



Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

A comprehensive analysis of the relationships between the built environment and traffic safety in the Dutch urban areas

Mehrnaz Asadi^{a,*}, Mehmet Baran Ulak^a, Karst T. Geurs^a, Wendy Weijermars^b, Paul Schepers^c

^a University of Twente, Department of Civil Engineering, Faculty of Engineering Technology, P.O. Box 217, 7500 AE Enschede, the Netherlands

^b SWOV Institute for Road Safety Research, P.O. Box 93113, 2509 AC The Hague, the Netherlands

^c Ministry of Infrastructure and the Environment, Rijkswaterstaat, P.O. Box 2232, 3500 GE Utrecht, the Netherlands

ARTICLE INFO

Keywords:

Traffic safety
Built-environment
Land-use design
Proximity to facilities
Spatial spillover effects
Hurdle Negative Binomial model

ABSTRACT

Built-environment factors potentially alleviate or aggravate traffic safety problems in urban areas. This paper aims to investigate the relationships of these factors with vehicle-bicycle and vehicle-vehicle property damage only (PDO) and killed and severe injury (KSI) crashes in urban areas. For this purpose, an area-level analysis using 100x100m² cells, along with a Spatial Hurdle Negative Binomial regression model were employed. The study area is composed of a selection of municipalities in the Netherlands-Randstad Area where major land-use developments have occurred since the 1970s. The study was conducted by developing a rich dataset composed of various national and local databases. The findings reveal that built-environment factors and land-use policies have substantial impacts on safety, which cannot be neglected. The factors explaining the land-use density and diversity in the area (e.g., urbanity and function mixing levels), as well as the land-use design characteristics (indicated by average age of the neighborhoods), traffic and road network characteristics, and proximity to different destinations influence the probability, frequency, and severity of crashes in urban areas. Furthermore, low socioeconomic levels are associated with a higher frequency of traffic crashes.

1. Introduction

Traffic safety policy and research have mostly focused on road network measures and road design, vehicle safety, and road user behavior in the past decades. However, the role of built-environment factors in traffic safety has not yet been fully examined. For the Netherlands, this is rather surprising given the importance attributed to urban spatial planning and it being one of the most planned countries in the world. Despite the considerable improvements that have taken place in traffic safety in urban areas over the decades, one-third of road fatalities still occur in the built-up areas of the Netherlands (SWOV, 2021). Considering the aim of construction of on average 75,000 new homes per year until 2025 (MinIenW, 2018), understanding the relationships between the built-environment factors and traffic safety is invaluable to improving traffic safety.

Built-environment characteristics are usually represented by indicators called “5Ds”, namely (*population*) *density*, (*land-use*) *diversity*, (*land-use*) *design*, *distance to transit*, and *accessibility to destinations* (Ewing & Cervero, 2010). Different built-environment factors can increase or decrease traffic crashes as they contribute to variations in exposure,

speeds, and conflicts between road users (Ewing & Dumbaugh, 2009; Merlin et al., 2020b; Saha et al., 2020). Previous studies about the effects of built-environment factors on traffic crashes have mainly accounted for the impacts of the road network characteristics and a limited number of other built-environment factors. The studies have resulted in mixed and sometimes contradicting findings on the magnitude and direction of the effects of the built-environment factors on traffic crashes. Some studies, for example, show an increase in population and employment density is associated with a higher frequency of vehicle and bicycle crashes (even after controlling for exposure) but a lower frequency of severe crashes (Obelheiro et al., 2020; Osama & Sayed, 2017). However, at the same time, higher residential density areas promote walking and cycling and reduce car-use which can lead to a decline in all types of traffic crashes including both vehicle only and bicycle involved crashes (Osama & Sayed, 2017; Saha et al., 2020; Schepers et al., 2019). Furthermore, some land-use characteristics such as commercial and educational zones have been reported to negatively affect cyclist safety (Mukoko & Pulugurtha, 2019; Osama & Sayed, 2017). High land-use diversity (i.e., mixed land-use index (MXI)) was shown to decrease (KSI) crash frequency in vehicle-only and vehicle-bicycle crashes (Chen

* Corresponding author.

E-mail address: m.asadi@utwente.nl (M. Asadi).

<https://doi.org/10.1016/j.aap.2022.106683>

Received 18 November 2021; Received in revised form 20 April 2022; Accepted 20 April 2022

Available online 28 April 2022

0001-4575/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

& Shen, 2016). This is because higher MXI leads to shorter travel distances which results in reduced car-use and consequently risk exposure specifically for the cyclists (Chen & Shen, 2016; Schepers et al., 2019). Proximity to facilities such as transit stops, commercial and service locations, on the other hand, were shown to increase the risk of crashes that involve vehicles and/or bicycles (Kim et al., 2010; Osama & Sayed, 2017; Schepers, 2021b; Vandenbulcke et al., 2014; Wei & Lovegrove, 2013). However, Kim et al. (2011) suggested that increased proximity to roads is associated with reduced crash severity.

This paper analyses the impact of a comprehensive set of built-environment variables on vehicle-only and bicycle-involved types of crashes. Most literature focuses on determining the impact of a specific or limited set of built-environment factors on different crash types. Only a few studies have examined the impact of a comprehensive set of built-environment factors on traffic safety. For instance, Obelheiro et al. (2020) analyzed the impact of built-environment factors on the number of crashes at newly developed traffic safety zones. However, the authors focused only on the “road network and infrastructure design” indicated by the density of different types of intersections, traffic signals, and the proportion of different types of roads. Similarly, Yu and Xu (2018) employed the “nonmotorized infrastructure design” features indicated by the ratio of sidewalk/bike lane length to the street length. Ukkusuri et al. (2012) accounted for different built-environment factors at the census tracts. However, only a limited number of land-use factors such as land-use density were utilized. Nonetheless, none of these studies examined the effects of the “land-use design” characteristics on traffic safety. As an exception, Xie et al. (2019) studied the effects of land-use changes and showed that conversions of the business, commercial and residential land-use classes could increase the severe crash frequency in urban areas.

Disregarding the above-mentioned collective effects of built-environment factors can potentially lead to an under/overestimation of their importance due to the unobserved effects of missing variables. Moreover, these factors do not have the same effect on different types of crashes (e.g., motor-vehicle vs. bicycle crashes). However, literature falls short of identifying and comparing the variations in the magnitude of these effects. For example, it is still not clear how different levels of MXI affect vehicle-bicycle crashes compared to vehicle-vehicle crashes. Similarly, the effects of proximity to different facilities on the severity level of crashes remain unclear.

From another perspective, built-environment characteristics change over time and space as a result of different land-use and transport policies. These changes naturally affect travel behavior and, in turn, traffic safety. For example, the Netherlands underwent major developments during the 1970s and 1990s, with policies such as “Dutch New Towns” (1960–1985) and “VINEX Areas” (1993). Safer road network designs, accessible public transport, and short distances to services (e.g., supermarkets) were some of the goals for the development of these newer areas. These spatial policies and developments together with transport planning have played a role in the declining numbers of fatal crashes (Schepers, 2021a; b; Schepers et al., 2019). Nonetheless, these changes add to the complexity of the problem of understanding the effect of the built-environment factors on safety because the homogeneity across different areas is lost.

To address the issues summarized in the preceding paragraphs, this study developed and utilized a wide-ranging set of built-environment, land-use, road network, and socioeconomic variables that have an effect on traffic safety. Moreover, the study accounts for real traffic exposure and transportation infrastructure characteristics in the analysis. To be specific, property damage only (PDO) and fatality/severe injury (KSI) Vehicle-Bicycle (V&B) and Vehicle-Vehicle (V&V) crashes were investigated. For this purpose, a rich dataset was compiled by integrating various national and local databases covering several domains ranging from built-environment characteristics to traffic features. A square grid-based macro-level analysis was adopted. The spatial contiguity effects of the built-environment factors were particularly

considered in the analysis to address the spatial heterogeneities and dependencies. The compiled rich dataset was analyzed using a Hurdle Negative Binomial model which can handle count datasets involving an excessive number of zero outcomes (a common issue in crash count data) (Hosseinpour et al., 2013). The analysis was conducted for a case study area covering a selection of municipalities in the most urbanized part of the Netherlands, the “Randstad Area” which has experienced several changes in the built-environment over the decades.

The contribution of this paper to the literature is threefold. Firstly, the study examines the collective effects of a comprehensive set of built-environment factors on traffic safety, exploiting a rich dataset compiled from different national and local databases. Secondly, the study compares the contribution of similar built-environment characteristics on both V&B and V&V crashes in PDO and KSI types. Thirdly, the study presents a spatial-statistical modeling approach that can address prevalent area-based modeling issues such as crash assignment (Kocatepe et al., 2019), spatial dependency and heterogeneity (Ziakopoulos & Yannis, 2020), and an excessive number of zeros in the crash counts (Lord & Mannering, 2010).

2. Study Area, geographical unit Selection, and modeling approaches

2.1. Study Area

This study particularly focuses on the urban parts of the cities that are demarcated as built-up and recreational areas by the Central Bureau of Statistics in the Netherlands (CBS, 2015) (Fig. 1-a). Furthermore, as this study only focuses on urban traffic crashes, motorways (i.e., inter-city roads and trunk links) and crashes that happened on these roads were excluded from the analysis.

The case study was conducted in the Randstad Area which is the most urbanised area in the Netherlands and includes the four major cities in the country, i.e., Amsterdam, Rotterdam, The Hague, and Utrecht (Fig. 1-b). These cities include new urban centers and housing developments that have been expanded from the city center towards the city boundaries since the 1970 s. Several new towns (built in the 1970 s), such as Almere and Houten were also included in the analysis (Fig. 1-c). The built-environment characteristics of these new urban areas (built after the 1970 s) mainly differ from historic city centers in terms of destination accessibility, distance to transit and infrastructure design with separated road space for different urban transport modes (Bach et al., 2006).

2.2. Geographical units of analysis

Analyzing traffic crash data requires specific considerations regarding the level and scale of the analysis, as well as the specification of the crash data. From the existing methods adopted to analyze the impacts of the “built-environment factors” on safety the area-level analysis is one of the most promising approaches. Depending on the study approach and availability of data, alternative units of analysis such as traffic analysis zones (TAZs) (Obelheiro et al., 2020; Osama & Sayed, 2017), administrative areas (e.g., census tracts or postcode areas (Abdel-Aty et al., 2013; Osama & Sayed, 2017; Saha et al., 2020)), or uniform formats (e.g., hexagons (Cui et al., 2021) and square grids (Gladhill & Monsere, 2012; Kim et al., 2010)) have been selected in the literature.

The uniform formats such as square grids are more advantageous than administrative areas considering that boundaries of these areas are demarcated by the roads where crashes occur. Therefore, there is always a problem of “crash assignment” when administrative areas are used (Gladhill & Monsere, 2012; Kocatepe et al., 2019). Utilizing the square grids as the unit of analysis allows the inclusion of all crashes without the need to give special consideration to crashes on “boundaries”. Moreover, the uniform units allow the researcher to control for the unit

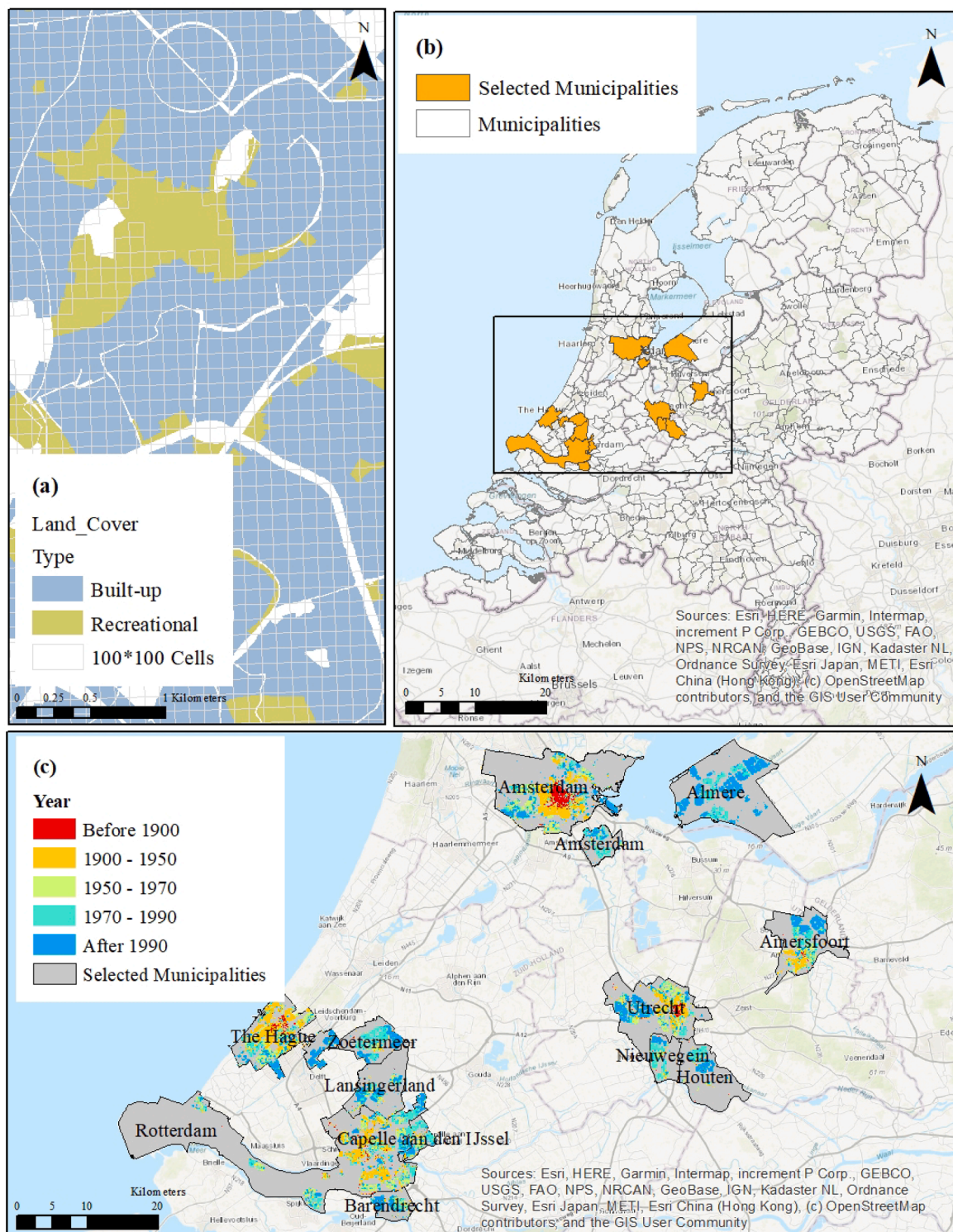


Fig. 1. (a) Representation of the studied urban areas (built-up and recreational) in the 100-cells; (b) Selected municipalities in Randstad region- the Netherlands; (c) Year of housing developments (weighted average) in the 100-cells.

size. Due to the practicality of this method, the availability of $100 \times 100m^2$ grid cells (100-cells), and the availability of extensive land-use and socioeconomic and demographic data in this format (CBS, 2017a) we divided the study area into 100-cells (Fig. 1-a).

2.3. Modeling spatial spillover effects

A major concern in traffic safety analysis is related to spatial dependency and heterogeneity biases. These biases affect the performance of crash analysis due to unobserved factors associated with the locations of crashes (Xu & Huang, 2015). To alleviate these issues, two alternative approaches are proposed in the literature (Cai et al., 2016): 1) spatial

autocorrelation effects: accounting for the effects of the unobserved exogenous variables by involving spatial autocorrelation among the units of analysis; 2) spatial spillover effects: accounting for the effects of the observed exogenous variables (e.g., built-environment factors) influencing the dependent variable (e.g., crash frequency) at the targeted and neighboring areas. The *spatial spillover effect* approach is appealing for macro-level studies due to the ease of modeling and its superior interpretability and communicability features (Cai et al., 2016).

To test the spatial dependency and heterogeneity, the Global Moran's I-tests (Ziakopoulos & Yannis, 2020) were conducted. The total number of PDO and KSI crashes in V&B and V&V types were tested. The tests results revealed that spatial autocorrelations exist between the crash frequencies of the neighboring analysis units. Therefore, we accounted for the spillover effects in the models. For this purpose, "queen contiguity" was used to identify the neighboring cells for each cell (eight neighbors in the case of square cells - Fig. 2). Then, the average values of identified "spillover variables" for these eight neighboring cells were computed and assigned to each cell. These *neighboring variables* were included as independent variables in the statistical models.

2.4. Hurdle Negative Binomial model

Several counts/continuous modeling techniques have been applied to analyze the crash frequencies and rates (Lord & Mannering, 2010). Crash data commonly include an excessive number of zeros due to the lack of crashes in the majority of the entities (e.g., roadway segments). The use of data with "excessive zeros" in modeling which does not address this issue can result in incorrect and invalid parameter estimations (Lord & Mannering, 2010). To address this problem, models such as the Tobit model for continuous dependent variables (e.g., crash rates) (Anastasopoulos et al., 2012; Chen et al., 2021; Ulak et al., 2018), as well as Hurdle and zero-inflated (ZI) models for discrete dependent variables (e.g., crash counts) (Cai et al., 2016; Chen et al., 2022; Katrakazas et al., 2021), have been proposed. Additionally, novel resampling data analysis methods are also introduced for handling unbalanced data in accident analysis studies. These methods are applied for developing predictive models using synthetic data for uncommon crash types (e.g., bus-involved crashes) and in the presence of excessive zeros (Ariannezhad et al., 2021; Chen et al., 2022; Morris & Yang, 2021; Yahaya et al., 2019).

The superiority of the dual-state models compared to the conventional Poisson and NB models has been previously considered for both

area-based and network-based analysis (Cai et al., 2016; Hosseinpour et al., 2013). The major difference between the ZI and the Hurdle models is that the latter only accounts for the sampling (i.e., random) zeros, whilst the ZI model accounts for both sampling zeros and structural zeros (Hosseinpour et al., 2013). The structural zeros indicate an "impossibility" of crash occurrence (e.g., when there is no traffic), which disregards the fact that traffic crashes can happen on any road regardless of the safety level (Hosseinpour et al., 2013; Yu et al., 2019). Therefore, the Hurdle model is considered the more appropriate approach in this study.

The Hurdle models are modified two-part count models of which the first part is a binary model that handles whether the dependent variable crosses the "hurdle" (i.e., whether a crash happens in a cell) and the second part consists of a count model which is truncated at zero (Mullahy, 1986). Alternative probability distributions such as Poisson, NB, and lognormal can be used in the count part of the model depending on the goodness-of-fit to the data (Cai et al., 2016; He et al., 2019; Ma et al., 2015). Hurdle model was utilized in traffic safety analysis by Hosseinpour et al. (2013) to investigate the effects of the road network characteristics on pedestrian-vehicle crash frequencies on the roads. Even though, they showed that the HP models offer better modeling performance compared to the HNB models. In this study, we used Hurdle NB Regression models, as not only the NB distribution can be generalized to Poisson distribution but also comparisons of the performance of the Hurdle Poisson (HP) and Hurdle NB (HNB) models showed a better goodness-of-fit for the HNB model. The HNB model can be described as Eq. (1):

$$f(y_i) = \begin{cases} p_i, y_i = 0 \\ (1 - p_i)g(y_i), y_i = 1, 2, \dots \end{cases} \quad (1)$$

$$g(y_i | \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1)\Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i}\right)^{\alpha^{-1}} \left(\frac{\mu_i}{\alpha^{-1} + \mu_i}\right)^{y_i} \quad (2)$$

where p_i specifies the probability of crash occurrence in the analysis unit ($0 < p_i < 1$); p_i can have Logit (or Probit) distribution and y_i indicates the number of events in the units of analysis; $g(y_i)$ is the negative Binomial distribution with mean μ_i and NB dispersion parameter α (Eq. (2)). The model parameters are estimated by the Maximum Likelihood method, where the specification of the likelihood function has the advantage of separately maximizing both the count and the hurdle components of the model (Zeileis et al., 2008). In this study, we used R software, "pscl" package (Jackman et al., 2015) for estimating the model parameters.

3. Data processing approach

For the analysis, a rich dataset was compiled using several sources. The final set of variables in this dataset and summary statistics of these variables are listed in Table 1. The variables in this table are grouped into seven categories based on the characteristics of the variables. In each cell, crashes (based on the type) were summed up to calculate the total number of crashes (Fig. 3). In addition, line and polygon features intersecting with each cell were also aggregated (Figs. 4 and 5). That is, for example, roadways that fall into a cell were combined to find the length of each roadway type (e.g., residential, etc.) in that cell. Then, the proportion of different types of roads in each cell was calculated and assigned to the cell. For this purpose, Eq. (3) was used.

$$R_i^r = \sum_{j=1}^n l_{ij}^r / \sum_{r=1}^k \sum_{j=1}^n l_{ij}^r \quad (3)$$

where R_i^r indicates the proportion of road type r in the corresponding cell i ; l_{ij}^r is the length of road segment j with type r in cell i . For polygon features, such as the area of different land-use classes, a simple proportioning method (Eq. (4)) was used.

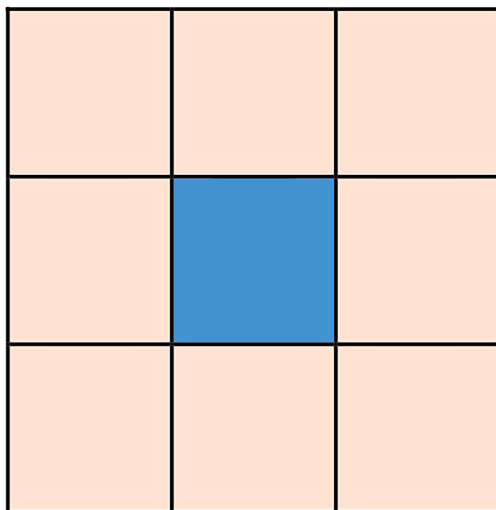


Fig. 2. Queen contiguity (light orange cells are considered neighbors of the blue cell in the center).

Table 1
Descriptive Statistics of the explanatory and response variables used in the models.

| Type of variable | Abbreviations | V&B-Models Variables N _{V&B} = 48,474 | | | | V&V-Models Variables N _{V&V} = 53,579 | | | | Description of the variable (Determined in the corresponding 100-cell) | source |
|---|-----------------------|---|-------|------|------|---|-------|------|--------------------------------------|---|--|
| | | min | max | mean | s.d. | min | max | mean | s.d. | | |
| Crash outcome | PDO | 0.00 | 15.00 | 0.11 | 0.46 | 0.00 | 81.00 | 0.49 | 1.83 | Number of PDO crashes. | (BRON, 2015-2019) |
| | KSI | 0.00 | 9.00 | 0.05 | 0.27 | 0.00 | 8.00 | 0.04 | 0.24 | Number of KSI crashes. | |
| Weighted average year of construction of housing properties | Before 1900 | 0.00 | 1.00 | 0.02 | 0.16 | 0.00 | 1.00 | 0.02 | 0.15 | Ave. year is before 1900 (True = 1, False = 0). | (BAG 2020) |
| | 1900–1950 | 0.00 | 1.00 | 0.14 | 0.35 | 0.00 | 1.00 | 0.13 | 0.34 | Ave. year is between 1900 and 1950 (True = 1, False = 0). | |
| | 1950–1970 | 0.00 | 1.00 | 0.12 | 0.32 | 0.00 | 1.00 | 0.11 | 0.31 | Ave. year is between 1950 and 1970 (True = 1, False = 0). | |
| | 1970–1990 | 0.00 | 1.00 | 0.18 | 0.38 | 0.00 | 1.00 | 0.16 | 0.37 | Ave. year is between 1970 and 1990 (True = 1, False = 0). | |
| Demographic and socioeconomic | After 1990 | 0.00 | 1.00 | 0.24 | 0.42 | 0.00 | 1.00 | 0.22 | 0.41 | Ave. year is after 1990 (True = 1, False = 0). | |
| | #Households | 0.00 | 7.55 | 0.31 | 0.42 | 0.00 | 7.55 | 0.28 | 0.41 | Number of households (x100). | (CBS, 2017a, 2017b) |
| | ∝Household with Child | 0.00 | 1.00 | 0.18 | 0.23 | 0.00 | 1.00 | 0.16 | 0.23 | Proportion of households with children. | |
| | %Rental Houses | 0.00 | 1.00 | 0.44 | 0.30 | 0.00 | 1.00 | 0.42 | 0.31 | Percentage of rental housing properties. | |
| | Value of Houses | 0.00 | 1.47 | 0.22 | 0.15 | 0.00 | 1.47 | 0.21 | 0.16 | Average value of housing properties (x 1 Million Euro). | |
| Traffic variables | %High Income | 0.00 | 0.80 | 0.17 | 0.16 | 0.00 | 0.80 | 0.16 | 0.16 | Percentage of high-income households. | |
| | log(VKT) | -10.20 | 10.44 | 5.54 | 1.39 | -10.20 | 10.82 | 5.52 | 1.41 | Log (total average daily VKT). | (Geofabrik, 2019; |
| | Speed Limit = 30 | 0.00 | 1.00 | 0.56 | 0.50 | - | - | - | - | Maximum speed limit = 30 (True = 1, False = 0). | (Jedlicka et al., 2015) |
| | Speed Limit = 50 | 0.00 | 1.00 | 0.33 | 0.47 | 0.00 | 1.00 | 0.30 | 0.46 | Maximum speed limit = 50 (True = 1, False = 0). | |
| | Speed Limit > 50 | - | - | - | - | 0.00 | 1.00 | 0.14 | 0.35 | Maximum speed limit > 50 (True = 1, False = 0). | Rijkswaterstraat, 2018, 2019a) |
| | ∝Primary Roads | - | - | - | - | 0.00 | 1.00 | 0.02 | 0.13 | Proportion of Primary roads. | |
| | ∝Secondary Roads | 0.00 | 1.00 | 0.07 | 0.21 | 0.00 | 1.00 | 0.07 | 0.21 | The proportion of Secondary roads. | Regional traffic model outputs: |
| | ∝Tertiary Roads | 0.00 | 1.00 | 0.09 | 0.23 | 0.00 | 1.00 | 0.08 | 0.22 | The proportion of Tertiary roads. | Amsterdam (2020), metropolitan region of Rotterdam-The Hague (2017), Province of Utrecht (2015), and city of Almere (2017) |
| ∝Residential Roads | 0.00 | 1.00 | 0.43 | 0.42 | 0.00 | 1.00 | 0.39 | 0.42 | The proportion of Residential roads. | (Leefomgeving, 2019) | |
| Land-use types and indices | ∝Accommodation Area | 0.00 | 2.14 | 0.00 | 0.04 | 0.00 | 2.14 | 0.00 | 0.04 | Proportion of Accommodation (e.g., hotels, hostels) areas. | |
| | ∝Sport Area | 0.00 | 2.73 | 0.00 | 0.02 | 0.00 | 2.73 | 0.00 | 0.02 | Proportion of Sports facility areas. | |
| | ∝Shop Area | 0.00 | 2.21 | 0.01 | 0.07 | 0.00 | 3.40 | 0.01 | 0.07 | Proportion of Shop areas. | |
| | ∝Meeting Area | 0.00 | 8.44 | 0.01 | 0.08 | 0.00 | 8.44 | 0.01 | 0.08 | Proportion of Meeting (e.g., cinema, church, etc.) areas. | |
| | ∝Office Area | 0.00 | 8.68 | 0.03 | 0.17 | 0.00 | 8.68 | 0.03 | 0.16 | Proportion of Office areas. | |
| | MXI | 0.00 | 1.00 | 0.60 | 0.41 | 0.00 | 1.00 | 0.55 | 0.42 | Mixed land-use function index. | |
| Proximity characteristics | Urbanity | 0.00 | 4.98 | 1.03 | 0.87 | 0.00 | 4.98 | 0.97 | 0.87 | Urbanity (number of registered address density × 100). | (CBS, 2017b) |
| | Secondary school | 0.00 | 10.00 | 0.72 | 1.01 | - | - | - | - | 1/distance to secondary schools. | (CBS, 2017a) |
| | Train Station | 0.00 | 10.00 | 0.35 | 0.52 | - | - | - | - | 1/distance to train stations. | |
| | Grocery Shops | 0.00 | 10.00 | 1.71 | 2.12 | 0.00 | 10.00 | 1.56 | 2.08 | 1/distance to supermarkets and daily shops. | |
| | Main Roads | 0.00 | 10.00 | 0.43 | 0.53 | 0.00 | 10.00 | 0.39 | 0.52 | 1/distance to main roads. | |

(continued on next page)

Table 1 (continued)

| Type of variable | Abbreviations | V&B-Models Variables N _{V&B} = 48,474 | | | | V&V-Models Variables N _{V&V} = 53,579 | | | | Description of the variable (Determined in the corresponding 100-cell) | source |
|-----------------------------------|-----------------------|---|-------|-------|------|---|------|------|------|---|--------------------------------|
| | | min | max | mean | s.d. | min | max | mean | s.d. | | |
| Cycling activities and facilities | log(BKT) | -7.05 | 9.52 | 2.74 | 4.14 | - | - | - | - | Log (total daily bicycle counts*kilometer-traveled (BKT)). | (Nationale Fietstelweek, 2016) |
| | Cycling Speed | 0.00 | 30.00 | 19.02 | 8.42 | - | - | - | - | Maximum average cycling speed. | (Fietstersbond, 2016) |
| | αSeparated Cycle Path | 0.00 | 0.51 | 0.04 | 0.07 | - | - | - | - | Proportion of separated cycling paths. | |
| | αSuggested Cycle lane | 0.00 | 0.39 | 0.01 | 0.03 | - | - | - | - | Proportion of cycle lanes. | |
| | αMixed Traffic | 0.00 | 0.79 | 0.15 | 0.10 | - | - | - | - | Proportion of other road types where cycling is allowed. | |
| Neighboring variables | #Households | 0.00 | 7.55 | 0.39 | 0.36 | 0.00 | 7.55 | 0.36 | 0.36 | Number of households (x 100). | - |
| | log(VKT) | 0.00 | 9.42 | 5.99 | 0.88 | -5.58 | 9.42 | 5.99 | 0.89 | Log (total VKT). | - |
| | αSecondary Roads | 0.00 | 0.51 | 0.04 | 0.07 | 0.00 | 0.51 | 0.04 | 0.07 | Proportion of Secondary roads. | - |
| | αTertiary Roads | 0.00 | 0.33 | 0.04 | 0.06 | 0.00 | 0.33 | 0.04 | 0.06 | Proportion of Tertiary roads. | - |
| | αResidential Roads | 0.00 | 0.36 | 0.11 | 0.07 | 0.00 | 0.36 | 0.10 | 0.07 | Proportion of Residential roads. | - |
| | αShop Area | 0.00 | 2.13 | 0.01 | 0.05 | 0.00 | 2.29 | 0.01 | 0.05 | Proportion of Shops areas. | - |
| | αOffice Area | 0.00 | 4.01 | 0.03 | 0.12 | 0.00 | 4.01 | 0.03 | 0.11 | Proportion of Office areas. | - |
| | αAccommodation Area | 0.00 | 5.21 | 0.42 | 0.40 | 0.00 | 5.21 | 0.39 | 0.40 | Accessibility to Main roads. | - |
| | log(BKT) | -5.11 | 8.97 | 4.55 | 1.52 | - | - | - | - | Log(total BKT). | - |

$$S_i^l = \frac{1}{100^2} \sum_{j=1}^n \left(\frac{S_{ij}^l}{B_j} \cdot a_{ij} \right) \tag{4}$$

where S_i^l is the proportion of the polygon feature S type l in the cell i; S_{ij}^l is the area of the S_i^l in building block j; B_j is the total area of the building block j, and a_{ij} is the area of the building block j that intersects with cell i.

These values were then used to estimate the mixed land-use index (MXI) of the cells, which indicates the ratio of housing function areas to all other functions (Van Den Hoek, 2008). It is worth noting that the utilized MXI uses a floor-space area of the land-use functions provided in the RUDIFUN database (Leefomgeving, 2019) rather than the ground-space area. This results in more accurate MXI values (Harbers et al., 2019).

3.1. Motor-vehicle and cycling volumes data preparation

One major issue in traffic safety studies is the lack of reliable exposure variables, e.g., traffic volume, data for the minor roads (Schepers, 2021b). Previous studies usually utilized “proxy variables” such as the trip production measures (i.e., population or employment), (bicycle) lane-kilometer, mode choice, or accessibility to train to account for the traffic exposure (Chen, 2015; Merlin et al., 2020b; Obelheiro et al., 2020; Wang et al., 2019). This means that traffic and cycling exposures were not explicitly controlled for whilst analyzing the effects of built-environment factors. Although it is clear that such proxy variables can lead to less accurate estimates of these effects, use of these variables can give a fair indication of the traffic/cycling volumes in the absence of high quality traffic volume data.

In this research, we used databases from different sources such as Open Transport Map (OTM) (Jackman et al., 2015) and Open Street Map (OSM) (Geofabrik, 2019). Furthermore, traffic count data were obtained from the traffic models of municipalities of Amsterdam, the metropolitan region of Rotterdam-The Hague, the Province of Utrecht, and the city of Almere. However, the traffic models had two shortcomings: 1) the traffic models provide weekday counts and do not model weekend counts, and 2) the traffic models mainly provide the vehicle counts on the major urban roads (arterials/distributor roads) and do not cover minor roads (e.g., residential or access roads), which constitute a major portion of the road network in urban areas.

To remedy these shortcomings, we used regression models (Table 2) based on the data for the selected cities from a database called “Intensities on WEGVakken” (INWEVA) provided by the ministry of infrastructure and the environment (Rijkswaterstraat, 2019a). This database consists of traffic intensities on major roads both for weekdays and weekends based on different vehicle categories. These models were used to estimate the vehicle counts at the weekend (WE_n) based on the weekday (WD_n) data. Estimated traffic volumes were then assigned to the minor roads. Finally, traffic volumes were used to calculate the total vehicle-kilometre-travelled (VKT) values in the cell i, using Eq. (5).

$$VKT_i = \sum_{j=1}^n v_{ij} \cdot l_{ij} \tag{5}$$

where VKT_i is the total VKT in cell i; v_{ij} is the average daily vehicle count on the road link j in the cell i; l_{ij} is the length of the road link j (in km) located in cell i.

For bicycle volume and average cycling speed, this study used data including the number of registered bicycle rides per road section during the entire Bicycle Counting Week (BCW) in 2016. This project had almost 51,000 participants in the Netherlands and was carried out from the 19th to 25th September 2016 (Nationale Fietstelweek, 2016). To check the validity of the BCW data, Van Petegem et al. (2021) conducted an analysis using data from manual bicycle counts, detector loops, and the BCW data in Amsterdam. The statistical tests found a high correlation between BCW data and manual counts (Pearson correlation = 0.78).

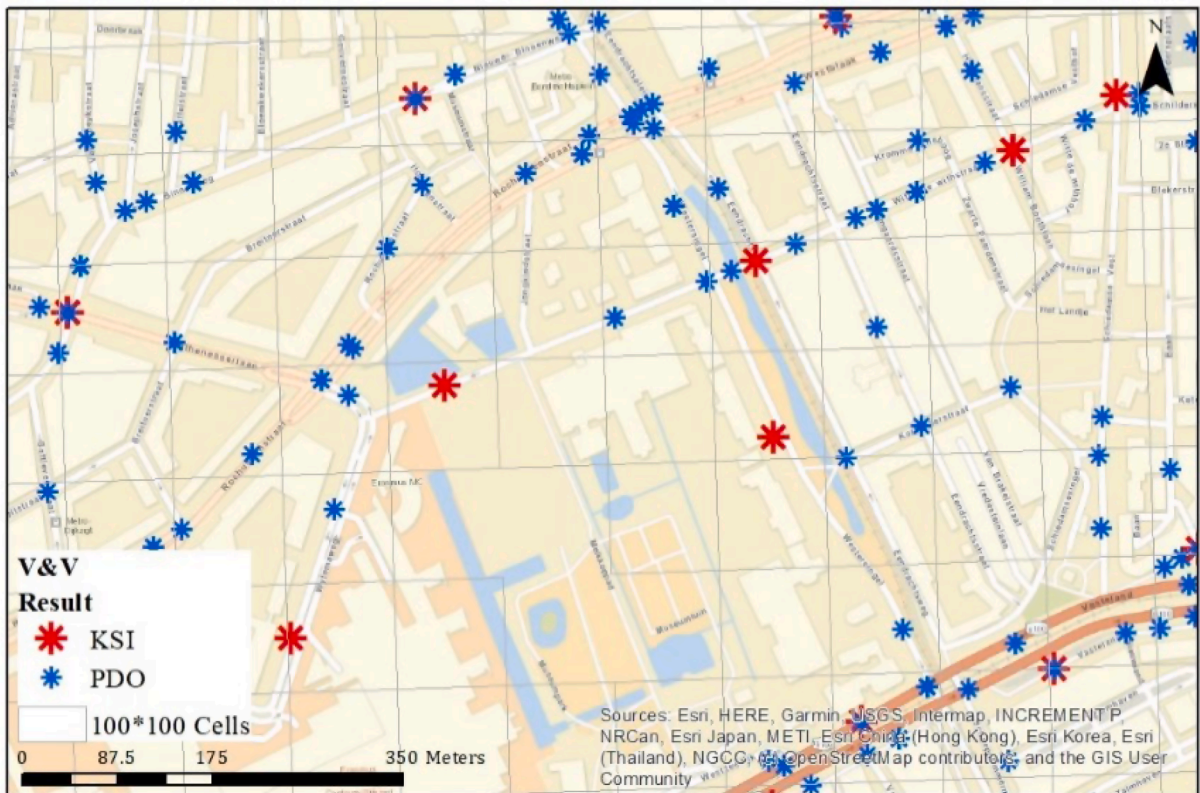


Fig. 3. Representation of the location of the V&V crash points in the 100-cells in the city of Rotterdam.

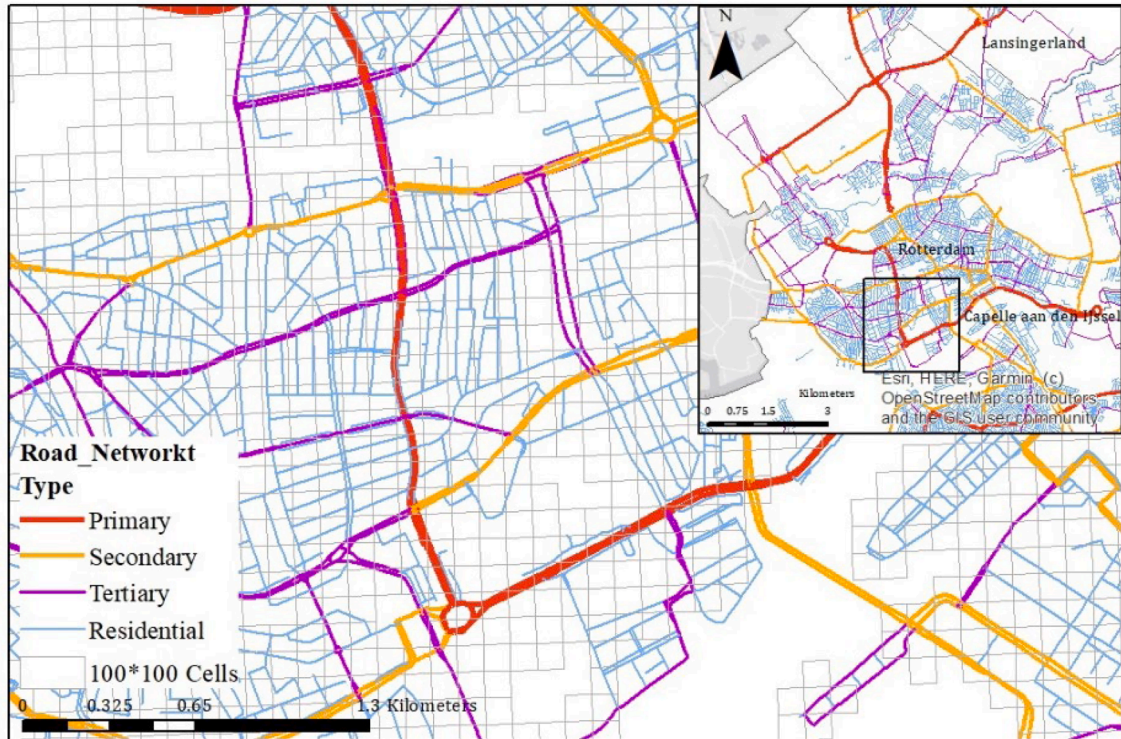


Fig. 4. Representation of the road network in the city of Rotterdam.

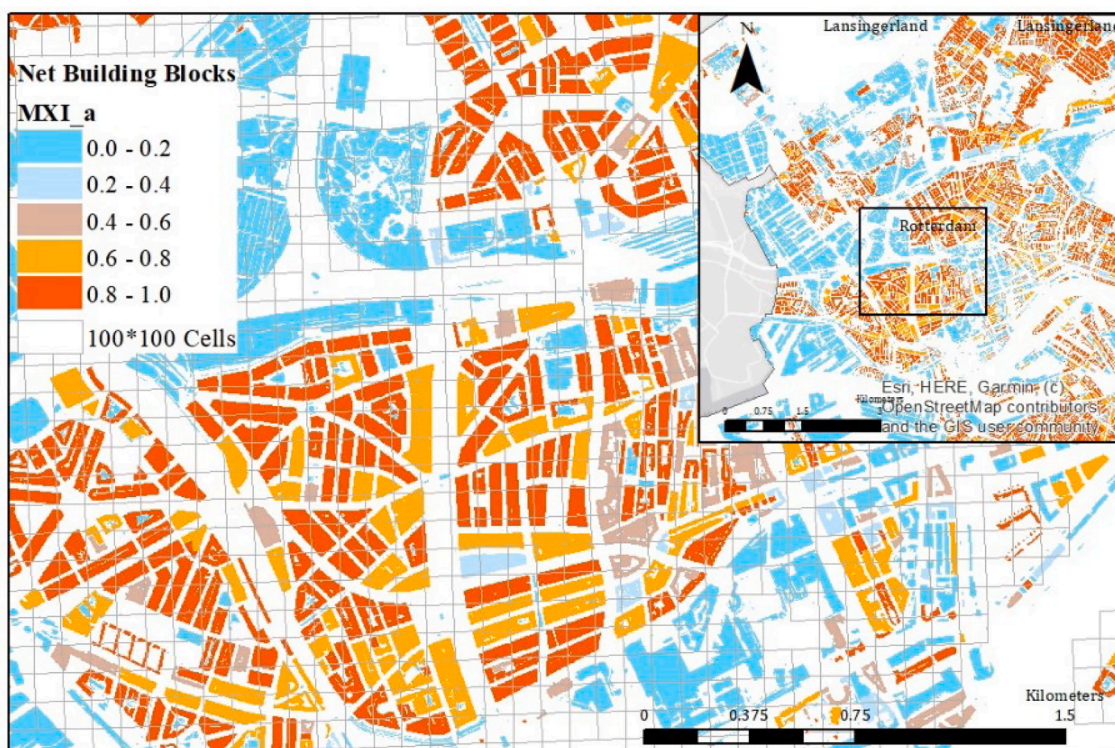


Fig. 5. Representation of land-use function mixing of the building blocks overlaid in the cells, in Rotterdam.

Table 2

Regression models for weekend/weekday vehicle count on major roads in the selected cities.

| n | City | Regression model | Adjusted R ² | Sample size |
|---|---|--|-------------------------|-------------|
| 1 | Amsterdam | $\log(\text{WE}) = -1.21 + 1.08 * \log(\text{WD})$ | 0.92 | 477 |
| 2 | Rotterdam | $\log(\text{WE}) = -1.26 + 1.08 * \log(\text{WD})$ | 0.93 | 560 |
| 3 | The Hague | $\log(\text{WE}) = -0.09 + 0.98 * \log(\text{WD})$ | 0.99 | 189 |
| 4 | Utrecht | $\log(\text{WE}) = -1.36 + 1.09 * \log(\text{WD})$ | 0.98 | 349 |
| 5 | Almere | $\log(\text{WE}) = -0.58 + 1.03 * \log(\text{WD})$ | 0.97 | 186 |
| 6 | Amersfoort | $\log(\text{WE}) = -1.09 + 1.08 * \log(\text{WD})$ | 0.99 | 179 |
| 7 | Houten, Nieuwegein | $\log(\text{WE}) = -1.13 + 1.07 * \log(\text{WD})$ | 0.99 | 100 |
| 8 | Zoetermeer, Lansingerland, Capelle, Barendrecht | $\log(\text{WE}) = -0.04 + 0.98 * \log(\text{WD})$ | 0.97 | 151 |

Moreover, Uijtdewilligen et al. (Accepted for publication) found that the ratio between the detector loops and BCW data is equal to 41.12. Thus, we calibrated the daily bicycle counts obtained from BCW by using a factor of 41.12. Similar to the VKT, the total Bicycle-Kilometer-Traveled (BKT) in the cells were calculated after this calibration.

3.2. Distance variables indicating the proximity to facilities

Proximity to facilities was evaluated using variables showing the distance to certain facilities. Information on the distance of each cell to the main roads was of interest in the V&V crash analysis because more motorized vehicles can be expected in the cells with a shorter distance to the main roads. In addition, the average distances to supermarkets and food-product shops were included as previous research in the Netherlands showed that these facilities are one of the main car/bicycle-

trip destinations (Veenstra et al., 2010). The average distance to different types of (secondary) schools and train stations was also included in the V&B crash analysis as cycling is especially prevalent in the Netherlands for trips to education facilities (Mobiliteitsbeleid, 2017b) and access/egress to train stations (Mobiliteitsbeleid, 2017a).

3.3. Crash data

This study uses data obtained from the database with crashes reported by the police (BRON, 2015-2019). Note that crash data commonly suffers from underreporting of minor injuries and PDO crashes and BRON crash dataset is no exception to this (SWOV, 2016). In general, severe crashes involving motor vehicles have a higher registration rate than less severe crashes and crashes which do not involve motor vehicles. Therefore, KSI crashes, defined as crashes with fatalities and/or injuries leading to hospitalization, are analyzed separately.

It is also worth mentioning that there can be inaccuracies in crash locations depending on the record type. In BRON data, crash locations are registered at four accuracy levels: 1) exact coordinates, 2) intersection level, 3) street level, and 4) municipality level (the least accurate level) (Please see (Rijkswaterstaat-CIV, 2021) for more details). In the BRON (2015-2019) data used in this study area, only 2.3% of crashes were registered at the municipality level. However, this part of the data was not discarded, due to its negligible share.

Despite these potential inaccuracies, the BRON dataset is still a high-quality dataset and the only available source of traffic crashes in the Netherlands. Of the final set of 100-cells included in the analysis, one cell had 491 registered V&V-PDO crashes and was discarded after a manual inspection due to erroneous records.

4. Results

4.1. Crash types and descriptive statistics

A total of 7,622 V & B crashes including 5,149 PDO and 2,473 KSI types were analyzed in this study. 26,443 PDO and 2,066 KSI crashes

involving motor-vehicles only (V&V) were included in the analysis. Fig. 6 shows that the majority of cells included in the analysis had zero crashes.

Figure 7 shows that the maximum number of V&B-PDO and V&B-KSI in 2015–2019 are 15 and 9 crashes respectively. For V&V crashes, the maximum number for PDO and KSI crashes are 81 and 8 respectively.

4.2. Modeling results

Spatial Hurdle Negative Binomial models were conducted to analyze the traffic crashes. Tables 3 and 4 present regression coefficients (β), standard error (SE), and the p-value of parameters of the count and zero-hurdle parts of the models as well as the goodness-of-fit (i.e., Log-Likelihood) of the models. In terms of comparing the modeling performances, we should note that the KSI crash models ($AIC_{KSI-V\&B} = 15,366.10$ and $AIC_{KSI-V\&V} = 13,917.21$) had a better performance than the PDO crash models ($AIC_{PDO-V\&B} = 24,888.37$ and $AIC_{PDO-V\&V} = 69,450.99$).

Figures 8 and 9 show the “standardized coefficients” (Siegel, 2016) for significant variables (95% CI). Standardized coefficients were chosen because they are scale-less and unit-less, making them more useful than regular coefficients for comparing the direction and specifically the magnitude of effects (Zhao et al., 2021).

In Figs. 8 and 9, the plain and diagonal shaded blue bars reflect the impacts of the factors on the PDO-crash “frequency” and “probability” respectively. “Frequency” refers to the expected number of crashes in the 100-cells and “probability” refers to the expected chance of observing one or more crashes in the 100-cells within a 5-year period. The orange bars, on the other hand, show the same values for KSI-crashes. Fig. 8 shows that a greater frequency of V&B crashes is expected in the areas with a larger number of old housing (built before 1900). A comparison between the size of the blue and orange bars reveals that the magnitude of this impact was smaller for PDO crashes than KSI ones. The age of the buildings implies policy changes towards land-use and roadway design, rather than the effects of the buildings themselves. That is, the weighted average year of build for residential

properties in the 100-cells can be treated as a proxy variable reflecting the variety in the design of the built-environment, which affects traffic safety.

In Fig. 8, it is also clear that the size of the plain orange bar corresponding to the 30 km/h variable is larger than the same bar for the 50 km/h. This indicates that 50 km/h roads have a larger impact on KSI crashes than 30 km/h roads. Whereas 30 km/h is more influential on PDO crashes than 50 km/h roads. These figures provide a clear illustration of the scale-free effects of variables on crashes and a comparison of these effects. However, the reader should refer to Tables 3 and 4 to see the unstandardized coefficients of these variables.

4.2.1. Effects of built-environment characteristics

The results show that fewer V&B crashes are likely in areas with a higher proportion of households with children (Fig. 8). This result was expected as the speed limit of the access roads in residential areas with housing for families is set at 30 km/h (15 km/h in home zones where the cyclists and pedestrians have priority over cars which makes the areas safer for pedestrians and cyclists). Moreover, relatively many families with children live in the VINEX areas that are designed (and also found) to be safer locations for cyclists.

Three groups of land-use characteristics affect traffic safety, land-use density, land-use mixture, and land-use classes (Xu et al., 2020). Our findings reveal that increasing the level of urbanization (indicated by the density of the registered addresses in the cells) raises the probability and frequency of traffic crashes. This might be due to the higher population and the conflicts caused by an increased rate of trip attractions/generations in these areas. In contrast, a reduced crash probability was expected and found at a higher level of MXI (representing land-use diversity). This relationship was also found in the literature (Chen & Shen, 2016). This relationship can be explained by the fact that locations (i.e., origins and destinations) are closer to each other in areas with larger MXI, which encourages the residents to walk or cycle.

Furthermore, the results show that the ages of the built-up areas used as proxy variables for various land-use designs impact traffic safety. These variables also indicate the implementation of different land-use

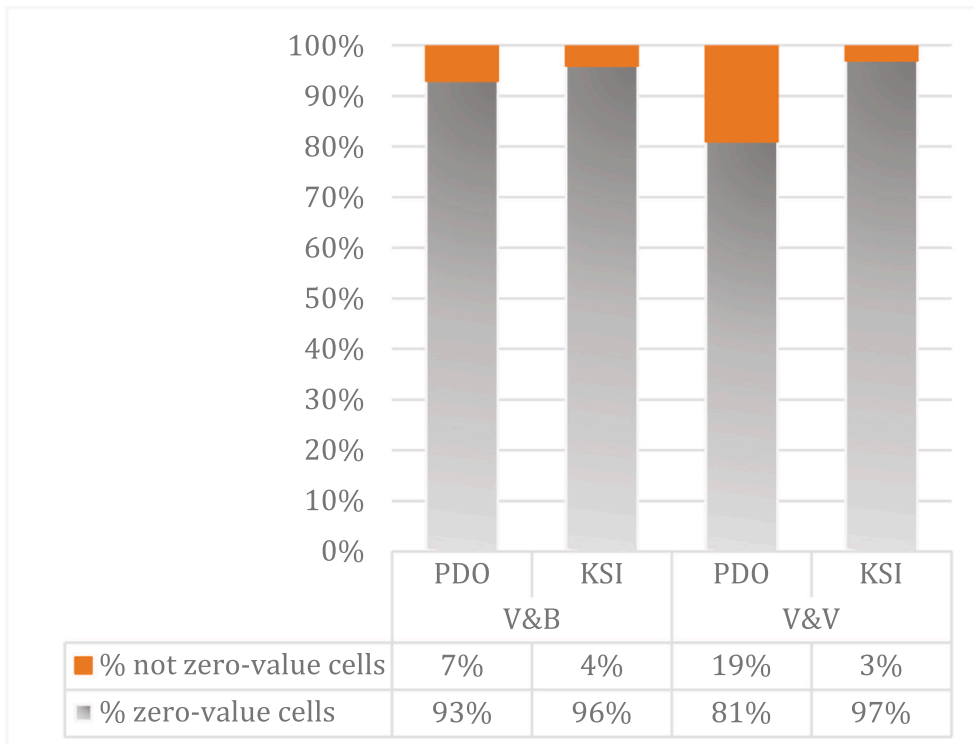


Fig. 6. Comparing the percentage of the cells with zero crashes and not-zero crashes.

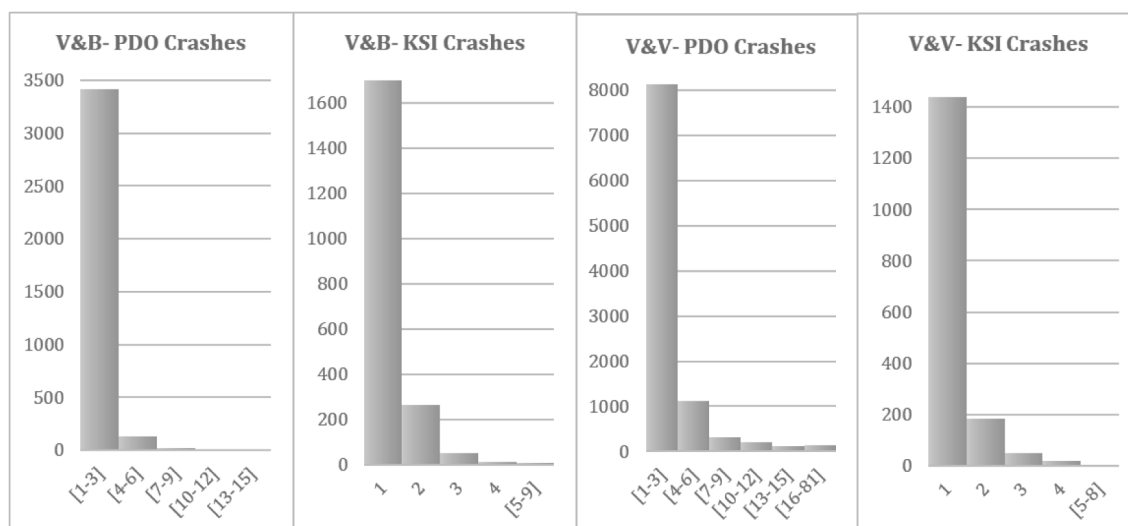


Fig. 7. Histogram graphs of the total number of V&B and V&V types of PDO and KSI crashes, in the cells.

policies which is very intriguing. As the results illustrate, areas built after 1950 are safer for cyclists. Also, the results imply that the older the area (i.e., built between 1900 and 1970), the higher the frequency of V&V-KSI crashes. These findings verify the results of Schepers (2021b) and Schepers et al. (2019) who found that the development of the Dutch new towns (in the 1970 s) and VINEX areas (in the 1990s) has had a positive impact in reducing road deaths.

Higher proximity to grocery stores increases the frequency and/or probability of KSI crashes in the area. In addition, the shorter the distance to the lower-ranked 50 km/h municipal roads (indicated by “distance to the main road” (Leeuwen & Venema, 2021)) the lower the probability of KSI crash occurrence. This result is in line with the findings of Merlin et al. (2020a) who showed that cumulative higher accessibility to traffic volume on arterials and freeways within a 30-minutes travel time thresholds had a positive correlation with crash frequency of all (injury and modes) types. The short distance to the main roads may indicate a better-designed infrastructure using new policies which consider traffic safety. Higher proximity to the main roads also indicates a longer distance to the city center. In these areas, the road network is less dense and modes other than motor-vehicles are less frequent. All these characteristics might lower traffic conflicts and crashes.

Figure 8 also illustrates that high proximity to schools increases the frequency of severe V&B crashes. These findings are consistent with studies in the Netherlands (Schepers, 2021b) and other regions (Hadayeghi et al., 2007) indicating that severe V&B crashes were more probable in the areas proximal to schools. One reason for this might be the tendency of secondary school students to cycle in groups (i.e., bunch cycling). Such behavior has been found to increase the risk of sports cycling crashes in the Netherlands (Wijlhuizen et al., 2016). Also, shorter distances to schools are common in areas with older buildings which were found to be riskier areas for cyclists.

4.2.2. Effects of traffic and infrastructure characteristics

The major role of exposure variables (i.e., VKT and BKT) in increasing the frequency and probability of crashes can be seen in Figs. 8 and 9. These findings broadly support the work of previous studies on traffic crashes (Mukoko & Pulugurtha, 2019; Tagar & Pulugurtha, 2021). Comparisons between the coefficients of BKT and VKT (Fig. 8 for KSI crashes) reflect “the safety-in-number effect”. The ratio of these coefficients ($VKT/BKT = 2.10/1.67$) is fairly compatible with what was found by Elvik and Bjørnskau (2017) (motor vehicle volume/bicycle volume = $0.5/0.43$). However, as Table 3 shows, the actual numerical estimates substantially vary in the Hurdle and count parts of the models.

The primary source of such variation in the safety-in-numbers effect in these models could be the models’ specifications and utilized explanatory variables (Elvik & Goel, 2019). The results also show that the impacts of exposure variables on KSI crashes were considerably larger than their impacts on PDO crashes.

Based on Fig. 8, KSI crashes are more frequent (and probable) in areas with 50 km/h speed limit roads compared to areas with 30 km/h roads. Such results were consistent with the fact that almost 80% of road deaths in Dutch cities occurred on roads with a 50 km/h (or 70 km/h) speed limit (Schepers, 2021a). Nonetheless, Fig. 9 shows that severe V&V crashes are less likely to occur in areas with high-speed (>50 km/h) roads. A similar result was found in the literature (Bao et al., 2021) indicating the possibility of drivers’ violence from the posted speed limits at the lower speeds as well as the complexity of driving conditions because of mixed traffic and more access points at these roads (Tignor & Warren, 1990). In line with previous studies (e.g., Adams & Aldred, 2020; Obelheiro et al., 2020; Schepers, 2021b) we found that a higher proportion of arterial and secondary roads is associated with an increased number/probability of the crashes. Moreover, the figures reflect that a higher density of residential and minor roads in the areas leads to a higher crash frequency in all types.

Interestingly, high cycling speeds were (on average) observed to reduce the frequency of V&B-PDO crashes. This relationship can be reasoned as higher cycling speed is possible on cycling paths where there are no conflicts with motorized traffic. Also, the average cycling speeds should be lower at intersections where more conflicts and consequently more crashes are expected.

Regarding the effects of the cycling facilities, the results confirm what was found in other studies indicating that V&B crashes are concentrated on the arterial roads that are adjacent to the high volume and high-speed roads (Schepers et al., 2013; Schepers et al., 2011). In this study, these facilities are characterized by suggested cycling lanes and other cycling roads (i.e., mixed traffic roads). Moreover, based on Fig. 8 the separated cycle paths may negatively affect the V&B crashes compared to the suggested cycle lanes and mixed traffic roads. This result seems to contradict the findings of Van Petegem et al. (2021) and other previous studies showing that separated cycle paths are usually safer than cycle lanes adjacent to 50 km/h streets and mixed traffic streets (Adams & Aldred, 2020; Aldred et al., 2021). Although the separated cycle paths prevent conflict with motor-vehicles, the presence of busy side roads, parked vehicles besides the cycle paths, as well as large intersections that need to be crossed can increase the probability and frequency of V&B crashes (Vandenbulcke et al., 2014).

Table 3

Results of the SHNBR Model on V&B crashes; Count model coefficients (truncated NB with log link) and Zero hurdle model coefficients (binomial with logit link).

| Groups | Explanatory variables | PDO Crash (Log-likelihood: -12,360 on 85 Df) Theta: count = 0.68 | | | | | | KSI Crash (Log-likelihood: -7,598 on 85 Df) Theta: count = 1.21 | | | | | |
|--|-------------------------|---|------|----------|-------------------|------|----------|--|-------|----------|-------------------|------|----------|
| | | Count Model | | | Zero hurdle model | | | Count Model | | | Zero hurdle model | | |
| | | β | SE | Pr(> z) | β | SE | Pr(> z) | β | SE | Pr(> z) | β | SE | Pr(> z) |
| Population density | (Intercept) | -4.52 | 0.66 | 0.00 | -7.87 | 0.29 | ≈0.00 | -5.31 | 1.04 | 0.00 | -8.4 | 0.36 | ≈0.00 |
| | #Households | 0.12 | 0.09 | 0.16 | -0.07 | 0.06 | 0.21 | -0.02 | 0.14 | 0.89 | -0.16 | 0.07 | 0.03 |
| | αHousehold with Child | -2.03 | 4.67 | 0.66 | 0.84 | 1.75 | 0.63 | -13.49 | 11.06 | 0.22 | -6.08 | 2.98 | 0.04 |
| Land-use density | αAccommodation Area | 0.31 | 0.40 | 0.44 | 1.15 | 0.36 | 0.00 | -0.83 | 1.00 | 0.40 | 0.33 | 0.42 | 0.43 |
| | αSport Area | -2.08 | 2.11 | 0.32 | 0.35 | 0.75 | 0.64 | -4.86 | 4.68 | 0.30 | 0.5 | 0.84 | 0.55 |
| | αShop Area | 0.88 | 0.44 | 0.05 | 0.19 | 0.25 | 0.45 | 0.54 | 0.83 | 0.52 | 0.11 | 0.33 | 0.74 |
| | αMeeting Area | 0.37 | 0.36 | 0.31 | 0.25 | 0.18 | 0.16 | 0.04 | 0.58 | 0.94 | 0.32 | 0.2 | 0.11 |
| | αOffice Area | -0.09 | 0.16 | 0.56 | 0.02 | 0.10 | 0.87 | -0.39 | 0.35 | 0.26 | -0.2 | 0.15 | 0.18 |
| Land-use diversity | Urbanity | 0.10 | 0.05 | 0.06 | 0.29 | 0.03 | ≈0.00 | 0.28 | 0.09 | 0.00 | 0.18 | 0.04 | 0.00 |
| | MXI | -0.03 | 0.12 | 0.83 | -0.37 | 0.10 | 0.00 | -0.10 | 0.28 | 0.72 | -0.14 | 0.12 | 0.26 |
| Land-use Design | Before 1900 | 0.37 | 0.19 | 0.05 | 0.32 | 0.12 | 0.01 | -0.10 | 0.29 | 0.74 | 0.3 | 0.15 | 0.05 |
| | 1900–1950 | 0.03 | 0.17 | 0.84 | 0.05 | 0.10 | 0.61 | -0.31 | 0.26 | 0.24 | 0.22 | 0.12 | 0.07 |
| | 1950–1970 | 0.25 | 0.17 | 0.16 | 0.11 | 0.10 | 0.28 | -0.46 | 0.28 | 0.10 | 0.19 | 0.12 | 0.12 |
| | 1970–1990 | 0.06 | 0.18 | 0.73 | 0.10 | 0.10 | 0.3 | -0.57 | 0.29 | 0.05 | -0.03 | 0.12 | 0.79 |
| | After 1990 | 0.01 | 0.18 | 0.95 | 0.20 | 0.10 | 0.04 | -0.71 | 0.30 | 0.02 | -0.06 | 0.12 | 0.62 |
| Proximity to | Train Station | 0.01 | 0.05 | 0.84 | 0.03 | 0.03 | 0.29 | 0.12 | 0.08 | 0.13 | 0.01 | 0.04 | 0.74 |
| | Secondary Schools | 0.05 | 0.03 | 0.09 | 0.08 | 0.02 | 0.00 | 0.03 | 0.05 | 0.48 | 0.04 | 0.02 | 0.11 |
| | Grocery Shopping | 0.03 | 0.01 | 0.02 | 0.07 | 0.01 | 0.00 | 0.02 | 0.03 | 0.56 | 0.05 | 0.01 | 0.00 |
| Traffic & Infrastructure characteristics | Main Roads | -0.19 | 0.12 | 0.12 | -0.15 | 0.07 | 0.04 | -0.04 | 0.19 | 0.83 | -0.08 | 0.08 | 0.34 |
| | log(VKT) | 0.17 | 0.06 | 0.00 | 0.47 | 0.03 | ≈0.00 | 0.18 | 0.10 | 0.08 | 0.46 | 0.04 | ≈0.00 |
| | Speed Limit ≤ 30 | 0.46 | 0.31 | 0.14 | 0.63 | 0.12 | 0.00 | 1.10 | 0.48 | 0.02 | 0.16 | 0.14 | 0.24 |
| | Speed Limit ≤ 50 | 0.52 | 0.30 | 0.09 | 0.74 | 0.12 | 0.00 | 0.93 | 0.45 | 0.04 | 0.47 | 0.12 | 0.00 |
| | αSecondary Roads | 0.36 | 0.22 | 0.10 | -0.03 | 0.12 | 0.81 | 0.90 | 0.33 | 0.01 | 0.1 | 0.14 | 0.46 |
| | αTertiary Roads | 0.15 | 0.19 | 0.45 | 0.50 | 0.11 | 0.00 | 0.26 | 0.30 | 0.38 | 0.81 | 0.13 | 0.00 |
| | αResidential Roads | -0.22 | 0.17 | 0.19 | 0.64 | 0.09 | 0.00 | 0.06 | 0.31 | 0.83 | 0.65 | 0.11 | 0.00 |
| Cycling Activities & Infrastructure | log(BKT) | 0.25 | 0.05 | 0.00 | 0.20 | 0.02 | ≈0.00 | 0.18 | 0.08 | 0.02 | 0.2 | 0.02 | ≈0.00 |
| | Cycling Speed | -0.01 | 0.01 | 0.17 | -0.01 | 0.00 | 0.09 | -0.03 | 0.02 | 0.03 | -0.01 | 0.01 | 0.14 |
| | αSeparated Cycle Path | 1.87 | 0.47 | 0.00 | 4.58 | 0.28 | ≈0.00 | 1.87 | 0.72 | 0.01 | 4.13 | 0.34 | ≈0.00 |
| | αSuggested Cycle lanes | 1.94 | 0.75 | 0.01 | 6.31 | 0.52 | ≈0.00 | 2.92 | 1.15 | 0.01 | 6.59 | 0.58 | ≈0.00 |
| | αMixed traffic | 0.53 | 0.47 | 0.25 | 3.15 | 0.25 | ≈0.00 | 1.01 | 0.77 | 0.19 | 3.14 | 0.31 | ≈0.00 |
| Socioeconomic characteristics | αRental Houses | 0.48 | 0.21 | 0.02 | 0.51 | 0.10 | 0.00 | 0.46 | 0.32 | 0.14 | 0.4 | 0.13 | 0.00 |
| | Housing Value | -0.91 | 0.43 | 0.03 | -0.13 | 0.20 | 0.52 | -1.84 | 0.69 | 0.01 | -0.26 | 0.24 | 0.28 |
| | αHigh Income | -0.92 | 0.53 | 0.08 | -0.25 | 0.22 | 0.26 | 1.27 | 0.84 | 0.13 | -0.23 | 0.28 | 0.4 |
| Neighboring variables | #Households | -0.21 | 0.12 | 0.08 | -0.09 | 0.04 | 0.05 | -0.63 | 0.21 | 0.00 | 0.01 | 0.09 | 0.88 |
| | αShop Area | 0.25 | 0.50 | 0.61 | 1.32 | 0.32 | 0.00 | -0.41 | 1.13 | 0.72 | 0.70 | 0.41 | 0.09 |
| | αOffice Area | 0.04 | 0.20 | 0.83 | 0.29 | 0.15 | 0.06 | -0.16 | 0.36 | 0.66 | 0.24 | 0.18 | 0.20 |
| | Proximity to Main Roads | -0.04 | 0.15 | 0.79 | 0.12 | 0.08 | 0.16 | 0.22 | 0.21 | 0.29 | -0.01 | 0.10 | 0.95 |
| | αSecondary Roads | -0.72 | 0.69 | 0.29 | 0.62 | 0.38 | 0.10 | -1.38 | 1.14 | 0.22 | -0.12 | 0.47 | 0.80 |
| | αTertiary Roads | 0.45 | 0.69 | 0.51 | 0.87 | 0.39 | 0.03 | 1.63 | 1.10 | 0.14 | 0.28 | 0.5 | 0.58 |
| | αResidential Roads | 0.79 | 0.78 | 0.31 | 0.17 | 0.43 | 0.70 | -0.06 | 1.31 | 0.96 | 0.05 | 0.54 | 0.93 |
| | log(VKT) | 0.03 | 0.09 | 0.74 | -0.09 | 0.04 | 0.05 | 0.12 | 0.14 | 0.39 | -0.01 | 0.05 | 0.93 |
| | log(BKT) | -0.38 | 0.26 | 0.14 | -0.09 | 0.03 | 0.00 | -0.03 | 0.11 | 0.79 | -0.03 | 0.04 | 0.37 |
| | Log(theta) | -0.38 | 0.26 | 0.14 | - | - | - | 0.19 | 0.52 | 0.71 | - | - | - |

4.2.3. Effects of socioeconomic and demographic characteristics

To identify the effects of socioeconomic levels on safety, variables including the percentage of rental properties, average of housing values, and the percentage of the high-income population were used in the model. Figs. 8 and 9 show the effects of these variables indicating that areas with high socioeconomic levels are safer places, confirming previous findings (Najaf et al., 2018; Osama & Sayed, 2017; Xie et al., 2019).

Our results show that severe V&B crashes are less likely in areas with a higher proportion of households with children. However, the associations between the number of households in the area and the analyzed crashes remained unclear in our study. It is also good to note that, we excluded demographic variables such as gender and age as they were highly correlated with the total number of households and the percentage of households with children in the study area.

4.2.4. Effects of characteristics of the neighboring areas (spillover effects)

As for the role of neighboring variables which reflect spatial spillover

effects, the results suggest that an increase in VKT in neighboring cells increases the probability (and to some extent the frequency) of PDO crashes. In contrast, high vehicle/bicycle volumes in neighboring cells were associated with a reduction in the frequency of severe V&B crashes and an increase in PDO-V&V crashes. This is rather counterintuitive but could indicate that the targeted areas are less likely to have high bicycle exposure. It is also worth noting that a higher proportion of shopping areas in the surrounding cells may raise the number of KSI crashes.

5. New insights, discussion, and future directions

5.1. New insights

Previous studies about the effects of built-environment factors on traffic crashes have shown mixed and contradicting outcomes on the magnitude of the effects of the built-environment factors. This study examined the effects of a comprehensive set of built-environment factors on vehicle-only and bicycle-vehicle crashes in the built-up areas. The

Table 4

Results of the SHNBR Model on V&V crashes; Count model coefficients (truncated NB with log link) and Zero hurdle model coefficients (binomial with logit link).

| Groups | Explanatory variables | PDO Crash (Log-likelihood: -34,650 on 71 Df) Theta: count = 0.21 | | | | | | KSI Crash (Log-likelihood: -6,888 on 71 Df) Theta: count = 0.00 | | | | | |
|--|-------------------------|---|-------|----------|-------------------|------|----------|--|---------|----------|-------------------|------|----------|
| | | Count Model | | | Zero hurdle model | | | Count Model | | | Zero hurdle model | | |
| | | β | SE | Pr(> z) | β | SE | Pr(> z) | β | SE | Pr(> z) | β | SE | Pr(> z) |
| Population Density | (Intercept) | -5.74 | 0.26 | ≈0.00 | -7.77 | 0.15 | ≈0.00 | -21.45 | 1172.70 | 0.99 | -9.96 | 0.31 | ≈0.00 |
| | #Households | -0.09 | 0.06 | 0.13 | 0.01 | 0.04 | 0.75 | -0.28 | 0.26 | 0.28 | -0.09 | 0.08 | 0.26 |
| | αHousehold with Child | -1.65 | 1.70 | 0.33 | -2.03 | 1.25 | 0.1 | -10.85 | 10.96 | 0.32 | -1.97 | 2.91 | 0.50 |
| Land-use Density | αAccommodation Area | 0.62 | 0.32 | 0.05 | 0.8 | 0.3 | 0.01 | -5.04 | 4.56 | 0.27 | 0.25 | 0.51 | 0.62 |
| | αSport Area | 0.48 | 1.21 | 0.69 | -0.03 | 0.65 | 0.96 | 5.02 | 5.34 | 0.35 | -0.38 | 1.36 | 0.78 |
| | αShop Area | 0.74 | 0.26 | 0.00 | 0.87 | 0.22 | 0 | -1.71 | 1.50 | 0.25 | 0.51 | 0.35 | 0.15 |
| | αMeeting Area | 0.04 | 0.22 | 0.86 | 0.2 | 0.14 | 0.15 | -1.48 | 1.24 | 0.23 | 0.20 | 0.25 | 0.41 |
| | αOffice Area | 0.03 | 0.11 | 0.82 | -0.24 | 0.09 | 0.01 | 0.50 | 0.43 | 0.24 | -0.23 | 0.16 | 0.16 |
| Land-use Diversity | Urbanity | 0.22 | 0.03 | 0.00 | 0.26 | 0.02 | ≈0.00 | 0.33 | 0.12 | 0.01 | 0.17 | 0.04 | 0.00 |
| | MXI | -0.24 | 0.09 | 0.01 | -0.37 | 0.06 | 0 | 0.05 | 0.33 | 0.88 | -0.23 | 0.16 | 0.16 |
| Land-use Design | Built before 1900 | -0.32 | 0.12 | 0.01 | 0.25 | 0.09 | 0.01 | -0.74 | 0.55 | 0.18 | 0.42 | 0.18 | 0.02 |
| | Built in 1900–1950 | -0.05 | 0.09 | 0.57 | 0.43 | 0.07 | 0 | -0.08 | 0.33 | 0.81 | 0.28 | 0.12 | 0.03 |
| | Built in 1950–1970 | -0.03 | 0.09 | 0.74 | 0.55 | 0.06 | ≈0.00 | -0.17 | 0.34 | 0.60 | 0.30 | 0.12 | 0.02 |
| | Built in 1970–1990 | -0.17 | 0.09 | 0.06 | 0.3 | 0.06 | 0 | -0.12 | 0.33 | 0.72 | 0.18 | 0.12 | 0.16 |
| | Built after 1990 | 0.01 | 0.09 | 0.95 | 0.22 | 0.06 | 0 | 0.30 | 0.31 | 0.33 | 0.11 | 0.12 | 0.34 |
| Proximity to | Grocery Shopping | 0.04 | 0.01 | 0.00 | 0.09 | 0.01 | ≈0.00 | 0.02 | 0.04 | 0.68 | 0.05 | 0.01 | 0 |
| | Main Roads | -0.06 | 0.05 | 0.18 | -0.08 | 0.04 | 0.05 | -0.17 | 0.19 | 0.38 | -0.17 | 0.08 | 0.04 |
| Traffic & Infrastructure characteristics | log(VKT) | 0.48 | 0.03 | ≈0.00 | 0.81 | 0.02 | ≈0.00 | 0.31 | 0.11 | 0.01 | 0.94 | 0.04 | ≈0.00 |
| | Speed Limit = 50 | 0.27 | 0.05 | 0.00 | 0.23 | 0.04 | 0 | -0.11 | 0.25 | 0.67 | 0.43 | 0.08 | 0.00 |
| | Speed Limit > 50 | 0.39 | 0.08 | 0.00 | -0.12 | 0.05 | 0.02 | 0.16 | 0.31 | 0.60 | 0.23 | 0.11 | 0.04 |
| | αPrimary Roads | 0.71 | 0.13 | 0.00 | 0.38 | 0.09 | 0 | 0.69 | 0.44 | 0.12 | 0.40 | 0.15 | 0.01 |
| | αSecondary Roads | 0.77 | 0.10 | 0.00 | 0.45 | 0.07 | 0 | 1.05 | 0.39 | 0.01 | 0.57 | 0.13 | 0.00 |
| | αTertiary Roads | 0.41 | 0.11 | 0.00 | 0.85 | 0.07 | ≈0.00 | 0.55 | 0.43 | 0.20 | 0.64 | 0.14 | 0.00 |
| | αResidential Roads | 0.30 | 0.09 | 0.00 | 0.73 | 0.06 | ≈0.00 | 0.76 | 0.42 | 0.07 | 0.74 | 0.13 | 0.00 |
| Socioeconomic characteristics | αRental Houses | 0.23 | 0.09 | 0.01 | 0.44 | 0.06 | 0 | -0.44 | 0.31 | 0.16 | 0.40 | 0.12 | 0.00 |
| | Housing Value | -1.31 | 0.19 | 0.00 | -0.6 | 0.12 | 0 | -0.63 | 0.73 | 0.39 | -1.49 | 0.29 | 0.00 |
| | αHigh Income | -0.32 | 0.20 | 0.11 | -0.99 | 0.13 | 0 | -0.61 | 0.78 | 0.43 | -0.33 | 0.30 | 0.26 |
| Neighboring characteristics | #Households | -0.14 | 0.07 | 0.06 | 0.18 | 0.05 | 0 | -0.10 | 0.27 | 0.70 | 0.02 | 0.1 | 0.80 |
| | αShop Area | 0.16 | 0.29 | 0.57 | 1.65 | 0.32 | 0 | -0.85 | 1.32 | 0.52 | -0.51 | 0.47 | 0.27 |
| | αOffice Area | 0.17 | 0.15 | 0.23 | 0.77 | 0.13 | 0 | -0.24 | 0.52 | 0.65 | 0.46 | 0.19 | 0.01 |
| | Proximity to Main Roads | -0.01 | 0.06 | 0.89 | 0.11 | 0.05 | 0.04 | 0.21 | 0.22 | 0.33 | 0.05 | 0.10 | 0.58 |
| | αSecondary Roads | -0.12 | 0.35 | 0.73 | 1.52 | 0.24 | 0 | -2.44 | 1.42 | 0.09 | 1.39 | 0.48 | 0.00 |
| | αTertiary Roads | 0.30 | 0.36 | 0.41 | 0.19 | 0.26 | 0.46 | 0.91 | 1.31 | 0.49 | 1.68 | 0.50 | 0.00 |
| | αResidential Roads | 0.83 | 0.42 | 0.05 | 1.06 | 0.28 | 0 | -0.88 | 1.61 | 0.58 | 2.01 | 0.56 | 0.00 |
| Log(VKT) | 0.20 | 0.04 | 0.00 | 0.03 | 0.03 | 0.16 | 0.39 | 0.16 | 0.02 | -0.08 | 0.06 | 0.18 | |
| Log(theta) | -1.57 | 0.14 | ≈0.00 | - | - | - | -15.51 | 1172.70 | 0.99 | - | - | - | |

new insights yielding from the results of this study are as follows:

1. Land use variables are often highly correlated which can result in biased effects on traffic crashes. By including land-use variables indicating the urbanity level, we observed that the proportion of different land-use types, e.g., accommodation, shop, etc., become less influential on the crash frequency and probability. Furthermore, we did not find significant associations between crash frequency (and probability) and the proportion of meeting and sports areas (which can be considered as leisure areas).
2. The effects of building density outweigh the effects of population density. This indicates that the contribution of number and location of human activities (i.e., buildings) and therefore land-use policies in traffic safety is greater than the number of residents in an area.
3. The majority of the built-environment factors have similar impacts in terms of increasing or decreasing V&B and V&V crashes. However, this study gives new insights about the magnitude of these impacts that vary based on the severity level of the crash and also the vulnerability level of the involved parties. For example, a greater level of MXI reduces the number of both types of crashes. In this study, we also observed that this impact on V&B crashes ($|std.coeff_{V\&B}^{MXI}| = 0.45$) is larger than that on V&V crashes ($|std.coeff_{V\&V}^{MXI}| = 0.40$).

4. A greater proportion of residential roads and separated cycling paths increases the probability (and frequency) of V&B crashes in Dutch cities. This result differs from findings of previous studies in other countries. E.g., [Osama and Sayed \(2017\)](#) found that an increase in the proportion of local roads and separate bicycle paths had reduced the number of V&B crashes in Vancouver, Canada. Findings for the Netherlands are likely to differ given the design and density of the cycling facilities and high bicycle volumes in Dutch cities ([Uijtwilligen et al., Accepted for publication](#)).

5.2. Discussion

The newly designed built-up areas, namely New Towns and VINEX areas developed in the 1970s and 1990s respectively, have improved the safety of cyclists in Dutch cities. Nonetheless, it is worth mentioning that another traffic safety policy, the Sustainable Safety Policy, was adopted after the 1990s, focusing on designing safer roadways ([SWOV, 2018](#)) and improving traffic safety ([Weijermars & Wegman, 2011](#)). As mentioned, we found some associations between traffic safety and the proxy variables indicating the implementation of different spatial and transport policies. This reveals that distinguishing the effects of such policies on traffic safety would be an interesting subject for future research. Contrary to cyclist safety, the above-mentioned urban policies might be less effective in reducing the frequency of serious car crashes.

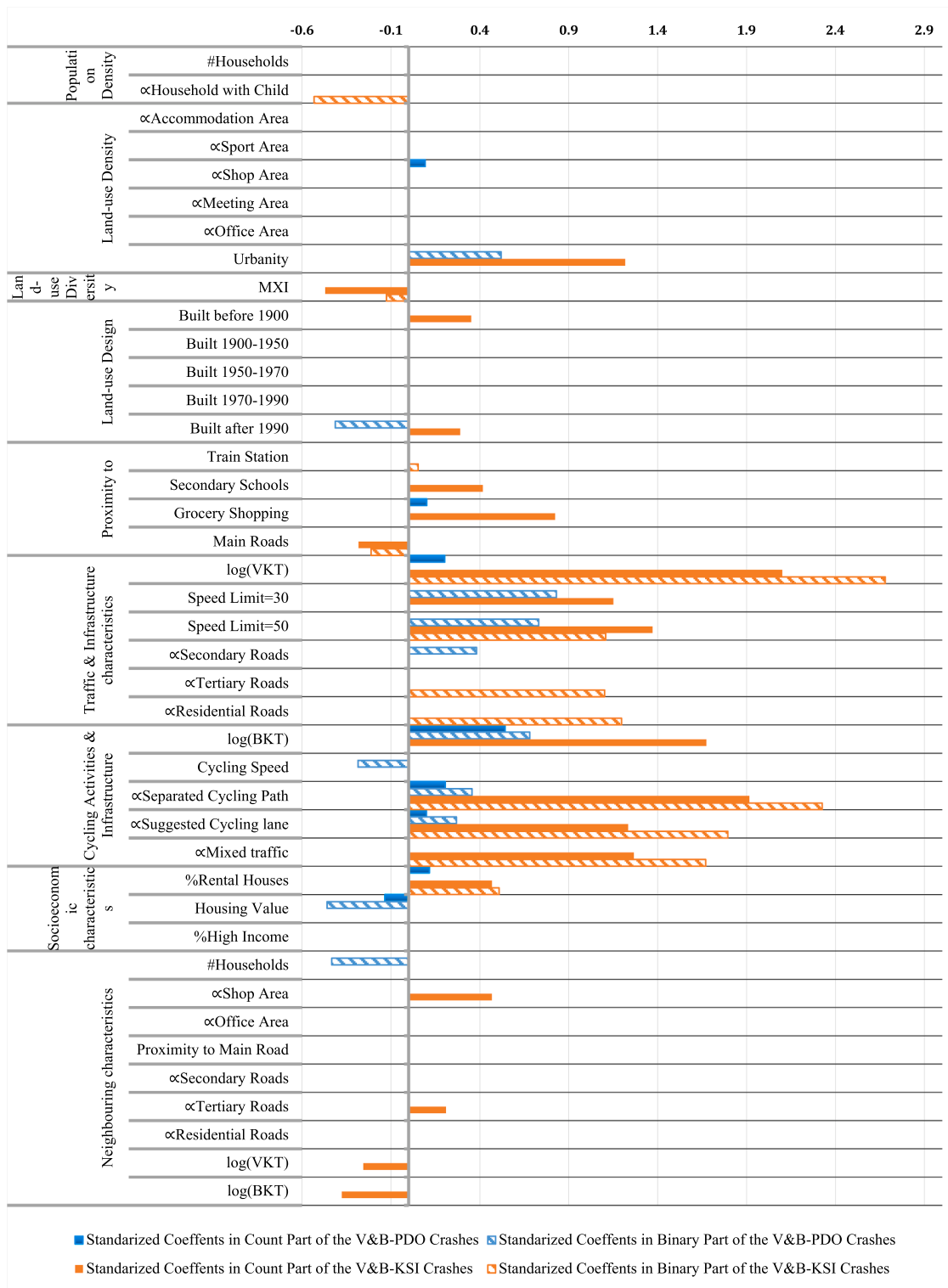


Fig. 8. Comparison between the standardized Count/Zero-Model coefficients of the V&B crashes.

Interestingly, older neighborhoods (e.g., built before 1900) appear to be safer than other areas. This is probably because these areas are usually located in the city centers (i.e., historic places); hence interactions between vehicles and bikes are fewer and less severe.

Newly designed areas were also devised to increase the accessibility of neighborhoods. In practice, accessibility is improved by shortening travel distances to various facilities, such as attraction locations and transport systems facilities. Although better accessibility is always

desirable, our findings revealed that higher proximity to some “attractions” such as grocery stores and schools aggravate traffic safety. Whilst, Merlin et al. (2020a) found a negative significant relationship between increased proximity to job locations from the residential location of the crash victim and vehicle crash frequency (per capita), only within a 10-minutes threshold (i.e., cumulative job-accessibility). They explained such a relationship based on the positive impacts of reduced distances to work on declined VKT. This raises a valid concern that needs to be



Fig. 9. Comparison between standardized Count/Zero-Model coefficients of the V&V crashes.

addressed by future research investigating the reasons behind such a deterioration in safety. Future studies should also investigate the impacts of accessibility indicators on traffic safety (and vice versa) to provide greater insight into the relationships between accessibility and traffic safety.

5.3. Limitations and future directions

This study adopted square grid cells to evaluate the impacts of the built-environment and land-use factors on traffic safety. The potential effects of this choice (i.e., square cells) on the results were not investigated as such an inquiry is beyond the scope of this paper. For example, the hexagon units could be a better alternative as they have the minimum total perimeter length which results in the least boundary effects

(Cui et al., 2021). Future research could focus on identifying the effects of using alternative spatial units such as honeycomb and hexagon cells, building blocks, neighborhoods, or postcode areas.

A limitation in this study was the crash data, which were compiled based on police reports which are known to underreport traffic crashes with minor injuries and to have inaccuracies in reporting the location of crashes. To alleviate this problem, safety officials are currently working on compiling crash data based on ambulance reports (Rijkswaterstraat, 2019b). Future traffic safety research in the Netherlands could benefit from such more accurate data in terms of the severity level of the crashes. This data is planned to become available in the coming years at the national level, based on a national action plan on traffic safety (Rijkswaterstraat, 2019b).

From the methodological perspective, the scope of this study is limited in terms of evaluating the unobserved spatial heterogeneity, as well as the spatial and temporal autocorrelations. Therefore, future research could take advantage of using Spatial Error and Random Parameter modeling approaches as used in previous studies (e.g., (Cui et al., 2021; Huang et al., 2018; Ulak et al., 2018; Wang et al., 2019; Xu et al., 2020; P. Xu & Huang, 2015; Ziakopoulos & Yannis, 2020)). Moreover, future studies may investigate the built-environment and traffic related factors that can contribute to the temporal heterogeneity in traffic crash frequency. Some examples for such time-varying explanatory factors are variation in traffic volumes and operation speeds (R. Yu et al., 2019), as well as variation in the role of commercial and residential areas during the hours of the day and days of a week.

Data processing and methodology approaches introduced in this study can be utilized for analyzing traffic crashes in different cities. However, the findings of this research are based on the data in the Netherlands. Thus, the results of this study cannot be directly generalized to other regions, such as urban areas in the US. Because the land-use and road infrastructure characteristics in the selected study area are different from other areas, especially in terms of the availability and design of the cycling infrastructure. Moreover, differences in these characteristics lead to differences in travel behavior and traffic/cycling volumes.

6. Conclusions

Various built-environment factors contribute to traffic safety in urban areas. However, there is a scarcity of literature on the effects of these factors on the severity and number of traffic crashes in the Netherlands. This study exploits a rich dataset to investigate how and to what extent different built-environment factors are associated with frequency and probability of vehicle-vehicle and bicycle-vehicle crashes. An area-level analysis utilizing $100 \times 100 \text{ m}^2$ cells along with Hurdle NB regression models was adopted to analyze PDO and KSI crash types. To account for the observed spatial correlation effects, the spillover effects of neighboring cells were included in the analysis.

The results of the HNB models show that in addition to road network and traffic characteristics, built-environment factors play a role in the number of crashes (i.e., crash frequency) as well as the chance of occurring a crash (i.e., crash probability), in the analysis units. It is noticeable that impacts of the built-environment factors are more dominant in KSI crashes compared to PDO crashes. Moreover, we observed that newly developed urban areas (built after 1990), that have different spatial designs, are safer for road-users (particularly cyclists). Increased land-use diversity (represented by MXI) is significantly associated with its positive impact on improving traffic safety. Whilst, a higher land-use density (represented by urbanity level) has negative effects on traffic safety. Population density, on the other hand, does not have significant effects although it was expected to have an impact on frequency and/or probability of traffic crashes in the analysis units. The modeling approach also addresses spatial dependency and heterogeneity through spatial spillover effects. Results indicate that traffic and land-use characteristics in surrounding areas (e.g., vehicle/bicycle

volumes, proportion of shopping areas) are associated with traffic safety in the targeted analysis unit.

One of the major contributions of this study is analyzing the collective impacts of a comprehensive set of built-environment variables along with the traffic and infrastructure-related factors. This comprehensive variable set became the cornerstone to examining a cumbersome topic in traffic safety: identifying the relationship between traffic safety and the built-environment.

CRedit authorship contribution statement

Mehrnaz Asadi: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Mehmet Baran Ulak:** Conceptualization, Data curation, Methodology, Supervision, Writing – review & editing. **Karst T. Geurs:** Conceptualization, Data curation, Resources, Supervision, Visualization, Writing – review & editing. **Wendy Weijermars:** Conceptualization, Supervision, Writing – review & editing. **Paul Schepers:** Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Abdel-Aty, M., Lee, J., Siddiqui, C., Choi, K., 2013. Geographical unit based analysis in the context of transportation safety planning. *Transp. Res. Part A: Policy Pract.* 49, 62–75. <https://doi.org/10.1016/j.tra.2013.01.030>.
- Adams, T., Aldred, R. (2020). Cycling Injury Risk in London: Impacts of Road Characteristics and Infrastructure. *Transp. Find.* 18226.
- Aldred, R., Kapousizis, G., Goodman, A., 2021. Association of infrastructure and route environment factors with cycling injury risk at intersection and non-intersection locations: a case-crossover study of Britain. *Int. J. Environ. Res. Public Health* 18 (6), 3060. <https://doi.org/10.3390/ijerph18063060>.
- Anastasopoulos, P.C., Mannering, F.L., Shankar, V.N., Haddock, J.E., 2012. A study of factors affecting highway accident rates using the random-parameters tobit model. *Accid. Anal. Prevent.* 45, 628–633. <https://doi.org/10.1016/j.aap.2011.09.015>.
- Ariannezhad, A., Karimpour, A., Qin, X., Wu, Y.-J., Salmani, Y., 2021. Handling imbalanced data for real-time crash prediction: application of boosting and sampling techniques. *J. Transp. Eng., Part A: Syst.* 147 (3), 04020165.
- Bach, B., van Hal, E., de Jong, M. I., & de Jong, T. M. (2006). *Urban Design and traffic; a selection form Bach's toolbox. Stedenbouw en verkeer; een selectie uit de gereedschapstijl van Bach*: Crow.
- BAG (2020). Retrieved from: <https://data.overheid.nl/en/dataset/basisregistratie-ad-ressen-en-gebouwen-bag>.
- Bao, J., Yang, Z., Zeng, W., Shi, X., 2021. Exploring the spatial impacts of human activities on urban traffic crashes using multi-source big data. *J. Transp. Geogr.* 94, 103118 <https://doi.org/10.1016/j.jtrangeo.2021.103118>.
- BRON. (2015-2019). *Verkeersveiligheid en ongevallencijfers*.
- Cai, Q., Lee, J., Eluru, N., Abdel-Aty, M., 2016. Macro-level pedestrian and bicycle crash analysis: incorporating spatial spillover effects in dual state count models. *Accid. Anal. Prevent.* 93, 14–22. <https://doi.org/10.1016/j.aap.2016.04.018>.
- CBS, 2015. CBS Bestand Bodemgebruik. Retrieved from: <https://www.pdok.nl/download/-/article/cbs-bestand-bodemgebruik>.
- Cbs, 2017. Statistische gegevens per 100 vierkant 2017. Retrieved from: <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/kaart-van-100-meter-bij-100-meter-met-statistieken>.
- Cbs, 2017. Statistische gegevens per 500 vierkant 2017. Retrieved from: <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/kaart-van-500-meter-bij-500-meter-met-statistieken>.
- Chen, P., 2015. Built environment factors in explaining the automobile-involved bicycle crash frequencies: a spatial statistic approach. *Safety Sci.* 79, 336–343. <https://doi.org/10.1016/j.ssci.2015.06.016>.
- Chen, P., Shen, Q., 2016. Built environment effects on cyclist injury severity in automobile-involved bicycle crashes. *Accid. Anal. Prevent.* 86, 239–246. <https://doi.org/10.1016/j.aap.2015.11.002>.
- Chen, T., Lu, Y., Fu, X., Sze, N., Ding, H., 2022. A resampling approach to disaggregate analysis of bus-involved crashes using panel data with excessive zeros. *Accid. Anal. Prevent.* 164, 106496 <https://doi.org/10.1016/j.aap.2021.106496>.
- Chen, T., Sze, N., Chen, S., Labi, S., Zeng, Q., 2021. Analysing the main and interaction effects of commercial vehicle mix and roadway attributes on crash rates using a Bayesian random-parameter Tobit model. *Accid. Anal. Prevent.* 154, 106089 <https://doi.org/10.1016/j.aap.2021.106089>.

- Cui, H., Xie, K.J.A.A., Prevention, 2021. An accelerated hierarchical Bayesian crash frequency model with accommodation of spatiotemporal interactions. *Accid. Anal. Prevent.* 153, 106018 <https://doi.org/10.1016/j.aap.2021.106018>.
- Elvik, R., Bjornskau, T., 2017. Safety-in-numbers: a systematic review and meta-analysis of evidence. *Safety Sci.* 92, 274–282. <https://doi.org/10.1016/j.ssci.2015.07.017>.
- Elvik, R., Goel, R., 2019. Safety-in-numbers: An updated meta-analysis of estimates. *Accid. Anal. Prevent.* 129, 136–147. <https://doi.org/10.1016/j.aap.2019.05.019>.
- Ewing, R., Cervero, R., 2010. Travel and the built environment: a meta-analysis. *J. Am. Plann. Assoc.* 76 (3), 265–294. <https://doi.org/10.1080/01944361003766766>.
- Ewing, R., Dumbaugh, E., 2009. The built environment and traffic safety: a review of empirical evidence. *J. Plann. Literature* 23 (4), 347–367. <https://doi.org/10.1177/0885412209335553>.
- Fietsersbond.* (2016). Retrieved from: <https://stichting.fietsersbond.nl/#onze-projecten>.
- Geofabrik. (2019). OpenStreetMap.
- Gladhill, K., Monsere, C.M., 2012. Exploring traffic safety and urban form in Portland, Oregon. *Transp. Res. Rec.* 2318 (1), 63–74. <https://doi.org/10.3141/2318-08>.
- Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2007. Safety prediction models: proactive tool for safety evaluation in urban transportation planning applications. *Transp. Res. Rec.* 2019 (1), 225–236. <https://doi.org/10.3141/2019-27>.
- Harbers, A., Spoon, M., van Amsterdam, H., van der Schuit, J. (2019). Ruimtelijke Dichtheden en Functiemenging in Nederland (Rudifun). *Planbureau voor de Leefomgeving PBL: The Hague, The Netherlands*.
- He, Y., Chen, H., Liu, C., Jiang, R., Ma, L., 2019. Analysis on the Influencing Factors of Traffic Accident Occurrence Based on Zero Expansion and Hurdle Model. Paper presented at the 2019 4th International Conference on Electromechanical Control Technology and Transportation (ICECTT).
- Hosseinpour, M., Prasetijo, J., Yahaya, A.S., Ghadiri, S.M.R.J.T., 2013. A comparative study of count models: application to pedestrian-vehicle crashes along Malaysia federal roads. *Traffic Injury Prevent* 14 (6), 630–638. <https://doi.org/10.1080/15389588.2012.736649>.
- Huang, Y., Wang, X., Patton, D., 2018. Examining spatial relationships between crashes and the built environment: a geographically weighted regression approach. *J. Transp. Geogr.* 69, 221–233. <https://doi.org/10.1016/j.jtrangeo.2018.04.027>.
- Jackman, S., Tahk, A., Zeileis, A., Maimone, C., Fearon, J., Meers, Z., . . . Imports, M. (2015). Package 'pscl'. *Political Science Computational Laboratory*, 18(04.2017).
- Jedlicka, K., Jezek, J., Kolovsky, F., Kozhukh, D., Martolos, J., Štastný, J., Charvát, K., Hájek, P., Beran, D. (2015). *Open Transport Map*. Retrieved from: <http://opentransportmap.info/>.
- Katrakazas, C., Theofilatos, A., Islam, M.A., Papadimitriou, E., Dimitriou, L., Antoniou, C., 2021. Prediction of rear-end conflict frequency using multiple-location traffic parameters. *Accident Analysis & Prevention* 152. <https://doi.org/10.1016/j.aap.2021.106007>.
- Kim, K., Pant, P., Yamashita, E., 2010. Accidents and accessibility: measuring influences of demographic and land use variables in Honolulu, Hawaii. *Transp. Res. Rec.* 2147 (1), 9–17. <https://doi.org/10.3141/2147-02>.
- Kim, K., Pant, P., Yamashita, E., 2011. Measuring influence of accessibility on accident severity with structural equation modeling. *Transp. Res. Rec.* 2236 (1), 1–10. <https://doi.org/10.3141/2236-01>.
- Kocatepe, A., Ulak, M.B., Ozguven, E.E., Horner, M.W., 2019. Who might be affected by crashes? Identifying areas susceptible to crash injury risk and their major contributing factors. *Transp. A: Transp. Sci.* 15 (2), 1278–1305. <https://doi.org/10.1080/23249935.2019.1581854>.
- Leefomgeving, P.v.d. (2019). *Ruimtelijke dichtheden en functiemenging in Nederland (RUDIFUN)*. Retrieved from: <https://dataportal.pbl.nl/downloads/RUDIFUN1/>.
- Leeuwen, N.V., Venema, J., 2021. Statistische gegevens per vierkant en postcode 2020–2019–2018. Retrieved from: <https://www.cbs.nl/nl-nl/longread/diversen/2021/statistische-gegevens-per-vierkant-en-postcode-2020-2019-2018/4-beschrijving-cijfers>.
- Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transp. Res. Part A: Policy Pract.* 44 (5), 291–305. <https://doi.org/10.1016/j.tra.2010.02.001>.
- Ma, L., Yan, X., Weng, J., 2015. Modeling traffic crash rates of road segments through a lognormal hurdle framework with flexible scale parameter. *J. Adv. Transp.* 49 (8), 928–940. <https://doi.org/10.1002/atr.1322>.
- Merlin, L., Cherry, C., Mohamadi-Hezaveh, A., Dumbaugh, E., 2020. Residential accessibility's relationships with crash rates per capita. *J. Transp. Land Use* 13 (1), 113–128. <https://doi.org/10.5198/jtu.2020.1626>.
- Merlin, L., Guerra, E., Dumbaugh, E., 2020. Crash risk, crash exposure, and the built environment: a conceptual review. *Accid. Anal. Prevent.* 134, 105244 <https://doi.org/10.1016/j.aap.2019.07.020>.
- MinlenW., 2018. National housing agenda (Nationale woonagenda) 2018–2021. Retrieved from: <https://www.rijksoverheid.nl/documenten/publicaties/2018/05/23/nationale-woonagenda-2018-2021>.
- Mobiliteitsbeleid, K.K.v. (2017a). Tussen 2005 en 2016 is het fietsgebruik naar reinstations toegenomen. *Mobiliteitsbeeld 2017*. Retrieved from <https://www.kimnet.nl/mobiliteitsbeeld/mobiliteitsbeeld-2017#personenvervoer-article8>.
- Mobiliteitsbeleid, K.K.v. (2017b). Tussen 2005 en 2016 reizen we steeds vaker en verder met de fiets. *Mobiliteitsbeeld 2017* Retrieved from <https://www.kimnet.nl/mobiliteitsbeeld/mobiliteitsbeeld-2017#personenvervoer-article5>.
- Morris, C., Yang, J.J., 2021. Effectiveness of resampling methods in coping with imbalanced crash data: crash type analysis and predictive modeling. *Accid. Anal. Prevent.* 159, 106240 <https://doi.org/10.1016/j.aap.2021.106240>.
- Mukoko, K.K., Pulugurtha, S.S., 2019. Examining the influence of network, land use, and demographic characteristics to estimate the number of bicycle-vehicle crashes on urban roads. *IATSS Res.* 44 (1), 8–16. <https://doi.org/10.1016/j.iatssr.2019.04.001>.
- Mullahy, J., 1986. Specification and testing of some modified count data models. *J. Econometrics* 33 (3), 341–365. [https://doi.org/10.1016/0304-4076\(86\)90002-3](https://doi.org/10.1016/0304-4076(86)90002-3).
- Najaf, P., Thill, J.-C., Zhang, W., Fields, M.G., 2018. City-level urban form and traffic safety: a structural equation modeling analysis of direct and indirect effects. *J. Transp. Geogr.* 69, 257–270. <https://doi.org/10.1016/j.jtrangeo.2018.05.003>.
- Nationale Fietsstreekweek* (2016). Retrieved from: <http://opendata.cyclingintelligence.eu/>.
- Obelheiro, M.R., da Silva, A.R., Nodari, C.T., Cybis, H.B.B., Lindau, L.A., 2020. A new zone system to analyze the spatial relationships between the built environment and traffic safety. *J. Transp. Geogr.* 84, 102699 <https://doi.org/10.1016/j.jtrangeo.2020.102699>.
- Osama, A., Sayed, T., 2017. Evaluating the impact of socioeconomics, land use, built environment, and road facility on cyclist safety. *Transp. Res. Rec.* 2659 (1), 33–42. <https://doi.org/10.3141/2659-04>.
- Rijkswaterstaat-CIV. (2021). *Handleiding product; Bestand geRegistreerde Ongevallen Nederland (BRON)*. Rijkswaterstaat Retrieved from <https://www.rijkswaterstaat.nl/apps/geoservices/geodata/dmc/bron/Documentatie/Handleiding%20product%20Bestand%20geRegistreerde%20Ongevallen%20Nederland.pdf>.
- Rijkswaterstaat, 2018. Wegvakken. Retrieved from: https://www.rijkswaterstaat.nl/apps/geoservices/geodata/dmc/nwb-wegen/geogegevens/shapefile/Nederland_tot_aal/.
- Rijkswaterstaat, 2019. Intensiteit WEGVakken (INWEVA). Retrieved from: <https://maps.rijkswaterstaat.nl/dataregister/srv/dut/catalog.search#/metadata/f58eacc9-ca69-487a-a53b-11efad0bbbc0>.
- Rijkswaterstaat. (2019b). Landelijk Actieplan Verkeersveiligheid 2019–2021.
- Saha, D., Dumbaugh, E., Merlin, L.A., 2020. A conceptual framework to understand the role of built environment on traffic safety. *J. Safety Res.* 75, 41–50.
- Schepers, P. (2021a). *Ruimtelijke inrichting en verkeersveiligheid* Retrieved from <https://www.swov.nl/publicatie/ruimtelijke-inrichting-en-verkeersveiligheid>.
- Schepers, P., 2021. Ruimtelijke inrichting en verkeersveiligheid in stedelijk gebied. Retrieved from: <https://www.swov.nl/publicatie/ruimtelijke-inrichting-en-verkeersveiligheid-stedelijk-gebied>.
- Schepers, P., Heinen, E., Methorst, R., Wegman, F., 2013. Road safety and bicycle usage impacts of unbundling vehicular and cycle traffic in Dutch urban networks. *Eur. J. Transp. Infrastruct. Res.* 13 (3).
- Schepers, P., Kroeze, P., Sweers, W., Wüst, J., 2011. Road factors and bicycle-motor vehicle crashes at unsignalized priority intersections. *Accid. Anal. Prevent.* 43 (3), 853–861. <https://doi.org/10.1016/j.aap.2010.11.005>.
- Schepers, P., Lovegrove, G., Helbich, M. (2019). Urban form and road safety: Public and active transport enable high levels of road safety *Integrating Human Health into Urban and Transport Planning* (pp. 383–408): Springer.
- Siegel, A., 2016. *Practical Business Statistics*. Academic Press.
- SWOV, 2016. Gegevensbronnen; Uitebreid overzicht. Retrieved from: https://www.swov.nl/sites/default/files/bestanden/wegwijzer/data_sources.pdf.
- SWOV, 2018. *Sustainable Safety 3rd edition – The advanced vision for 2018–2030* (978–90-73946-17-0). Retrieved from: <https://www.swov.nl/en/publication/sustainable-safety-3rd-edition-advanced-vision-2018-2030>.
- SWOV, 2021. Road deaths in the Netherlands. Retrieved from: <https://www.swov.nl/en/facts-figures/factsheet/road-deaths-netherlands>.
- Tagar, S., Pulugurtha, S.S., 2021. Effect of increasing the freeway posted speed limit on entry ramp speed-change lane crash frequency. *Transp. Eng.* 4, 100067 <https://doi.org/10.1016/j.treng.2021.100067>.
- Tignor, S.C., Warren, D., 1990. *Driver Speed Behavior on US Streets and Highways*. Compendium of Technical Papers, Institute of Transportation Engineers, Washington, DC.
- Uijtendewilgen, T., Ulak, M.B., Wijnhuizen, G.J., Bijleveld, F., Dijkstra, A., Geurs, K.T., 2022. How does hourly variation in exposure to cyclists and motorised vehicles affect cyclist safety? A case study from a Dutch cycling capital. *Safety Sci.* 152, 105740.
- Ukkusuri, S., Miranda-Moreno, L.F., Ramadurai, G., Isa-Tavarez, J., 2012. The role of built environment on pedestrian crash frequency. *Safety Sci.* 50 (4), 1141–1151. <https://doi.org/10.1016/j.ssci.2011.09.012>.
- Ulak, M.B., Ozguven, E.E., Vanli, O.A., Dulebenets, M.A., Spainhour, L., 2018. Multivariate random parameter Tobit modeling of crashes involving aging drivers, passengers, bicyclists, and pedestrians: spatiotemporal variations. *Accid. Anal. Prevent.* 121, 1–13. <https://doi.org/10.1016/j.aap.2018.08.031>.
- Van Den Hoek, J.W., 2008. The MXI (Mixed-use Index) as tool for urban planning and analysis. *Corporations and Cities: Envisioning Corporate Real Estate in the Urban Future*.
- Van Petegem, J.H., Schepers, P., Wijnhuizen, G.J. (2021). The safety of physically separated cycle tracks compared to marked cycle lanes and mixed traffic conditions in Amsterdam. *Eur. J. Transp. Infrastruct. Res.* 21(3), 19–37. <https://doi.org/10.18757/ejitr.2021.21.3.5283>.
- Vandenbulcke, G., Thomas, I., Int Panis, L., 2014. Predicting cycling accident risk in Brussels: a spatial case-control approach. *Accid. Anal. Prevent.* 62, 341–357. <https://doi.org/10.1016/j.aap.2013.07.001>.
- Veenstra, S., Thomas, T., Tutert, B., 2010. Trip generation of grocery shopping trips using survey data: A case study in the Netherlands. Paper 10–2484. Paper presented at the 89th Transportation Research Board (TRB) Annual Meeting 2010.
- Wang, H., Chang, S.K.J., De Backer, H., Lauwers, D., De Maeyer, P., 2019. Integrating spatial and temporal approaches for explaining bicycle crashes in high-risk areas in Antwerp (Belgium). *Sustainability* 11 (13), 3746. <https://doi.org/10.3390/su11133746>.
- Wei, F., Lovegrove, G., 2013. An empirical tool to evaluate the safety of cyclists: community based, macro-level collision prediction models using negative binomial regression. *Accid. Anal. Prevent.* 61, 129–137. <https://doi.org/10.1016/j.aap.2012.05.018>.

- Weijermars, W., Wegman, F., 2011. Ten years of sustainable safety in the Netherlands: an assessment. *Transp. Res. Rec.* 2213 (1), 1–8. <https://doi.org/10.3141/2213-01>.
- Wijlhuizen, G.J., Van Gent, P., Stipdonk, H., 2016. Sport cycling crashes among males on public roads, the influence of bunch riding, experience and competitiveness. *Safety* 2 (2), 11. <https://doi.org/10.3390/safety2020011>.
- Xie, B., An, Z., Zheng, Y., Li, Z., 2019. Incorporating transportation safety into land use planning: pre-assessment of land use conversion effects on severe crashes in urban China. *Appl. Geogr.* 103, 1–11. <https://doi.org/10.1016/j.apgeog.2018.12.003>.
- Xu, C., Wang, Y., Ding, W., Liu, P., 2020. Modeling the spatial effects of land-use patterns on traffic safety using geographically weighted Poisson regression. *Netw. Spatial Economics* 20 (4), 1015–1028. <https://doi.org/10.1007/s11067-020-09509-2>.
- Xu, P., Huang, H., 2015. Modeling crash spatial heterogeneity: random parameter versus geographically weighting. *Accid. Anal. Prevent.* 75, 16–25. <https://doi.org/10.1016/j.aap.2014.10.020>.
- Yahaya, M., Jiang, X., Fu, C., Bashir, K., Fan, W., 2019. Enhancing crash injury severity prediction on imbalanced crash data by sampling technique with variable selection. Paper presented at the 2019 IEEE Intelligent Transportation Systems Conference (ITSC).
- Yu, C., Xu, M., 2018. Local variations in the impacts of built environments on traffic safety. *J. Plann. Educ. Res.* 38 (3), 314–328. <https://doi.org/10.1177/0739456X17696035>.
- Yu, R., Wang, Y., Quddus, M., Li, J., 2019. A marginalized random effects hurdle negative binomial model for analyzing refined-scale crash frequency data. *Analyt. Methods Accid. Res.* 22, 100092. <https://doi.org/10.1016/j.amar.2019.100092>.
- Zeileis, A., Kleiber, C., Jackman, S., 2008. Regression models for count data in R. *J. Stat. Software* 27 (8), 1–25. <https://doi.org/10.18637/jss.v027.i08>.
- Zhao, Y., Miyahara, T., Mizuno, K., Ito, D., Han, Y., 2021. Analysis of car driver responses to avoid car-to-cyclist perpendicular collisions based on drive recorder data and driving simulator experiments. *Accid. Anal. Prevent.* 150, 105862. <https://doi.org/10.1016/j.aap.2020.105862>.
- Ziakopoulos, A., Yannis, G., 2020. A review of spatial approaches in road safety. *Accid. Anal. Prevent.* 135, 105323. <https://doi.org/10.1016/j.aap.2019.105323>.