logit model accounting for heterogeneity in means and variances (RPBL-HMV). The XGBoost technique, which has gained popularity recently and is one of the best-performing machine learning methods in crash investigation research, serves as a base model for feature selection for the RPBL-HMV model. XGBoost, an improved machine learning method from boosting techniques, has shown good performance compared to other machine learning techniques in recent literature (Lin *et al.* 2020, Zhou *et al.* 2020). Furthermore, the RPBL-HMV, a subset of random parameters logit with heterogeneity in means and variances for binary outcome models, is known for its superior ability to handle dataset heterogeneity compared to other statistical models in the literature (Behnood and Mannering 2017). Moreover, temporal instability tests have been conducted to investigate the transferability between three important time periods in the dataset: pre-COVID, during COVID, and post-COVID.

The subsequent sections of this paper will include an in-depth review of existing literature. Additionally, the study's methodology will be provided and described. This will be succeeded by a detailed overview of the data utilized, emphasizing the variables incorporated into the model. Following the temporal instability assessment, the estimation results from the models will then be presented, complemented by a thorough analysis of the findings, and the paper will discuss the outcomes obtained from the model estimations. To conclude, inferences will be drawn based on the analysis findings, summarizing the key implications.

1. **Literature Review**

Despite the increasing popularity of mobility scooters in everyday use, there is a scarcity of studies focusing on mobility scooters, particularly from a transportation perspective. Johnson *et al.* (2013) conducted in-depth interviews with key stakeholders in Australia to identify the advantages and challenges of mobility scooters as a growing mode of transportation. Moreover, Toosizadeh *et al.* (2014) conducted a literature review to identify factors enhancing mobility scooter rider skills and performance, proposing riders' training, infrastructure evaluation, and the use of assistive tools as three key approaches to achieving this goal. In another study, Isaacson and Barkay (2020) explored gaps in the mobility scooter literature, highlighting three areas that require examination: the impact of mobility scooter usage on older individuals, the relationship between mobility scooter usage and urban environments, and safety issues associated with mobility scooters. Furthermore, Peters *et al.* (2022) utilized 9-year hospital data from the Netherlands to investigate mobility scooter-related mechanisms, severity, and injury localization. Due to the high rates of admission and surgery, the study emphasized the significance of mobility scooter-related injuries, identifying this population as particularly vulnerable and in a similar vein, Krahelski *et al.* (2022) investigated mobility scooter-associated injuries among 1504 hospital patients in the UK. The study underscored the considerable issue of mobility scooter-related injuries, with collisions with vehicles being the most common and severe crash mechanism, and limbs identified as the most common site of injury. As can be seen, existing research has predominantly explored the medical perspectives of mobility scooter crashes, highlighting their importance for various reasons. However, investigations from a transportation viewpoint are lacking, which is crucial for informing policymakers and implementing necessary measures to mitigate the severity of mobility scooter crashes.

The identification of crash-contributing factors and the efforts to mitigate their impact on road users crash severity have been pivotal topics in road safety literature. The modeling of crashes has undergone significant advancements, yet there is ample room for further progress.

In the early studies, primary statistical models utilizing limited available datasets, as demonstrated by Shankar and Mannering (1996), employed a multinomial logit model, which was considered a promising approach at the time. Subsequent models, such as ordered logit/probit (Khattak *et al.* 1998), and nested logit (Chang and Mannering 1998), gained widespread usage in the literature.

Recognizing the need to enhance model performance in handling unobserved heterogeneity—variabilities and complexities not reflected by the available variables in the dataset—researchers turned to more advanced statistical models and latent class models (Park and Lord 2009) and random parameters (Milton *et al.* 2008) were introduced to mitigate estimated parameter biases and improve overall efficiency. These models have served as common tools for estimating factors associated with crash severity until studies by Seraneeprakarn *et al.* (2017) and Behnood and Mannering (2017) highlighted the superior performance of a random parameter model which accounts for heterogeneity in the means and variances of random parameters. However, these models often rely on restrictive distributional assumptions, making predictions challenging due to the complexity of the models and the estimation of observation-specific parameters (Mannering *et al.* 2020).

As statistical models faced challenges in making accurate predictions, especially with the increasing detail and complexity of datasets, data-driven methods gained popularity among researchers. These methods demonstrated greater capability in modeling high-dimensional datasets and provided improved options for prediction. For example, Chong *et al.* (2005) employed four machine learning algorithms—Artificial Neural Networks (ANN), Decision Trees, Support Vector Machines, and Hybrid Decision Tree-ANN—to predict driver severity in United States road crash data from 1995 to 2000. The study reported that, for lower levels of severity such as 'No Injury' or 'Possible Injuries', decision trees exhibited better performance, while for higher levels of severity, a hybrid decision tree-ANN outperformed other methods. In recent years, numerous studies have explored different data-driven models. For example, Lin *et al.* (2020) conducted a comparison between random forest and XGBoost, with the latter demonstrating superior predictive performance. This finding aligns with Zhou *et al.* (2020) research, where XGBoost outperformed multinomial logistic regression, Naive Bayes, Classification and Regression Tree, and support vector machine. The consensus on XGBoost's effectiveness has prompted other researchers to adopt it as the primary method for analyzing various road user crashes (Parsa *et al.* 2020, Yang *et al.* 2021). Furthermore, Goswamy *et al.* (2023) employed XGBoost to identify contributing factors in pedestrian crash severity, highlighting its superior performance in injury severity prediction compared to a random parameter discrete outcome model. The ongoing comparison between these two methods presents researchers with a dilemma—choosing between the robust identification and interpretation of contributing factors in statistical models or the strong prediction performance of data-driven methods (Mannering *et al.* 2020, Ali *et al.* 2024).

As Mannering *et al.* (2020) suggests, a promising approach involves adopting a hybrid method that combines both data-driven and statistical methods. Sun *et al.* (2022) implemented such an approach by merging a random parameter logit model, which identifies key factors and explores potential heterogeneity, with a Bayesian network to examine statistical relationships between injury severity and significant factors obtained from the random parameter model.

In another study, Sun *et al.* (2023) employed a hybrid approach using a random forest and a random parameters logit model to analyze crashes involving vulnerable road users in Shenyang, China. The study reported consistency in results between the two methods but noted suboptimal performance of the random forest, potentially attributed to a small sample size and the importance of future research to expand the scope of hybrid methods for improved modeling and performance was highlighted. The usage of hybrid methods has seen a significant rise in recent literature as they leverage the predictive ability of machine learning models in various ways and combine this advantage with the strength of statistical models (Samerei and Aghabayk 2024). This approach enables researchers to investigate the effect of each variable on outcome severity more effectively and interpret the models with greater clarity (Ali *et al.* 2021, Azmeri Khan *et al.* 2023).

Building upon these insights, our study adopts a hybrid method combining XGBoost, known for its effectiveness in modeling small datasets, with Random Parameter Binary Logit with heterogeneity in means and variances, recognized as one of the best statistical models. This hybrid approach aims to model the available dataset, assess their integrated performance, and identify factors influencing mobility scooter crashes severity.

1. **Data**

A five-year dataset of UK road crashes from 2018 to 2022 was obtained from the UK Department for Transport DfT (2023) which include three distinct datasets: Collision, Vehicle, and Casualty. These datasets were merged using unique identifier variables named "accident reference" and "accident index". In the next stage, the data was filtered to create a dataset exclusively focused on crashes involving mobility scooter users as casualties. This refined dataset provides detailed information about the involved participants, including sex, age, and residential area type. Furthermore, the dataset encompassed temporal attributes, such as season, day of the week, and time of day, as well as environmental attributes, comprising light and weather conditions, junction type, and junction control measures. Additionally, vehicle-related attributes such as vehicle location and type, maneuvers executed before the crash, and the initial point of impact were included for comprehensive analysis.

The original data categorizes the crash severity into three levels: (i) slight injury, (ii) serious injury and (iii) fatality. The general dataset from 2018 to 2022 recorded a total of 786 slight injuries, 250 serious injuries, and 44 fatal injuries. However, due to the limited number of observations at the higher severity levels, crashes resulting in severe or fatal outcomes were combined into a single category termed as severe injury.

Table 1 provides a summary description of the dataset variables. Notably, the dataset recorded only 111 crashes where a mobility scooter rider was involved in a singular crash, resulting in injury, and without any interaction with other vehicles. A specific variable has created for this particular category of crashes. The limited number of observations in this category could be attributed to the potential under-reporting of crashes with no visible injury. The underreporting of crashes, particularly in micro-mobility crash datasets, is a widespread issue (Schumann *et al.* 2023). Individuals may refrain from reporting such crashes to law enforcement or other data collection agencies, especially when there are no visible or severe injuries.

Moreover, the dataset includes valuable insights into the attributes of the second vehicle involved in each crash, and these attributes were duly examined.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table1** | | | | | | |
| Descriptive statistics: Injury severity level distribution of mobility scooter crash data for variables. (Percentage in parenthesis) | | | | | | |
| **Variables** | **Slight Injury** | | **Severe Injury** | | **Total (Percentage of total crashes)** | |
| **Mobility Scooter rider characteristics** | | | | | | |
| *Sex* | | | | | | |
| Male | 478 | (72.5%) | 181 | (27.5%) | 659 | **(61.0%)** |
| Female | 308 | (73.2%) | 113 | (26.8%) | 421 | **(39.0%)** |
| *Age* | | | | | | |
| Young adults (Age ≤ 25) | 48 | (88.9%) | 6 | (11.1%) | 54 | **(5.0%)** |
| Adults (25 < Age ≤ 40) | 71 | (86.6%) | 11 | (13.4%) | 82 | **(7.6%)** |
| Middle aged adults (40 < Age ≤65) | 188 | (77.4%) | 55 | (22.6%) | 243 | **(22.5%)** |
| Elder (65 < Age ≤80) | 307 | (72.6%) | 116 | (27.4%) | 423 | **(39.2%)** |
| Seniors (80 < Age) | 172 | (61.9%) | 106 | (38.1%) | 278 | **(25.7%)** |
| *Residential area type* | | | | | | |
| Urban | 694 | (73.8%) | 247 | (26.2%) | 941 | **(87.1%)** |
| Rural | 92 | (66.2%) | 47 | (33.8%) | 139 | **(12.9%)** |
| *Deprivation status* | | | | | | |
| Deprived rider  (Individual IMD\* decile ranked between 1 to 5) | 551 | (74.3%) | 191 | (25.7%) | 742 | **(68.7%)** |
| Non-Deprived rider  (Individual IMD decile ranked between 6 to 10) | 235 | (69.5%) | 103 | (30.5%) | 338 | **(31.3%)** |
| *Maneuver* | | | | | | |
| Going straight | 615 | (73.6%) | 221 | (26.4%) | 836 | **(77.4%)** |
| Reversing | 11 | (100.0%) | 0 | (0.0%) | 11 | **(1.0%)** |
| Turning | 33 | (66.0%) | 17 | (34.0%) | 50 | **(4.6%)** |
| Overtaking | 7 | (77.8%) | 2 | (22.2%) | 9 | **(0.9%)** |
| Moving off | 120 | (69.0%) | 54 | (31.0%) | 174 | **(16.1%)** |
| *Skidded/Overturned* | | | | | | |
| Yes | 131 | (64.2%) | 73 | (35.8%) | 204 | **(18.9%)** |
| No | 655 | (74.8%) | 221 | (25.2%) | 876 | **(81.1%)** |
| *First point of impact* | | | | | | |
| No impact | 54 | (67.5%) | 26 | (32.5%) | 80 | **(7.4%)** |
| Front impact | 293 | (76.3%) | 91 | (23.7%) | 384 | **(35.6%)** |
| Side impact | 391 | (71.6%) | 155 | (28.4%) | 546 | **(50.5%)** |
| Back impact | 48 | (68.6%) | 22 | (31.4%) | 70 | **(6.5%)** |
| **Crash characteristics** | | | | | | |
| *Junction control* | | | | | | |
| Traffic signal | 77 | (64.7%) | 42 | (35.3%) | 119 | **(11.0%)** |
| Give-way or uncontrolled | 389 | (75.2%) | 128 | (24.8%) | 517 | **(47.9%)** |
| Other | 12 | (75.0%) | 4 | (25.0%) | 16 | **(1.5%)** |
| Not at Junction | 308 | (72.0%) | 120 | (28.0%) | 428 | **(39.6%)** |
| *Mobility Scooter on footway* | | | | | | |
| Yes | 144 | (78.3%) | 40 | (21.7%) | 184 | **(17.0%)** |
| No | 642 | (71.7%) | 254 | (28.3%) | 896 | **(83.0%)** |
| *Crash type* | | | | | | |
| Multi-vehicle crash | 724 | (74.9%) | 243 | (25.1%) | 969 | **(89.5%)** |
| Single-vehicle crash | 62 | (54.9%) | 51 | (45.1%) | 111 | **(10.5%)** |
| *Light condition* | | | | | | |
| Daylight | 684 | (72.7%) | 257 | (27.3%) | 941 | **(87.1%)** |
| Illuminated darkness | 78 | (70.9%) | 32 | (29.1%) | 110 | **(10.2%)** |
| Darkness | 24 | (82.8%) | 5 | (17.2%) | 29 | **(2.7%)** |
| *Weather condition* | | | | | | |
| Clear weather | 664 | (71.6%) | 263 | (28.4%) | 927 | **(85.8%)** |
| Inclement weather (rain and snow) | 122 | (79.7%) | 31 | (20.3%) | 153 | **(14.2%)** |
| **Other-party vehicle characteristics** | | | | | | |
| *Driver's sex* | | | | | | |
| Male | 473 | (71.5%) | 189 | (28.5%) | 662 | **(68.3%)** |
| Female | 233 | (75.9%) | 74 | (24.1%) | 307 | **(31.7%)** |
| *Driver's age* | | | | | | |
| Young (16 < Age ≤ 30) | 100 | (65.4%) | 53 | (34.6%) | 153 | **(15.8%)** |
| Adults (30 < Age ≤ 60) | 496 | (74.8%) | 167 | (25.2%) | 663 | **(68.4%)** |
| Elder (60 < Age) | 107 | (69.9%) | 46 | (30.1%) | 153 | **(15.8%)** |
| *Driver's residential area type* | | | | | | |
| Urban | 618 | (73.1%) | 228 | (26.9%) | 846 | **(87.3%)** |
| Rural | 86 | (69.9%) | 37 | (30.1%) | 123 | **(12.7%)** |
| *Driver's deprivation status* | | | | | | |
| Deprived driver  (Individual IMD decile ranked between 1 to 5) | 294 | (71.9%) | 115 | (28.1%) | 409 | **(42.2%)** |
| Non-Deprived driver  (Individual IMD decile ranked between 6 to 10) | 411 | (73.3%) | 149 | (26.7%) | 560 | **(57.8%)** |
| *Vehicle type* | | | | | | |
| Bicycle | 8 | (80.0%) | 2 | (20.0%) | 10 | **(1.0%)** |
| Motorcycle | 87 | (58.6%) | 65 | (41.4%) | 152 | **(15.7%)** |
| Passenger car | 486 | (75.5%) | 157 | (24.5%) | 643 | **(66.3%)** |
| Van | 87 | (81.3%) | 20 | (18.7%) | 107 | **(11.1%)** |
| Heavy duty vehicles | 35 | (61.4%) | 22 | (38.6%) | 57 | **(5.9%)** |
| *Maneuver* | | | | | | |
| Going straight | 320 | (66.4%) | 162 | (33.6%) | 482 | **(49.7%)** |
| Reversing | 78 | (84.8%) | 14 | (15.2%) | 92 | **(9.5%)** |
| Turning | 153 | (79.3%) | 40 | (20.7%) | 193 | **(19.9%)** |
| Overtaking | 16 | (59.3%) | 11 | (40.7%) | 27 | **(2.8%)** |
| Moving off | 145 | (82.9%) | 30 | (17.1%) | 175 | **(18.1%)** |
| **Temporal characteristics** | | | | | | |
| *Time of day* | | | | | | |
| Morning rush hour (7-10 A.M) | 92 | (73.0%) | 34 | (27.0%) | 126 | **(11.7%)** |
| Afternoon rush hour (16-19 P.M) | 143 | (78.1%) | 40 | (21.9%) | 183 | **(16.9%)** |
| No rush hour | 551 | (71.5%) | 220 | (28.5%) | 771 | **(71.4%)** |
| *Day of week* | | | | | | |
| Weekend | 168 | (69.1%) | 75 | (30.9%) | 243 | **(22.5%)** |
| Weekday | 618 | (73.8%) | 219 | (26.2%) | 837 | **(77.5%)** |
| *Season* | | | | | | |
| Autumn | 237 | (77.4%) | 69 | (22.6%) | 306 | **(28.4%)** |
| Winter | 151 | (65.7%) | 79 | (34.3%) | 230 | **(21.2%)** |
| Spring | 176 | (73.6%) | 63 | (26.4%) | 239 | **(22.1%)** |
| Summer | 222 | (72.9%) | 83 | (27.1%) | 305 | **(28.3%)** |
| **\*** **IMD: Index of Multiple Deprivation (**measures relative deprivation in small areas across the UK, ranking them from most to least deprived based on factors like income, employment, and education (Scarano *et al.* 2023).) | | | | | | |

1. **Methodological Approach: Machine Learning and Statistical Methods**

In contemporary crash investigations, a wealth of factors is systematically collected and reported, resulting in wide-ranging, detailed, and consequently high-dimensional datasets. The expansive nature of these datasets enhances model predictive performance, contributing to more accurate predictions. Data-driven methods, adept at handling large volumes of data, have become increasingly popular compared to statistical models for modeling such datasets (Mannering *et al.* 2020).

This advantage of having numerous independent variables poses challenges for statistical methods, which struggle to handle datasets with numerous explanatory variables, complicating the process (Scarano *et al.* 2023). Conversely, data-driven methods, often labeled as 'black box,' pose challenges in interpreting outcomes. This challenge in interpretation holds particular significance in road safety research, where the goal is to identify the impact of each crucial variable on outcomes and propose safety countermeasures.

To reconcile these challenges, a hybrid approach is employed. Initially, a data-driven tool, specifically the eXtreme Gradient Boosting (XGBoost) machine learning algorithm, is used along with the Shapley Additive exPlanations (SHAP) technique for outcome interpretation. XGBoost has demonstrated good performance for datasets with a limited number of observations and assists in identifying important crash factors while reducing the dataset dimensions (Zhou *et al.* 2020, Chang *et al.* 2022).

Subsequently, the refined dataset is employed in an advanced statistical model named the random parameter binary logit with heterogeneity in means and variances (Behnood and Mannering 2017, Seraneeprakarn *et al.* 2017). This study hybrid approach, depicted schematically in Figure 2, aims to capitalize on the strengths of each method while addressing their respective challenges.

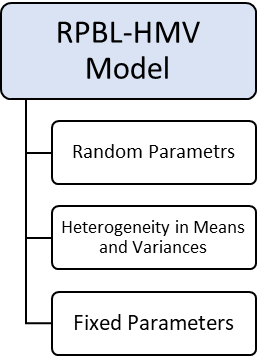
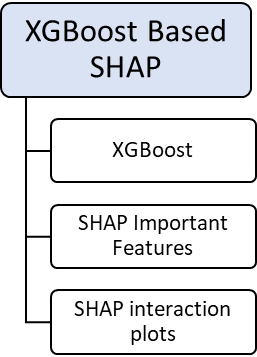
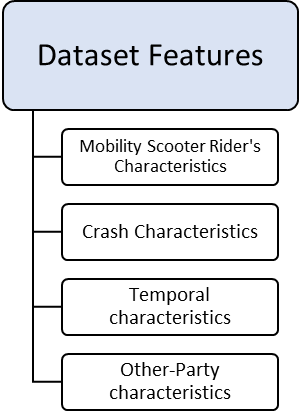


Figure 2: Schematic Process of Modeling

***4.1 XGBoost-based SHAP***

XGBoost, a tree-based machine learning technique, belongs to the boosting class of ensemble learning methods, specifically designed to enhance basic decision trees by combining weak classifiers to improve the final prediction (Chen and Guestrin 2016). Decision trees operate as single models that predict the final outcome based on all available data. When integrated with bagging approaches, decision trees produce multiple classifiers by training them with randomly selected training data. This process aggregates predictions from all trained classifiers and is known as a different method in machine learning models, termed Random Forest. In the final stage of model evaluation, using Random Forests with boosting methods results in gradient boosting methods, of which XGBoost is a subset, offering improved performance in model speed and prediction ability (Yuan *et al.* 2022).

When predicting a target variable Yi, these models utilize a training dataset containing various explanatory variables Xi. Central to their evaluation is an objective function (), essential for assessing the model's goodness of fit on the training dataset. This function comprises two key terms: the training loss function and the regularization term, as outlined below (Chen and Guestrin 2016):

(1)

In this context, denotes the training loss function, which assesses model's predictive performance on the training set where represents the actual label, with *i* representing the *i*-th sample in the dataset. Furthermore, stands for the predicted values given by the formula:

(2)

Here t denotes the number of iterations, K represents number of additive functions and each corresponds to an independent tree structure q and leaf weights :

(3)

Unlike decision trees, each tree in this context contains a continuous score on each leaf, represented by . Returning to equation (1), represents the regularization term governing the model's complexity, formulated as:

(4)

Where is the number of leaf nodes, j is the index of each leaf node, and and λ represent the regularization parameters.

To optimize the model-building process, it is crucial to fine-tune the XGBoost hyperparameters. This was accomplished through the grid search technique (Yang and Shami 2020), which systematically tests a predefined range of hyperparameters to identify the combination that yields the best performance based on cross-validation results. Specifically, we tuned parameters such as the learning rate, maximum depth, number of estimators, and regularization parameters. The grid search method ensures a thorough exploration of the hyperparameter space, allowing us to select the most effective model configuration for our dataset. The results of the grid search process, including the primary hyperparameter values used, are presented in Table 2.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table2** | | | |
| Results of Grid search in model optimization process | | | |
| Primary Hyperparameters | Models | | |
| Pre-COVID | COVID | Post-COVID |
| Number of estimators | 150 | 250 | 100 |
| Learning rate | 0.01 | 0.10 | 0.10 |
| Maximum depth | 3 | 3 | 4 |

Moreover, to enhance generalization and mitigate the risks of overfitting, a 10-fold cross-validation is implemented, coupled with an 80-20 split for training and testing data classification (Ali *et al.* 2024). Furthermore, the interpretation of model outcomes involves visualizing SHAP values (Lundberg and Lee 2017), which provides a comprehensive global influence and importance ranking diagram for the impact of each variable on the severity of traffic crashes, along with interaction plots between some selected variables.

The Python software (version 3.10) was employed for the analysis, utilizing various packages such as XGBoost and SHAP which were instrumental in achieving the research objectives.

* 1. ***Random parameter*** ***binary logit with heterogeneity in means and variances***

In the preceding stage, significant categorical variables were analyzed using the XGBoost technique. These variables were subsequently transformed into dummy variables to enhance their applicability in the random parameters approach, facilitating the assessment of their impacts. Within the framework of random parameters, the severity function (Y) is defined as binary to determine the severity level (L = 0: slight injury, L =1: severe injury) corresponding to each individual i:

(5)

The vector XLi comprises explanatory variables, where βL is a vector containing estimable parameters associated with discrete outcome L, and the error term is denoted as εLi. In Equation 5, the consideration of variability in means and variances is achieved by allowing βL to vary among observations. To estimate βL in this context, Equation 6 is employed (Seraneeprakarn *et al.* 2017, Washington *et al.* 2020).

(6)

In this context, β denotes the mean parameter estimate observed across all observations. ZLi is a vector comprising explanatory variables that account for variations in means, while θLi represents the corresponding vector of estimable parameters, WLi is a vector of explanatory variables capturing variations in the standard deviation σLi using the parameter vector ωLi, and νLi denotes the disturbance term.

The probability of injury severity category L, attributed to observation i, is expressed as:

(7)

Here, the density function of β is denoted by f(β|φ), with the other parameters having been previously discussed and due to binary nature of outcome severity in this study can be written as follow:

(8)

In this study, the simulated maximum likelihood technique, employing 1000 Halton draws, is utilized to estimate the models (Train 2009), taking into consideration both efficiency and accuracy. During the model estimation process, a normal distribution is chosen based on its demonstrated appropriateness in comparison to other distributions (Waseem *et al.* 2019). Furthermore, a thorough examination of all significant explanatory variables is conducted to ascertain their impact on the means and variances of the random parameters. Lastly, to interpret the influences of significant variables on the injury severity outcomes of mobility scooter users, marginal effects are computed, providing insights into the impact of these variables on the outcome severity (Washington *et al.* 2020).

1. **Temporal Instability**

In recent safety literature, there has been widespread emphasis on the potential for model outcomes to vary across different time periods, necessitating temporal instability tests to assess the transferability of results. Additionally, drawing from behavioral research fields such as psychology, economics, and neural and cognitive sciences, Mannering (2018) suggested that this instability is likely a result of various aspects of individual decision-making processes. Such processes are important and repetitive during driving, as the task involves continual trade-offs between different utilities influenced by the driver's experiences and information gathered from the environment, including other drivers, road conditions, and social interactions.

Given the potential impact of COVID-19 on societies and individuals' driving behaviors (Dong *et al.* 2022), it is important to investigate the possible instability in crash severities before, during, and after the pandemic. Three distinct groups representing important time periods of equal length, specifically 17 months, were extracted from the five-year dataset to check the possible impact of COVID-19 on mobility scooter crashes. These periods are pre-COVID, during COVID, and post-COVID. The pre-COVID period includes crashes from October 1, 2018, to February 29, 2020, before the declaration of coronavirus outbreak by the World Health Organization and the onset of COVID-related social restrictions in the UK (Sherrington 2022). The during-COVID period covers crashes from March 1, 2020, to July 31, 2021, coinciding with significant vaccination efforts and the easing of restrictions in the last months (IFG 2022). Finally, the post-COVID period comprises crashes from August 1, 2021, to December 31, 2022, following the relaxation of most legal limits on social contacts and the reopening of closed sectors of the economy.

This separation of data resulted in approximately balanced distinct datasets, with 285, 252, and 395 crashes, respectively, for the pre-COVID, during-COVID, and post-COVID periods. This setup enables the examination of the potential effect of COVID-19 on road crash severity in the UK. To assess the transferability across different periods, likelihood ratio tests were employed to examine the null hypothesis that models across different mentioned periods are similar and separate specifications are not needed. Equation 9 was used to conduct these tests as follows (Washington *et al.* 2020):

(9)

Where is the log-likelihood at convergence of the model parameters derived from period P2 data while using period P1 data, and is the log-likelihood at convergence of the model based on period P1 data. The values of P1 and P2 refer to different aforementioned periods of time. To conduct comparisons across periods, these tests were also conducted by replacing P1 and P2 values with each other. The resulting has a chi-squared distribution with a degree of freedom equal to the number of estimated parameters and are represented in Table 3. The table shows that the null hypothesis is rejected at the 99% confidence interval.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 3** | | | |
| Likelihood ratio test results between different periods. Confidence levels in the bracket with degrees of freedom in parenthesis. | | | |
|  |
| **P1** | **P2** | | |  |
| Pre-COVID | COVID | Post-COVID |  |
| Pre-COVID | - | 100.2 (12) | 92.7 (12) |  |
| [ at 99%] | [at 99%] |  |
| COVID | 77.3 (10) | - | 66.6 (12) |  |
| [at 99%] | [at 99%] |  |
| Post-COVID | 35.1 (10) | 41.8 (12) | - |  |
| [at 99%] | [at 99%] |  |

As this study investigates mobility scooter crash severity for the first time in the transportation era, the upcoming results section places greater emphasis on shedding light on specific influential variables.

1. **Results**

***6.1 Important factors based on XGBoost-SHAP***

The XGBoost-SHAP results for each period model are visually presented in Figure 3-5 (and for the 2018-2022 model, it is presented in appendix figure A.1), depicting the 20 most important factors according to the XGBoost model, ranked in order of SHAP values. Higher SHAP values indicate greater importance of the variable in influencing mobility scooter crash severity. Additionally, a percentile change line is included, representing the percentage difference between each variable's SHAP values mean. This percentile change line assists in identifying the variables intended for the RPBL-HMV model (Sun *et al.* 2023). The peaks in the percentile line are calculated by dividing the difference of two neighboring SHAP values by the upper factor value.

For instance, in Figure 3, the first significant variation is observed between the Mobility Scooter rider's age and Other-party driver’s age variable. The percentile line exhibits five peaks, indicating extreme increases in percentile line values in the Figure 3. Using this information, the first 15 variables based on five peaks in the percentile line are separated from the dataset and utilized in the RPBL-HMV model to examine their effects on crash severity for the pre-COVID period. These variables include age, sex, deprivation status, first point of impact, and skidded/overturned from Mobility Scooter rider characteristics; and maneuver, first point of impact, driver's age, driver's sex, driver's deprivation status, and vehicle type from second party characteristics. Additionally, junction control is considered from crash characteristics, while season, day of week, and daytime rush hour are incorporated from temporal characteristics for analysis in the RPBL-HMV model for the pre-COVID period. Similarly, in the process of conducting hybrid models on divided datasets for temporal instability tests, a group of variables from the first five peaks in the percentile line are selected for other periods to build the RPBL-HMV model. Based on figures 4 and 5, this procedure results in 16 and 14 variables for the during and post-COVID periods, respectively (and also 13 variables for the 2018-2022 model, presented in the appendix figure A.1).

Variables such as mobility scooter rider's age and first point of impact, other-party vehicle type, vehicle maneuver, first point of impact, driver's age, and deprivation status, as well as season, daytime rush hour, and junction control are mutual across the three periods, indicating their consistent importance over time. It is worth noting that most of the other important variables for each period also have high SHAP values; however, as we only used variables before the fifth peak in SHAP values, they were not transferred into the RPBL-HMV models.

Additionally, this approach of selecting variables based on their percentile change in the XGBoost-SHAP results contributes to a reduction in the dataset dimension. This reduction is advantageous as it helps enhance the framework of the RPBL-HMV. A more focused dataset with the most influential variables allows for a more efficient and targeted analysis within the RPBL-HMV model.

Figure 3: Top 20 Features Based on SHAP Values for Pre-COVID Period

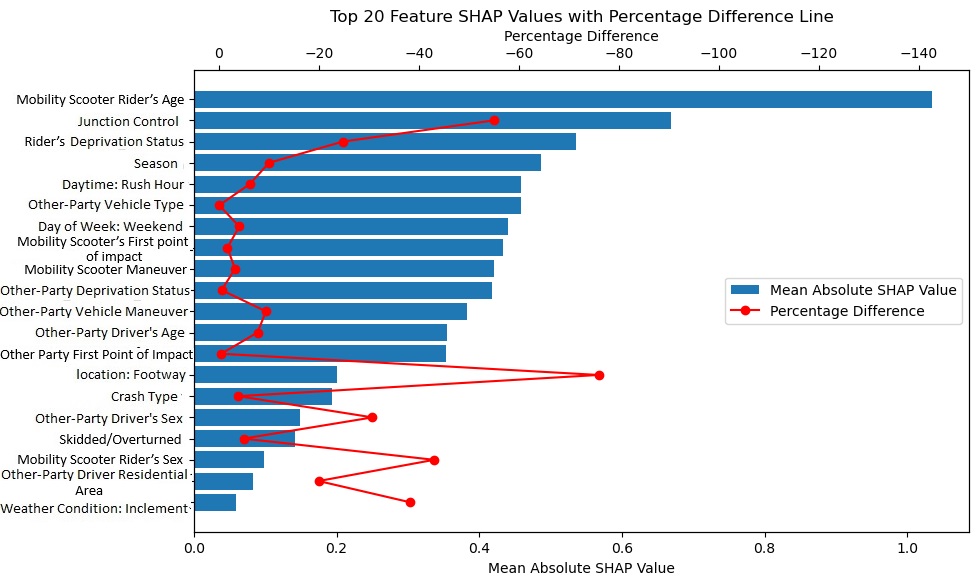
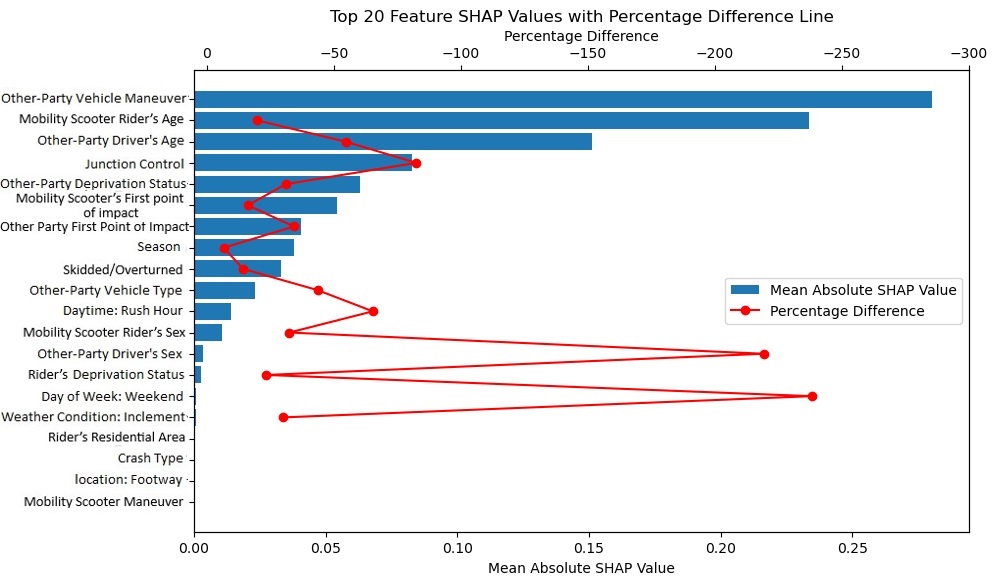


Figure 4: Top 20 Features Based on SHAP Values for COVID Period

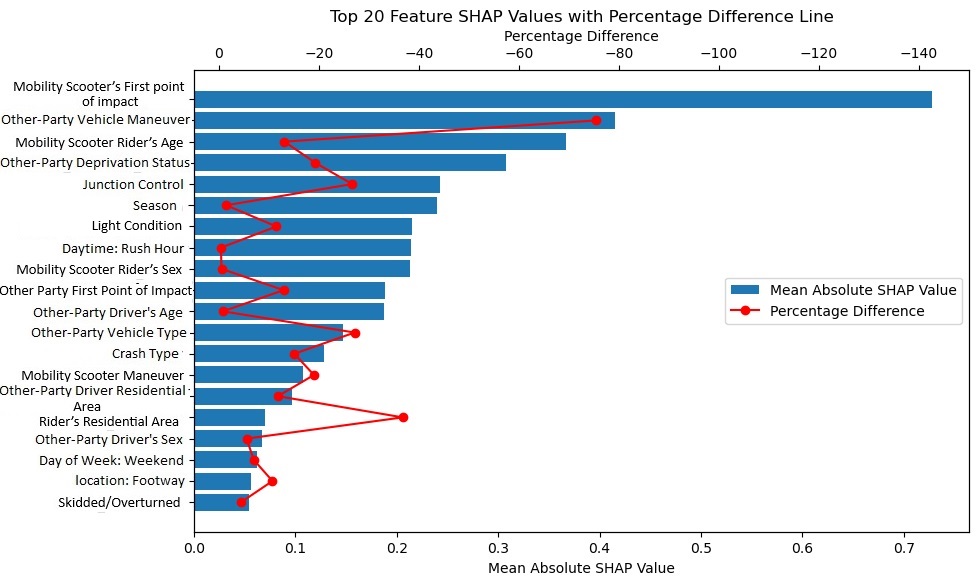


Figure 5: Top 20 Features Based on SHAP Values for Post-COVID Period

Moreover, the goodness-of-fit measures for XGBoost temporal models are presented in Table 4. Although the datasets for each period are relatively small, the models exhibited satisfactory goodness-of-fit measures for the test dataset.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table4** | | | |
| Goodness-of-fit values for temporal models | | | |
| Measures | Models | | |
| Pre-COVID | COVID | Post-COVID |
| Accuracy | 0.61 | 0.80 | 0.64 |
| Precision | 0.61 | 0.68 | 0.60 |
| Recall (Sensitivity) | 0.65 | 0.65 | 0.63 |

***6.2 Random parameter binary logit with heterogeneity in means and variances model outcome***

The selected variables from XGBoost are employed to construct the RPBL-HMV model using mobility scooter crash data to explore their impact and significance on crash severity. Depicted in Table 5, only variables with a significance level of 0.1 or below are retained in the final models, which is a common practice in crash severity literature (Michalaki *et al.* 2015, Se *et al.* 2021, Salehian *et al.* 2023).

While all variables were tested for possible heterogeneity in means and variances, significant heterogeneity was found only in the variance of the 2018-2022 model’s random parameter (see Table A.1 in the appendix). All significant items are presented in Table 5 and Tables A.1-A.4 in the appendix. Furthermore, to examine the effect of each variable and interpret the model, the marginal effects of each explanatory variable have been calculated and are detailed in outcome tables A.1-A.4 in the appendix.

6.2.1 Random parameters with heterogeneity in mean and variance

As illustrated in Table 5, the variable representing give-way or uncontrolled junctions generates a random parameter with a parameter mean of -2.285 and a standard deviation of 3.253 in the post-COVID model. This denotes a decrease in the probability of crashes occurring in these types of junctions resulting in the mobility scooter rider being killed or severely injured for 76% of the individuals (see Figure 6). Similarly, this variable exhibited random effects on outcome severity in the 2018-2022 model (see Table A.1.), reducing severe injuries for 70% of observations. This outcome may be attributed to the tendency of mobility scooter riders, often individuals with various movement difficulties, tend to exercise extra caution while navigating roads, particularly at junctions without traffic lights or other facilities for vulnerable road users. Additionally, other road users, such as car drivers, may exhibit increased attention and caution when approaching junctions with the presence of vulnerable road users such as mobility scooters (Hussain *et al.* 2023).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table5** | | | | | | | | | | | | | | |
| RPBL-HMV Model Results for Mobility Scooter Rider Injury Severity (Z-values are presented in parenthesis and random parameters means and standard deviations are bolded) | | | | | | | | | | | | | | |
| Variable | Parameter estimate | | | | | | | | | |
| Pre-COVID | | | COVID | | | Post-COVID | | | | |
| Constant | -2.104 (-5.58) | | | -4.085 (-4.00) | | | | -1.591 (-5.48) | | | | |
| Adult Rider (1 if the rider was over the age of 25 and under the age of 40; 0 otherwise) | 0.727 (1.83) | | | - | | | | - | | | | |
| Senior Rider (1 if the rider was over the age of 80; 0 otherwise) | 1.359 (3.36) | | | 1.638 (2.45) | | | | 1.247 (3.24) | | | | |
| Deprived Rider (1 if the mobility scooter rider was deprived; 0 otherwise) Mean | - | | | **-2.631 (-1.77)** | | | | - | | | | |
| standard deviation | - | | | **4.266 (2.24)** | | | | - | | | | |
| Mobility Scooter skidded/overturned | 0.680 (1.80) | | | - | | | | - | | | | |
| Mobility Scooter Maneuver = Move Off (1 if the mobility scooter was moving off; 0 otherwise) | - | | | 1.523 (1.96) | | | | - | | | | |
| Rear-end Collision (1 if the first point of impact for mobility scooter was from the back of their vehicle; 0 otherwise) | - | | | - | | | | 1.243 (2.26) | | | | |
| Sideswipe Collision (1 if the first point of impact for mobility scooter was on the side of their vehicle; 0 otherwise) | - | | | 1.776 (2.40) | | | | - | | | | |
| Junction Control = Traffic signal (1 if the junction was controlled by traffic signal; 0 otherwise) | 1.678 (3.67) | | | - | | | | - | | | | |
| Give-way or Uncontrolled Junction (1 if the junction was a Give-way or uncontrolled junction; 0 otherwise) | - | | | 1.179 (1.83) | | | | **-2.285 (-1.72)** | | | | |
| standard deviation of Give-way or Uncontrolled Junction | - | | | - | | | | **3.253 (1.77)** | | | | |
| Footway (1 if the mobility scooter was on the footway; 0 otherwise) | - | | | -2.815 (-2.44) | | | | - | | | | |
| Single Vehicle Crash (1 if only one motorized vehicle was involved in the crash; 0 otherwise) | - | | | 2.969 (3.25) | | | | 1.344 (2.89) | | | | |
| Darkness (1 if there was no lightning at the crash location; 0 otherwise) | - | | | - | | | | -2.187 (-1.86) | | | | |
| Other Party Vehicle = Young Driver (1 if the second vehicle's driver was over the age of 16 and under the age of 30; 0 otherwise) | 1.052 (2.82) | | | - | | | | - | | | | |
| Other Party Vehicle = Elder Driver (1 if the second vehicle's driver was aged over 60; 0 otherwise) | - | | | 2.342 (2.50) | | | | - | | | | |
| Other Party Vehicle = Deprived Driver (1 if the second vehicle's driver was deprived; 0 otherwise) | - | | | - | | | | 0.569 (1.68) | | | | |
| Other Party Vehicle Maneuver = Reversing (1 if the second vehicle was reversing; 0 otherwise) | - | | | - | | | | -1.192 (-1.95) | | | | |
| Other Party Vehicle Maneuver = Turning (1 if the second vehicle was turning; 0 otherwise) | -1.627 (-3.38) | | | -2.127 (-2.01) | | | | - | | | | |
| Other Party Vehicle Maneuver = Move Off (1 if the second vehicle was moving off; 0 otherwise) | -1.434 (-3.11) | | | - | | | | - | | | | |
| Morning Rush Hour (1 if the crash occurred between 7-10 A.M.; 0 otherwise) | 0.994 (2.16) | | | - | | | | - | | | | |
| Weekend (1 if the crash occurred during weekend days; 0 otherwise) | - | | | 1.801 (2.36) | | | | - | | | | |
| Winter (1 if the crash occurred during winter months; 0 otherwise) | 0.544 (1.74) | | | - | | | | 1.401 (3.26) | | | | |
| ***Heterogeneity in the mean of the Random parameters*** | | | | | | | | | | | | | | |
| Give way or Uncontrolled Junction: Other Party Vehicle = Deprived Driver | | - | | | - | | | | -3.494 (-1.76) | | | | |
| Give way or Uncontrolled Junction: Summer | | - | | | - | | | | 2.120 (1.78) | | | | |
| ***Heterogeneity in the variance of the Random parameters*** | | | | | | | | | | | | | | |
| Number of observations | | | 285 | | | 252 | | | | 395 | | | | |
| Log-likelihood at convergence | | | -142.04 | | | -108.10 | | | | -208.27 | | | | |
| McFadden Pseudo R2 *= 1 − (LL(β))/(LL(0))* | | | 0.281 | | | 0.381 | | | | 0.239 | | | | |
| Adjusted R2 *= 1 − (LL(β)−N)/(LL(0))* | | | 0.230 | | | 0.312 | | | | 0.195 | | | | |

N = Number of estimated parameters

Moreover, in the post-COVID model, the variable representing crashes occurring during the summer season, and in the 2018-2022 model, the variable representing motorcycles as the second party in the crash, introduce heterogeneity in the mean of the give-way or uncontrolled junctions parameter. This increases the mean of the random parameter, resulting in a higher probability of the mobility scooter rider being severely injured or killed. Furthermore, the variable indicating the location of the mobility scooter on the footway causes heterogeneity in the variance of the 2018-2022 model random parameter, widening the distribution of the variable. This indicates that crashes at give-way or uncontrolled junctions have a higher degree of randomness when the mobility scooter is on the footway.

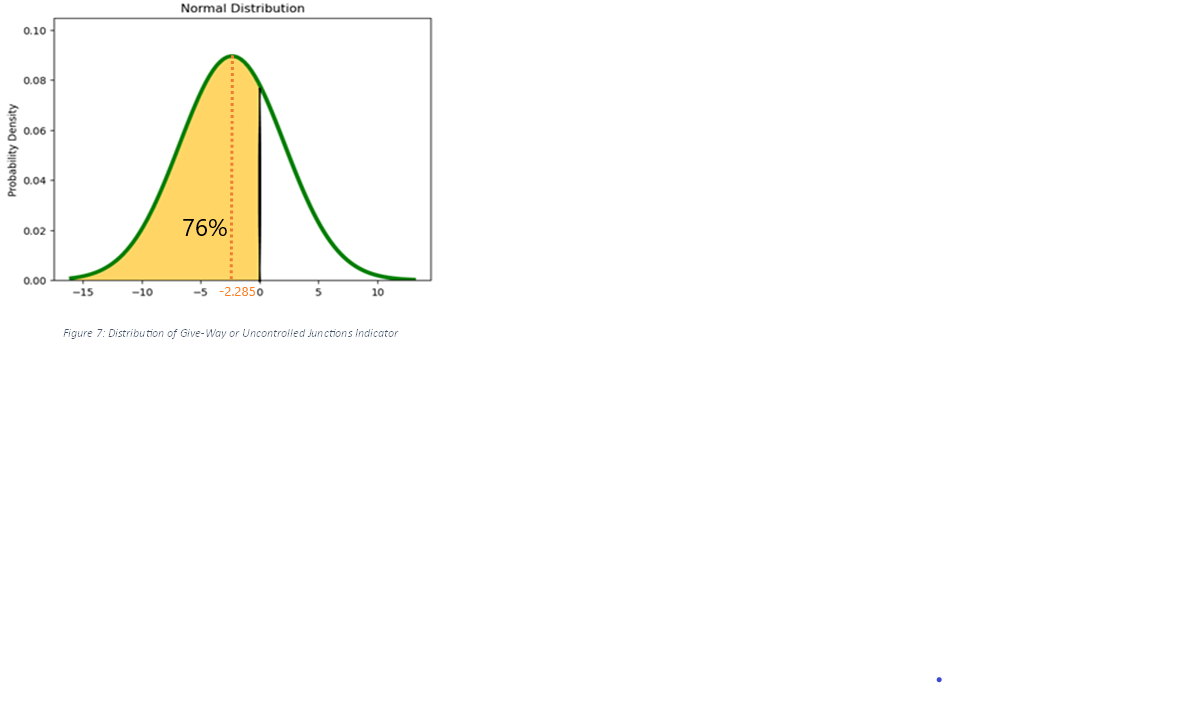
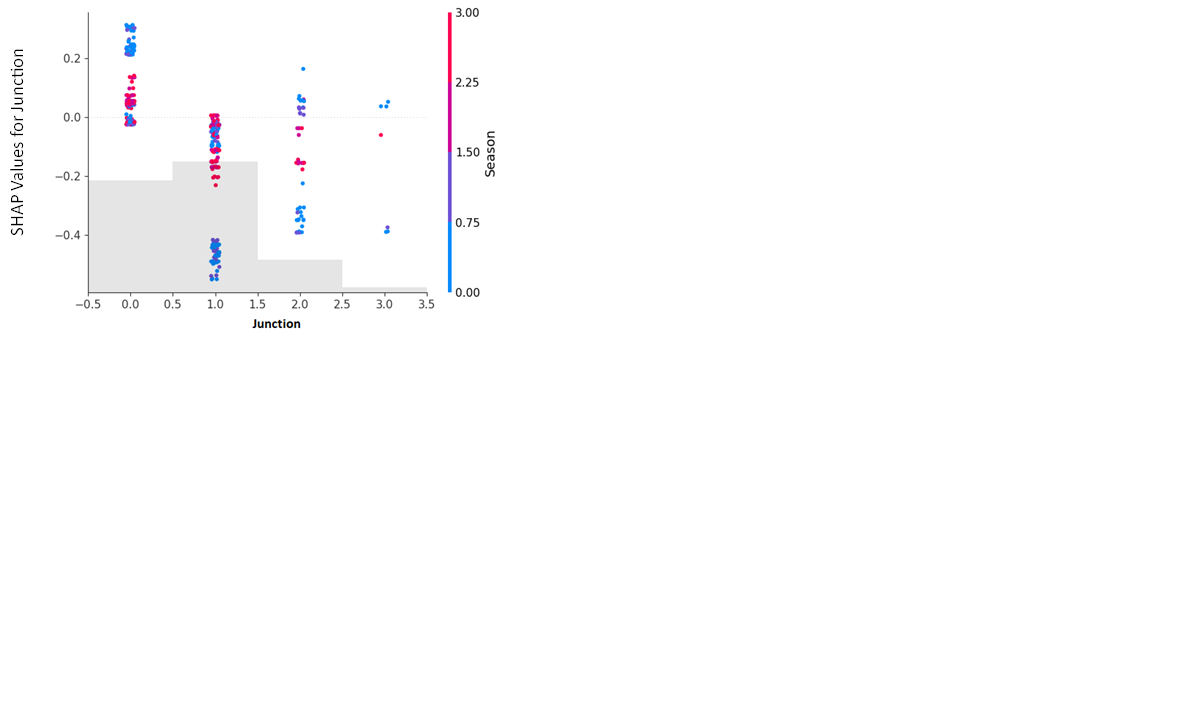


Figure 6: Distribution of Give-Way or Uncontrolled Junctions parameter for Post-COVID model

Figure 7 and Figure A.2 are SHAP interaction plots used to compare the results of the XGBoost model and the RPBL-HMV model. In the post-COVID model, two variables generated heterogeneity in the mean of the random parameter: Season (summer) and Other-Party Deprived Driver. These variables influenced the mean of the variable representing give-way or uncontrolled junctions. The interaction between the random parameter and each of these variables is examined in Figures 7 and A.2 by utilizing the SHAP values of the XGBoost model.

The X axis of plot shows different categories of “Junction type” variable, the left-side of the Y axis is SHAP values of them and their effect on the output of XGBoost model and the right-side of the Y axis which shows different colors of dots, represents different categories of variable Season (Wen et al. 2021). Depicted in Fig7, Give-way or uncontrolled junctions (Junction=1) have negative SHAP values generally, which means that they have general of negative impact on crashes with severe injuries which is approximately equal to results of RPBL model. Moreover, for crashes during summer (Season = 3, shown with red dots), the SHAP values of give-way or uncontrolled junctions are higher, indicating a greater probability of severe crashes compared to autumn and winter seasons (Season = 0 and 1). This seasonal variation in interactions at give-way or uncontrolled junctions, which aligns with the outcomes of the RPBL model, may be attributed to precipitation-induced factors such as reduced visibility and poor road conditions. These factors likely prompt drivers to exercise greater caution (Konstantopoulos et al., 2010).

Figure 7: SHAP Interaction between Junction Type and Season from XGBoost model



**1**

**3**

**0**

**2**

6.2.2. Fixed parameters of RPBL-HMV model

Examining the fixed coefficients of the models, several factors related to the age of mobility scooter riders significantly influenced the severity of crashes. Beginning with the variable representing senior riders (aged over 80 years old) exhibited a direct relation with a higher likelihood of being killed or seriously injured in all models. This finding aligns with the research by Krahelski *et al.* (2022) and can be attributed to various aspects of their condition, such as poor physical condition (Hwang and Hong 2018, Adebisi *et al.* 2019) or higher reaction times (Anstey *et al.* 2012, Ayuso *et al.* 2020). These factors contribute to their reduced ability to prevent collisions and increased vulnerability during collisions. Additionally, as indicated by Oxley *et al.* (2010) injuries to senior individuals can complicate or extend the rehabilitation and hospitalization process, significantly impacting their post-crash quality of life (Peters *et al.* 2022). Moreover, rear-end collisions in post-COVID and 2018-2022 models, as well as sideswipe collisions in during COVID and 2018-2022 models, both impact the severity of mobility scooter users by heightening the probability of severe injuries. This could be attributed to the limited protection for mobility scooter riders in such collisions, affecting their balance and potentially causing falls from the vehicle, resulting in different injuries to vital body parts.

Additionally, crashes occurring during winter months tend to increase the severity of mobility scooter crashes in all models except during COVID model, possibly due to factors such as increased precipitation leading to slippery roads, resulting in more loss of control for mobility scooter riders and other vehicle drivers.

In the COVID period model, several variables emerged as significant that were not found to be significant in other period models. This could be attributed to various factors. For instance, the variable representing other-party elder drivers and crashes occurring on weekends had a positive effect on the probability of severe injuries, increasing their likelihood. During the COVID period, which was marked by different regulations at different times, lockdowns often necessitated individuals to fulfill their needs independently, including elderly individuals. This increased their presence in society and, consequently, their involvement in crashes. Additionally, with remote work becoming prevalent during the pandemic, weekends may have been the only opportunity for individuals to leave their homes. This, coupled with the perception of fewer vehicles on the streets, might have led to a false sense of security and potentially increased instances of distracted or unsafe driving, whether intentional or accidental. These behaviors could have contributed to a higher probability of crashes resulting in severe injuries for mobility scooter users during weekends in the COVID period.

On the other hand, crashes occurring on footways diminished the likelihood of severe injuries or deaths in the during COVID and 2018-2022 models, as these crashes typically involve collision between mobility scooters and other vulnerable road users. In the context of collision with other motorized vehicles, the separation between on-footway users and other road vehicles in these cases usually results in less transfer of impact energy to mobility scooter vehicles and riders. This separation helps mobility scooter users maintain better balance and tolerate fewer strikes compared to collisions with other vehicles on the road. Furthermore, based on Table 5, the outcomes of attributes related to the involvement of other parties, including the effect of the second-party vehicle maneuver, indicate that both variables representing turning and reversing as second vehicle maneuvers decreased the probability of severe injuries to the mobility scooter rider. This can be attributed to the lower speed of the second vehicle during these maneuvers. (Tefft 2013).

1. Policy Implications

Based on the models’ outcomes, there are several interventions that could be implemented to reduce the severity of crashes involving mobility scooter riders. Regarding the age of mobility scooter riders, while mobility scooters are intended to assist individuals with mobility issues, there are fundamental abilities required for users to prevent harm to themselves and others on the road. One approach is to impose stricter regulations regarding the minimum abilities and skills of mobility scooter riders. These regulatory requirements could help prevent individuals with significant impairments from driving on roads, especially in shared spaces with vulnerable road users such as pedestrians. However, it is important to consider alternatives for those who are unable to meet these requirements. For instance, older riders experiencing serious issues with body control, movement, and cognitive function could benefit from advancements in technology. Technologies such as those used in the autonomous car industry could potentially compensate for the lacking abilities of riders, thereby enhancing their driving skills and safety.

Moreover, these findings regarding infrastructure effects on crash severity, such as footways, demonstrates the importance of separating mobility scooters from road vehicles by utilizing footways. In the UK, there are two types of legitimations for mobility scooters, allowing specific kinds of mobility scooters to drive on roads. These findings suggest that mobility scooters on footways, including class 2 mobility scooters, have a lower probability of suffering from severe injuries. This highlights for policymakers the importance of fully separating mobility scooters from vehicles on the road, when possible, to protect them from injuries. Although sharing space between mobility scooters and other footway users, such as pedestrians, may introduce some challenges (Jancey *et al.* 2013), policymakers can address serious concerns through proper regulations (Laverdet *et al.* 2023), rider training to improve skills (Toosizadeh *et al.* 2014), and the provision of adequate and safe infrastructure (Jancey *et al.* 2013). These measures can help create a safer shared space for all vulnerable road users.

1. **Limitations**

The data employed in this study has certain limitations, emphasizing the necessity for more detailed information. Specifically, acknowledging the intoxication status of crash participants and obtaining more information about crash participants, safety tools used by mobility scooter users, such as helmets or specific vehicle abilities, would enhance understanding of their impact on crash severity. Additionally, identifying the exact class of mobility scooters (Class 2 or Class 3) could offer additional insights for further research, given that these classes are subject to different regulations. Unfortunately, such details were unavailable in the current dataset.

Datasets on mobility scooter crashes should become more detailed to enable researchers to investigate these crashes more effectively, examine specific features comprehensively, and ultimately improve the safety of these vehicle users.

1. **Conclusion**

This study identified the factors influencing crash severity among mobility scooter users on United Kingdom roads from 2018 to 2022. The research employed a hybrid methodology that combines a machine learning method, XGBoost based SHAP, with an advanced statistical model known as the random parameter binary logit model, accounting for heterogeneity in means and variances. By integrating these two distinct methods from different model categories—data-driven and statistical—the study leveraged the strengths of each method while mitigating their respective weaknesses. This includes harnessing dimension reduction and predictive capabilities from data-driven methods and gaining interpretability and the ability to account for unobserved heterogeneity from statistical models (Mannering *et al.* 2020).

The XGBoost-based SHAP models were instrumental in identifying the most crucial variables, ranking them in terms of importance. The study extracted the variables based on the difference in SHAP values, which encompass various aspects of crashes and the involved participants. Subsequently, these variables were used to construct the random parameter binary models, accounting for heterogeneity in means and variances.

Furthermore, temporal instability tests were employed, utilizing a likelihood ratio test, to examine the transferability of variables affecting the crash severity involving riders of mobility scooters across three different periods: Pre, during, and Post COVID. The results indicated that the null hypothesis, which suggested the transferability of model outcomes between different periods, was rejected. This finding implies that despite the mutual features, each year's model and estimated parameters were temporally unstable during the study's covered data.

Among these variables, crashes occurring at give-way or uncontrolled junctions generated random parameters in the post-COVID period model, decreasing the probability of crashes resulting in severe injury or fatality for 76% of the population. However, the variable representing crashes occurring during summer months in the post-COVID model caused heterogeneity in the random parameter's mean, thereby increasing the likelihood of more severe crashes. Moreover, fixed variables such as a senior mobility scooter rider (aged over 80), rear-end and sideswipe collisions, crashes during winter months, increased the probability of severe injuries for mobility scooter riders. Conversely, variables such as crashes with mobility scooters on the footway and second vehicle maneuvers such as reversing and turning reduced the likelihood of severe injuries.

In conclusion, the study's outcomes offer valuable insights regarding the significance of different variables in influencing the severity of mobility scooter crashes. The findings emphasize the importance of conducting further studies with more comprehensive crash information, alongside utilizing the efficiency of the applied methodological approaches.

**Declaration of Competing Interests**

The authors declare that they do not have any known competing financial interests or personal relationships that could have influenced the work reported in this paper.

**Funding**

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

**Data availability**

Available online

**References**

Adebisi, A., Ma, J., Masaki, J. , Sobanjo, J. 2019. Age-related differences in motor-vehicle crash severity in california. *Safety* [Online], 5.10.3390/safety5030048

Ali, Y., Haque, M. M., Zheng, Z. , Bliemer, M. C. J., 2021. Stop or go decisions at the onset of yellow light in a connected environment: A hybrid approach of decision tree and panel mixed logit model. Analytic Methods in Accident Research31, 100165. <https://doi.org/https://doi.org/10.1016/j.amar.2021.100165>

Ali, Y., Hussain, F. , Haque, M. M., 2024. Advances, challenges, and future research needs in machine learning-based crash prediction models: A systematic review. Accid Anal Prev194, 107378. <https://doi.org/10.1016/j.aap.2023.107378>

Anstey, K. J., Horswill, M. S., Wood, J. M. , Hatherly, C., 2012. The role of cognitive and visual abilities as predictors in the multifactorial model of driving safety. Accident Analysis & Prevention45, 766-774. <https://doi.org/https://doi.org/10.1016/j.aap.2011.10.006>

Ayuso, M., Sánchez, R. , Santolino, M., 2020. Does longevity impact the severity of traffic crashes? A comparative study of young-older and old-older drivers. Journal of Safety Research73, 37-46. <https://doi.org/https://doi.org/10.1016/j.jsr.2020.02.002>

Azmeri Khan, S., Yasmin, S. , Mazharul Haque, M., 2023. Effects of design consistency measures and roadside hazard types on run-off-road crash severity: Application of random parameters hierarchical ordered probit model. Analytic Methods in Accident Research40, 100300. <https://doi.org/https://doi.org/10.1016/j.amar.2023.100300>

Bækgaard, E. S., 2017. Mobility scooter accidents - need for preventative action? Clinical Medical Reviews and Case Reports4 (2). <https://doi.org/10.23937/2378-3656/1410158>

Behnood, A. , Mannering, F., 2017. Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research16, 35-47. <https://doi.org/10.1016/j.amar.2017.08.001>

Chang, I., Park, H., Hong, E., Lee, J. , Kwon, N., 2022. Predicting effects of built environment on fatal pedestrian accidents at location-specific level: Application of xgboost and shap. Accid Anal Prev166, 106545. <https://doi.org/10.1016/j.aap.2021.106545>

Chang, L.-Y. , Mannering, F. L., 1998. Predicting vehicle occupancies from accident data: An accident severity approach. Transportation Research Record1635 (1), 93-104. <https://doi.org/10.3141/1635-13>

Chen, T. , Guestrin, C., Year. Xgboost: A scalable tree boosting system. In: Proceedings of the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785-794.

Chong, M., Abraham, A. , Paprzycki, M., 2005. Traffic accident analysis using machine learning paradigms. Informatica29 (1).

Delbosc, A. , Currie, G., 2011. Exploring the relative influences of transport disadvantage and social exclusion on well-being. Transport Policy18 (4), 555-562. <https://doi.org/https://doi.org/10.1016/j.tranpol.2011.01.011>

Dft. 2023. *Road safety data*. Department for Transport. Available: <https://www.data.gov.uk> [Accessed 2023-10-18].

Dong, X., Xie, K. , Yang, H., 2022. How did covid-19 impact driving behaviors and crash severity? A multigroup structural equation modeling. Accident Analysis Prevention172, 106687. <https://doi.org/10.1016/j.aap.2022.106687>

Goswamy, A., Abdel-Aty, M. , Islam, Z., 2023. Factors affecting injury severity at pedestrian crossing locations with rectangular rapid flashing beacons (rrfb) using xgboost and random parameters discrete outcome models. Accid Anal Prev181, 106937. <https://doi.org/10.1016/j.aap.2022.106937>

Gov.Uk. 2023. *Mobility scooters and powered wheelchairs: The rules*. UK Government. Available: <https://www.gov.uk/mobility-scooters-and-powered-wheelchairs-rules> [Accessed 2023-11-26].

Hussain, Q., Alhajyaseen, W. K. M., Kharbeche, M. , Almallah, M., 2023. Safer pedestrian crossing facilities on low-speed roads: Comparison of innovative treatments. Accident Analysis & Prevention180, 106908. <https://doi.org/10.1016/j.aap.2022.106908>

Hwang, Y. , Hong, G.-R. S., 2018. Predictors of driving cessation in community-dwelling older adults: A 3-year longitudinal study. Transportation Research Part F: Traffic Psychology and Behaviour52, 202-209. <https://doi.org/https://doi.org/10.1016/j.trf.2017.11.017>

Ifg. 2022. *Timeline of uk government coronavirus lockdowns and restrictions*. INSTITUTE FOR GOVERNMENT. Available: <https://www.instituteforgovernment.org.uk/data-visualisation/timeline-coronavirus-lockdowns> [Accessed 10-11-2023].

Isaacson, M. , Barkay, D., 2020. Mobility scooters in urban environments: A research agenda. Journal of transport & health18, 100917.

Jancey, J., Cooper, L., Howat, P., Meuleners, L., Sleet, D. , Baldwin, G., 2013. Pedestrian and motorized mobility scooter safety of older people. Traffic Injury Prevention14 (6), 647-653. <https://doi.org/10.1080/15389588.2012.749465>

Johnson, M., Rose, G. , Oxley, J., Year. Motorised mobility scooters—understanding a growing transport mode for older australians. In: Proceedings of the Proceedings of the 36th Australasian Transport Research Forum (ATRF), Brisbane, Australia, October 2, pp. 2013.

Khattak, A. J., Kantor, P. , Council, F. M., 1998. Role of adverse weather in key crash types on limited-access: Roadways implications for advanced weather systems. Transportation Research Record1621 (1), 10-19. <https://doi.org/10.3141/1621-02>

Kirk-Wade, E. 2023. *Uk disability statistics: Prevalence and life experiences*. House of Common. Available: <https://commonslibrary.parliament.uk/research-briefings/cbp-9602/> [Accessed 19-01-2024].

Konstantopoulos, P., Chapman, P. , Crundall, D., 2010. Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers’ eye movements in day, night and rain driving. Accident Analysis & Prevention 42 (3), 827-834. <https://doi.org/10.1016/j.aap.2009.09.022>

Krahelski, O., Sivarajah, S., Eardley, W., Smith, T. O. , Hing, C. B., 2022. Major trauma associated with mobility scooters: An analysis of the trauma audit research network. Injury53 (9), 3011-3018. <https://doi.org/https://doi.org/10.1016/j.injury.2022.06.019>

Laverdet, C., Malola, P., Meyer, T. , Delhomme, P., 2023. Electric personal mobility device driver behaviors, their antecedents and consequences: A narrative review. Journal of Safety Research86, 274-285. <https://doi.org/https://doi.org/10.1016/j.jsr.2023.07.006>

Lin, C., Wu, D., Liu, H., Xia, X. , Bhattarai, N. 2020. Factor identification and prediction for teen driver crash severity using machine learning: A case study. *Applied Sciences* [Online], 10.10.3390/app10051675

Luiu, C. , Tight, M., 2021. Travel difficulties and barriers during later life: Evidence from the national travel survey in england. Journal of Transport Geography91, 102973. <https://doi.org/https://doi.org/10.1016/j.jtrangeo.2021.102973>

Lundberg, S. M. , Lee, S.-I., 2017. A unified approach to interpreting model predictions. Advances in neural information processing systems30.

Mannering, F., 2018. Temporal instability and the analysis of highway accident data. Analytic Methods in Accident Research17, 1-13. <https://doi.org/10.1016/j.amar.2017.10.002>

Mannering, F., Bhat, C. R., Shankar, V. , Abdel-Aty, M., 2020. Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis. Analytic Methods in Accident Research25, 100113. <https://doi.org/https://doi.org/10.1016/j.amar.2020.100113>

Mannering, F. L., Shankar, V. , Bhat, C. R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic Methods in Accident Research11, 1-16. <https://doi.org/10.1016/j.amar.2016.04.001>

Michalaki, P., Quddus, M. A., Pitfield, D. , Huetson, A., 2015. Exploring the factors affecting motorway accident severity in england using the generalised ordered logistic regression model. Journal of Safety Research55, 89-97. <https://doi.org/https://doi.org/10.1016/j.jsr.2015.09.004>

Milton, J. C., Shankar, V. N. , Mannering, F. L., 2008. Highway accident severities and the mixed logit model: An exploratory empirical analysis. Accident Analysis & Prevention40 (1), 260-266. <https://doi.org/10.1016/j.aap.2007.06.006>

Nordbakke, S. , Schwanen, T., 2014. Well-being and mobility: A theoretical framework and literature review focusing on older people. Mobilities9 (1), 104-129. <https://doi.org/10.1080/17450101.2013.784542>

Oxley, J., Langford, J. , Charlton, J., 2010. The safe mobility of older drivers: A challenge for urban road designers. Journal of Transport Geography18 (5), 642-648. <https://doi.org/https://doi.org/10.1016/j.jtrangeo.2010.04.005>

Paparone, P., 2013. Lower extremity ulceration caused by medical scooter injury: A case series. Geriatric Nursing34 (1), 25-29. <https://doi.org/https://doi.org/10.1016/j.gerinurse.2012.06.011>

Park, B.-J. , Lord, D., 2009. Application of finite mixture models for vehicle crash data analysis. Accident Analysis & Prevention41 (4), 683-691. <https://doi.org/https://doi.org/10.1016/j.aap.2009.03.007>

Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S. , Mohammadian, A., 2020. Toward safer highways, application of xgboost and shap for real-time accident detection and feature analysis. Accident Analysis & Prevention136, 105405. <https://doi.org/https://doi.org/10.1016/j.aap.2019.105405>

Peters, A., Versteegen, M. G. J., Van Osch, F., Janzing, H. M. J. , Barten, D. G., 2022. Mechanism and severity of mobility scooter–related injuries. Traffic Injury Prevention23 (2), 112-117. <https://doi.org/10.1080/15389588.2021.1998469>

Rica. 2014. *Mobility scooters: A market study*. Research Institute for Consumer Affairs. Available: <https://www.gov.uk/government/publications/mobility-scooters-a-market-study> [Accessed 2024-01-11].

Salehian, A., Aghabayk, K., Seyfi, M. , Shiwakoti, N., 2023. Comparative analysis of pedestrian crash severity at united kingdom rural road intersections and non-intersections using latent class clustering and ordered probit model. Accident Analysis & Prevention192, 107231. <https://doi.org/https://doi.org/10.1016/j.aap.2023.107231>

Samerei, S. A. , Aghabayk, K., 2024. Analyzing the transition from two-vehicle collisions to chain reaction crashes: A hybrid approach using random parameters logit model, interpretable machine learning, and clustering. Accident Analysis & Prevention202, 107603. <https://doi.org/https://doi.org/10.1016/j.aap.2024.107603>

Scarano, A., Rella Riccardi, M., Mauriello, F., D'agostino, C., Pasquino, N. , Montella, A., 2023. Injury severity prediction of cyclist crashes using random forests and random parameters logit models. Accid Anal Prev192, 107275. <https://doi.org/10.1016/j.aap.2023.107275>

Schumann, H.-H., Haitao, H. , Quddus, M., 2023. Passively generated big data for micro-mobility: State-of-the-art and future research directions. Transportation Research Part D: Transport and Environment121. <https://doi.org/10.1016/j.trd.2023.103795>

Se, C., Champahom, T., Jomnonkwao, S., Karoonsoontawong, A. , Ratanavaraha, V., 2021. Temporal stability of factors influencing driver-injury severities in single-vehicle crashes: A correlated random parameters with heterogeneity in means and variances approach. Analytic Methods in Accident Research32, 100179. <https://doi.org/https://doi.org/10.1016/j.amar.2021.100179>

Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N. , Milton, J., 2017. Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research15, 41-55. <https://doi.org/10.1016/j.amar.2017.05.003>

Shankar, V. , Mannering, F., 1996. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. Journal of Safety Research27 (3), 183-194. <https://doi.org/https://doi.org/10.1016/0022-4375(96)00010-2>

Sherrington, A. 2022. *2 years of covid-19 on gov.Uk*. Government Digital Service. Available: <https://gds.blog.gov.uk/2022/07/25/2-years-of-covid-19-on-gov-uk/> [Accessed 20-2-2024].

Sun, Z., Wang, D., Gu, X., Abdel-Aty, M., Xing, Y., Wang, J., Lu, H. , Chen, Y., 2023. A hybrid approach of random forest and random parameters logit model of injury severity modeling of vulnerable road users involved crashes. Accid Anal Prev192, 107235. <https://doi.org/10.1016/j.aap.2023.107235>

Sun, Z., Xing, Y., Wang, J., Gu, X., Lu, H. , Chen, Y., 2022. Exploring injury severity of vulnerable road user involved crashes across seasons: A hybrid method integrating random parameter logit model and bayesian network. Safety Science150. <https://doi.org/10.1016/j.ssci.2022.105682>

Tefft, B. C., 2013. Impact speed and a pedestrian's risk of severe injury or death. Accident Analysis & Prevention50, 871-8. <https://doi.org/10.1016/j.aap.2012.07.022>

Toosizadeh, N., Bunting, M., Howe, C., Mohler, J., Sprinkle, J. , Najafi, B., 2014. Motorized mobility scooters: The use of training/intervention and technology for improving driving skills in aging adults-a mini-review. Gerontology60 (4), 357-365.

Train, K. E., 2009. Discrete choice methods with simulation Cambridge university press. 1139480375

Waseem, M., Ahmed, A. , Saeed, T. U., 2019. Factors affecting motorcyclists’ injury severities: An empirical assessment using random parameters logit model with heterogeneity in means and variances. Accident Analysis & Prevention123, 12-19. <https://doi.org/10.1016/j.aap.2018.10.022>

Washington, S., Karlaftis, M. G., Mannering, F. , Anastasopoulos, P., 2020. Statistical and econometric methods for transportation data analysis CRC press. 0429520751

Wen, X., Xie, Y., Wu, L. , Jiang, L., 2021. Quantifying and comparing the effects of key risk factors on various types of roadway segment crashes with lightgbm and shap. Accident Analysis & Prevention159, 106261. <https://doi.org/https://doi.org/10.1016/j.aap.2021.106261>

Who. 2013. *Disability*. Available: <https://apps.who.int/gb/ebwha/pdf_files/WHA66/A66_12-en.pdf> [Accessed 2024-01-17].

Who. 2022. *Ageing and health*. Available: <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health> [Accessed 2024-01-17].

Who. 2023. *Disability*. Available: <https://www.who.int/news-room/fact-sheets/detail/disability-and-health> [Accessed 2024-01-20].

Yang, C., Chen, M. , Yuan, Q., 2021. The application of xgboost and shap to examining the factors in freight truck-related crashes: An exploratory analysis. Accident Analysis & Prevention158, 106153. <https://doi.org/https://doi.org/10.1016/j.aap.2021.106153>

Yang, L. , Shami, A., 2020. On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing415, 295-316. <https://doi.org/https://doi.org/10.1016/j.neucom.2020.07.061>

Yuan, C., Li, Y., Huang, H., Wang, S., Sun, Z. , Li, Y., 2022. Using traffic flow characteristics to predict real-time conflict risk: A novel method for trajectory data analysis. Analytic Methods in Accident Research35. <https://doi.org/10.1016/j.amar.2022.100217>

Zhou, B., Wang, X., Zhang, S., Li, Z., Sun, S., Shu, K. , Sun, Q., 2020. Comparing factors affecting injury severity of passenger car and truck drivers. IEEE Access8, 153849-153861. <https://doi.org/10.1109/ACCESS.2020.3018183>

**Appendix:** **Estimation results for temporal models.**

See Figures A.1, A.2 and Tables A.1–A.4.

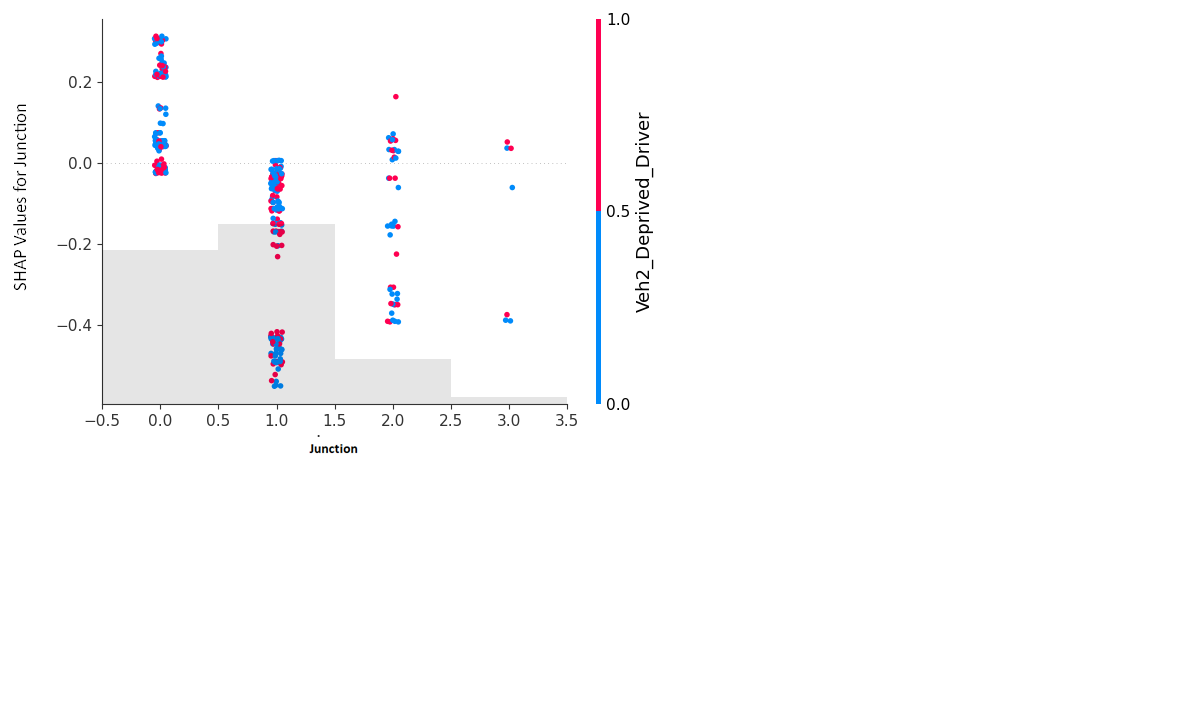


Figure A.2: SHAP interaction between Junction Type and Other-party deprived driver from XGBoost model

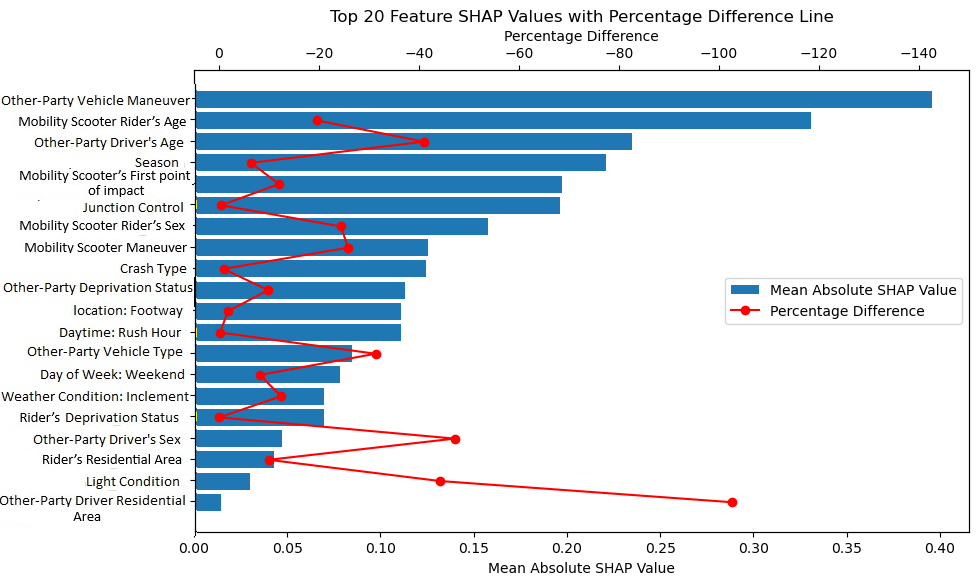


Figure A.1: Top 20 Features Based on SHAP Values for 2018-2022 model

|  |  |  |  |
| --- | --- | --- | --- |
| **Table A.1** | | | |
| RPBL-HMV Model Results for Mobility Scooter Rider Injury Severity 2018-2022 | | | |
| Variable | Parameter estimate | Z-value | Marginal effect |
| Severe Injury |
| Constant | -1.624 | -9.04 | |
| Senior Rider (1 if the rider was over the age of 80; 0 otherwise) | 0.858 | 4.02 | 0.0238 |
| Adult Rider (1 if the rider was over the age of 25 and under the age of 40; 0 otherwise) | -1.281 | -2.64 | -0.0104 |
| Young Rider (1 if the rider was under the age of 25; 0 otherwise) | -1.110 | -1.81 | -0.0060 |
| Rear-end Collision (1 if the first point of impact for mobility scooter was from the back of their vehicle; 0 otherwise) | 0.823 | 2.35 | 0.0057 |
| Sideswipe Collision (1 if the first point of impact for mobility scooter was on the side of their vehicle; 0 otherwise) | 0.586 | 2.60 | 0.0142 |
| Footway (1 if the mobility scooter was on the footway; 0 otherwise) | -0.523 | -1.86 | -0.0096 |
| Other Party Vehicle Type = Motorcycle (1 if the second vehicle was a motorcycle; 0 otherwise) | 0.978 | 3.88 | 0.0158 |
| Other Party Vehicle Type = Heavy vehicles (1 if the second vehicle was a heavy vehicle; 0 otherwise) | 1.056 | 2.92 | 0.0059 |
| Winter (1 if the crash occurred during winter months; 0 otherwise) | 1.010 | 4.43 | 0.0231 |
| ***Random parameters*** | | | |
| Give-way or Uncontrolled Junction (1 if the junction was a Give-way or uncontrolled junction; 0 otherwise) Mean | -2.331 | -1.76 | 0.0227 |
| Standard deviation | 4.456 | 2.22 | |
| ***Heterogeneity in the mean of the Random parameters*** | | | |
| Give way or Uncontrolled Junction : Other Party Vehicle Type = Motorcycle | 2.353 | 1.75 | |
| ***Heterogeneity in the variance of the Random parameters*** | | | |
| Give way or Uncontrolled Junction : Footway | -0.570 | -1.82 | |
| ***Model summary*** | | | |
| Number of observations | 1080 | | |
| Log-likelihood at convergence | -582.92 | | |
| McFadden Pseudo R2 *= 1 − (LL(β))/(LL(0))* | 0.221 | | |
| Adjusted R2 *= 1 − (LL(β)−N)/(LL(0))* | 0.203 | | |

N = Number of estimated parameters

|  |  |  |  |
| --- | --- | --- | --- |
| **Table A.2** | | | |
| RPBL-HMV Model Results for Mobility Scooter Rider Injury Severity: Pre-COVID Period | | | |
| Variable | Parameter estimate | Z-value | Marginal effect |
| Severe Injury |
| Constant | -2.104 | -5.58 |  |
| Senior Rider (1 if rider was over the age of 80; 0 otherwise) | 1.359 | 3.36 | 0.2440 |
| Adult rider (1 if rider was over the age of 25 and under the age of 40; 0 otherwise) | 0.727 | 1.83 | 0.1305 |
| Mobility Scooter skidded/overturned | 0.680 | 1.80 | 0.1221 |
| Junction Control = Traffic signal (1 if the junction was controlled by traffic signal; 0 otherwise) | 1.678 | 3.67 | 0.3014 |
| Other Party Vehicle = Young Driver (1 if the second vehicle's driver was aged between 16 and 30; 0 otherwise) | 1.052 | 2.82 | 0.1889 |
| Other Party Vehicle Maneuver = Turning (1 if the second vehicle was turning; 0 otherwise) | -1.627 | -3.38 | -0.2922 |
| Other Party Vehicle Maneuver = Move Off (1 if the second vehicle was moving off; 0 otherwise) | -1.434 | -3.11 | -0.2575 |
| Morning Rush Hour (1 if the crash occurred between 7-10 A.M.; 0 otherwise) | 0.994 | 2.16 | 0.1785 |
| Winter (1 if collision occurred during winter months; 0 otherwise) | 0.544 | 1.74 | 0.0976 |
| ***Model summary*** | | | |
| Number of observations | 285 | | |
| Log-likelihood at convergence | -142.04 | | |
| McFadden Pseudo R2 *= 1 − (LL(β))/(LL(0))* | 0.281 | | |
| Adjusted R2 *= 1 − (LL(β)−N)/(LL(0))* | 0.230 | | |

N = Number of estimated parameters

|  |  |  |  |
| --- | --- | --- | --- |
| **Table A.3** | | | |
| RPBL-HMV Model Results for Mobility Scooter Rider Injury Severity: COVID Period | | | |
| Variable | Parameter estimate | Z-value | Marginal effect |
| Severe Injury |
| Constant | -4.085 | -4.00 | |
| Senior Rider (1 if the rider was over the age of 80; 0 otherwise) | 1.638 | 2.45 | 0.0244 |
| Mobility Scooter Maneuver = Move Off (1 if the mobility scooter was moving off; 0 otherwise) | 1.523 | 1.96 | 0.0166 |
| Sideswipe Collision (1 if the first point of impact for mobility scooter was on the side of their vehicle; 0 otherwise) | 1.776 | 2.40 | 0.0614 |
| Give-way or Uncontrolled Junction (1 if the junction was a Give-way or uncontrolled junction; 0 otherwise) | 1.179 | 1.83 | 0.0398 |
| Footway (1 if the mobility scooter was on the footway; 0 otherwise) | -2.815 | -2.44 | -0.0359 |
| Single Vehicle Crash (1 if only one motorized vehicle was involved in the crash; 0 otherwise) | 2.969 | 3.25 | 0.0229 |
| Other Party Vehicle = Elder Driver (1 if the second vehicle's driver was aged over 60; 0 otherwise) | 2.342 | 2.50 | 0.0193 |
| Other Party Vehicle Maneuver = Turning (1 if the second vehicle was turning; 0 otherwise) | -2.127 | -2.01 | -0.0221 |
| Weekend (1 if the crash occurred during weekend days; 0 otherwise) | 1.801 | 2.36 | 0.0282 |
| ***Random parameters*** | | | |
| Deprived Rider (1 if the mobility scooter rider was deprived; 0 otherwise) Mean | -2.631 | -1.77 | 0.0571 |
| standard deviation | 4.266 | 2.24 | |
| ***Model summary*** | | | |
| Number of observations | 252 | | |
| Log-likelihood at convergence | -108.10 | | |
| McFadden Pseudo R2 *= 1 − (LL(β))/(LL(0))* | 0.381 | | |
| Adjusted R2 *= 1 − (LL(β)−N)/(LL(0))* | 0.312 | | |

N = Number of estimated parameters

|  |  |  |  |
| --- | --- | --- | --- |
| **Table A.4** | | | |
| RPBL-HMV Model Results for Mobility Scooter Rider Injury Severity: Post-COVID Period | | | |
| Variable | Parameter estimate | Z-value | Marginal effect |
| Severe Injury |
| Constant | -1.591 | -5.48 |  |
| Senior Rider (1 if the rider was over the age of 80; 0 otherwise) | 1.247 | 3.24 | 0.0457 |
| Rear-end Collision (1 if the first point of impact for mobility scooter was from the back of their vehicle; 0 otherwise) | 1.243 | 2.26 | 0.0138 |
| Darkness (1 if there was no lightning at the crash location; 0 otherwise) | -2.187 | -1.86 | -0.0117 |
| Single Vehicle Crash (1 if only one motorized vehicle was involved in the crash; 0 otherwise) | 1.344 | 2.89 | 0.0216 |
| Other Party Vehicle = Deprived Driver (1 if the second vehicle's driver was deprived; 0 otherwise) | 0.569 | 1.68 | 0.0339 |
| Other Party Vehicle Maneuver = Reversing (1 if the second vehicle was reversing; 0 otherwise) | -1.192 | -1.95 | -0.0177 |
| Winter (1 if the crash occurred during winter months; 0 otherwise) | 1.401 | 3.26 | 0.0432 |
| ***Random parameters*** | | | |
| Give way or Uncontrolled Junction (1 if the junction was a give way or uncontrolled junction; 0 otherwise) Mean | -2.285 | -1.72 | 0.0550 |
| standard deviation | 3.253 | 1.77 | |
| ***Heterogeneity in the mean of the Random parameters*** | | | |
| Give way or Uncontrolled Junction: Summer | 2.120 | 1.78 | |
| Give way or Uncontrolled Junction: Other Party Vehicle = Deprived Driver | -3.494 | -1.76 | |
| ***Model summary*** | | | |
| Number of observations | 395 | | |
| Log-likelihood at convergence | -208.27 | | |
| McFadden Pseudo R2*= 1 − (LL(β))/(LL(0))* | 0.239 | | |
| Adjusted R2 *= 1 − (LL(β)−N)/(LL(0))* | 0.195 | | |

N = Number of estimated parameters