Heliyon 10 (2024) e39459

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Impact of urban road characteristics on vehicle speed: Insights from Brescia, Italy

Stefano Raccagni, Roberto Ventura *, Benedetto Barabino

Department of Civil, Environmental, Architectural Engineering and Mathematics (DICATAM), University of Brescia, Brescia, Italy

ARTICLE INFO

Keywords:
Operating speed
Multiple linear regression models
Urban area
Road infrastructures
Road management

ABSTRACT

In managing road infrastructures, a key benchmark is the 85th percentile of vehicle speeds (V_{85}). While V_{85} can be derived from spot speed samples, these are often lacking on each urban road. Thus, prediction models become valuable tools for examining the relationship between V_{85} and road characteristics. Although various models exist for rural roads, the impact of roadside characteristics and markings on V_{85} in urban road networks has been partially investigated, and the effect of traffic calming measures remains fragmented.

This study aims to address these gaps by applying a methodology that sheds light on the effects of some variables that influence V_{85} on urban roads. Specifically, the methodology selects and segments roads along the urban road network of the municipality of Brescia (Italy) and collects data on both road characteristics and 48,000+ spot speed information. Following data cleansing, it processes these data according to three different multiple regression models to analyse the influence of various predictors on V_{85}. Once the best model is estimated, its performance is evaluated, and the final list of significant predictors is obtained.

The results revealed that V_{85} increases with longer homogeneous segments, greater distance to successive intersections, bituminous conglomerate roads with more lanes, and the presence of trees, visible road markings, and posted speed limits. Conversely, V_{85} decreases in the presence of on-street parking and other obstacles (e.g., walls and road posts), when the density of road intersections and pedestrian crossings increases, when the left crossbar width increases and when the land use crossed is commercial or office, residential or industrial. Nevertheless, no significant effect was found for several traffic calming measures included in the model.

These findings can assist road authorities in verifying road operating conditions and planning infrastructure interventions to reduce speeds, thereby creating a safer urban environment for all users.

1. Introduction

High vehicle speed represents a primary element that impacts the likelihood and severity of crashes [1]. For instance, in 2022, 165, 889 crashes occurred in Italy, most of them (73.4 %) in urban areas. 8.1 % of these crashes were attributed to driving behaviour caused by inappropriate vehicular speed [2]. Furthermore, crashes frequently involve vulnerable road users. To contrast dangerous driving behaviour, urban road managers increasingly adopt a range of solutions, such as optimal speed limits, road section transformation, and

* Corresponding author.
E-mail address: roberto.ventura@unibs.it (R. Ventura).

https://doi.org/10.1016/j.heliyon.2024.e39459
Received 7 March 2024; Received in revised form 7 October 2024; Accepted 15 October 2024
Available online 17 October 2024

Traffic Calming Measures (TCMs) [3–9]. These are some of the potential approaches to achieve safer speeds [10]. As a result, speed management strategies are expected to achieve the road safety targets set at the European level (i.e., zero deaths on the European roads in 2050).

The study of the influence of geometric characteristics and the road environment on speed is a well-known topic. Several models have been proposed, which analysed the effects of predictors such as road axis geometry, cross section geometry, land use on speed (V_{85} in Ref. [11,12]), and average speed [13]. Moreover, most of the literature has focused on the estimation of speeds in rural areas to the detriment of the urban areas [13]. There could be various reasons for this gap, such as less complex driver behaviour [12,14] or managing higher rural speeds with the aim of promoting safety in rural areas [15,16]. Conversely, the urban environment exhibits a high degree of heterogeneity. Many factors (e.g., diverse vehicle categories, land uses neighbouring the road, and the presence of vulnerable users) increase or diminish urban speed and, consequently, the risk of crashes [6,11,17–24]. According to this perspective, the provision of speed prediction models affords competent road management authorities a valuable and practical decision-support tool. In fact, practitioners can plan interventions by focusing on elements most impactful on speed while simultaneously enhancing safety.

A careful review of the literature shows that most of these models estimate operating speed. This specific speed is conventionally used as a benchmark for managing existing road infrastructures. It is defined as the 85th percentile of speed distributions in free-flowing traffic conditions and is labelled as V_{85} [12,24–29]. The computation of V_{85} is straightforward, assuming that speed data are available. However, road authorities often lack speed samples for each road, making it challenging to estimate V_{85} [13]. Further, this underscores the need to develop speed prediction models that are accessible and easily applicable.

This paper aims to assess how geometric road characteristics and road environments affect V_{85}. Specifically, spot speed measurements were taken along a large portion of the urban road network of Brescia (Italy), resulting in 48,000 + values that were helpful in calculating V_{85}. Next, V_{85} was modelled as a random variable with a normal distribution and regressed by multiple linear regression (MLR) models against variables (e.g., parameters, predictors, or attributes) describing the road characteristics and the related environment. Next, the model is refined to enhance the accuracy of estimated values.

The contribution of this paper involves both theory and practice. From the theoretical side, novel predictors are introduced and analysed in this study that previous literature had not yet considered. Special attention is paid to the possible presence of TCMs. Individual variables are used to assess the effect of different measures, such as chicanes, speed bumps, and traffic islands, on V_{85}. From a practical perspective, the proposed model enables the determination of the operating speed for each urban road. Additionally, it identifies the road variables that exert the most influence on driver behaviour, on which the road authority can take action to improve safety. Consequently, the prediction model could benefit public authorities and road managers in revamping and enhancing the infrastructure with the purpose of (i) reducing speed in urban areas, (ii) protecting the most vulnerable users, and (iii) increasing urban road safety.

The rest of the paper is structured as follows. Section 2 reviews various models that estimate operating speed. Section 3 outlines the primary concepts and the methodology applied in the research. Section 4 presents the key findings and analyses them in the context of previous literature. Finally, Section 5 provides conclusions and proposes future developments.

2. Literature review

This section briefly reviews some prediction models for the estimation of V_{85} in an urban environment. The modelling of V_{85} has been studied for approximately thirty years. Most of the studies come from the United States (US), but recently, this modelling has also been addressed by non-US research groups, including Italian ones. All studies have focused on understanding the main factors influencing vehicular speed, collecting data using different tools (e.g., GPS or laser devices), and implementing different analysis methods. They agree that operating speed V_{85} is influenced by several factors, specifically by the road geometry, the interaction of the vehicle with other road users, and the surrounding environment [11,27,30,31]. Research has mainly focused on modelling speed adopted in rural areas [12,14,16,32,[33]]. However, much of it remains to be addressed in the urban context. The reasons for this divergence could be traced back to a much more heterogeneous urban environment and, consequently, to more complex modelling than in the rural case [13].

Table 1 concisely summarises some of these studies and is briefly commented on in what follows for the main points.

According to HCM [40], some authors have focused on a specific road type (urban collector by Ref. [30,29]; residential streets by Ref. [14]), while others have considered a generic urban street [36,37] or a kind of road [11,31]

Most studies have generally analysed segments and portions of roads according to different author definitions. For instance, Martinelli et al. [13] considered the portion of road comprised between two consecutive signalised intersections or roundabouts. Al-Bahr et al. [39] established seven segmentation criteria. The segmentation is necessary because road characteristics change quickly in urban environments, which could make prediction models more accurate.

The literature outlines two types of speed sampling methods: spot speed measurements and continuous measurements. Spot speed measurements are recorded at specific fixed points on the road network, involving the immediate speed assessment of each vehicle passing through a designated section. Various devices, e.g., video recorders or laser scanners, are employed for this purpose. However, the measured speed value may be inaccurate due to improper positioning of the instrument. Moreover, the presence of such equipment could influence driver behaviour [13]. As a result, the measured speed values would not be representative and, therefore, not useable in the modelling process without preprocessing and cleaning. Conversely, continuous speed measurements are carried out by equipping vehicles with mobile devices (e.g., GPS) that can geo-localise and determine the speed of each vehicle [28,37,34]. However, some drawbacks should be noted. Firstly, a substantial number of drivers is required for diverse road assessments under optimal

Table 1
Summary of the considered literature.

Author, year	Country	Road type	Survey locations	Locations number	Observed variable	Survey instrument	Sample size	Model applied	Estimated variable	Significant group of variables	R^2 (or R^2_{adj})
[30]	United States	UC	S	27	SS	Traffic counter (magnetic sensors)	2500/2700 for each sensor	MLR	OS	Cross section geometry; Horizontal and vertical alignment	0.67
[29]	United States	UC	S	27	SS	Traffic counter (magnetic sensors)	2667	SLR	OS	Horizontal and vertical alignment	0.63/0.82/0.80
[25]	United States	SA	C and T	14	SS	Laser and radar gun	N/A	MLR	OS	Horizontal and vertical alignment	0.71/0.83/0.72
[26]	United States	UC/RS	C	27	SS	Traffic counter (magnetic sensors)	2601	MEM	OS	Horizontal and vertical alignment; Cross section geometry	0.96
[27]	United States	SA	C and T	19 ÷ 36	SS	Laser gun	N/A	MLR	OS	Other variables	0.75 (0.71)/ 0.54 (0.53)
[31]	United States	SA/UA/ UC/LR	T	79	SS	Laser gun	35/22/13	MLR	OS	Other variables	0.86/0.41/0.14
[28]	United States	UC/LR	C and T	35	CS	GPS device	>200	MLR	OS	Other variables	0.67
[24]	Italy	UA/UC	S	24	SS	Laser gun and Video recording	>1600	MLR	OS	Cross section geometry	0.97/0.96/0.96
[34]	Scotland	UA	S	107	CS	GPS device	N/A	MLR	OS	Roadside configuration	(0.72)/(0.97)
[14]	Japan	RS	S	85	CS	Radar gun	5359	R with 3SLS	OS and MS	Cross section geometry; Roadside configuration	0.56 (0.55)/ 0.58 (0.57)
[35]	Canada	LR/UA	S	58	SS	NC100 traffic sensors	N/A	OP	DS	Cross section geometry	N/A
[11]	Italy	US/UC	S	7	SS	Laser gun and Video recording	5339	MLR/ MEM	OS	Cross section geometry; Roadside configuration	(0.93)/(0.99)/ (0.99)
[12]	Canada	UA/UC	S	249	SS	Traffic Analyzer	~80000 per site	GLM	OS	Cross section geometry; Roadside configuration; Other variables	0.78/0.84/0.77
[36]	Croatia	UR	S	24	SS	Traffic counter	N/A	SLR	OS	Other variables	0.74
[37]	Ecuador	UR	T	45	CS	Video VBOX Lite	21/67/45/13/ 90/125	MLR	OS	Other variables	(0.94)
[38]	Canada	RS	S	140	SS	Pneumatic tubes	N/A	MLR	OS	Cross section geometry; Roadside configuration	N/A
[39]	Malaysia	UR	S	197	CS	MOM (Test vehicles and GPS device)	11812	MLR	ATS	Cross section geometry; Other variables	Range [0.62–0.97]
[13]	Italy	UC	S	52	SS	Laser traffic counter	11466	MLR	ATS and DSS	Cross section geometry; Roadside configuration	0.81/0.36

As for the symbol '/' reported in the last column, it means that there is more than one model estimated.
Abbreviations given reading by row.
UC = urban collector; S = segments; SS = spot speed; MLR = multiple linear regression; OS = operating speed; SLR= Simple Linear Regression; SA = suburban arterial; C = curve; T = tangents; N/A = not available; RS = Residential street; MEM = mixed effect models; UA = Urban arterial; LR = local road; CS = continuous speed; R with 3SLS = simultaneous regression with a three-stage least square (3SLS) estimator; MS = mean speed; OP = ordered probit; DS = discretized speed; GLM = generalized linear model; UR = urban road; MOM = moving observer method; ATS = average traffic speed; DSS = speed standard deviation;.

conditions ([28] with 35 drivers [37], with 45 drivers). Furthermore, the presence of a GPS device may induce an "unnatural" driving behaviour in the driver.

Various modelling tools were available. The primary modelling techniques have employed MLR models [13,30,31,37]. These models are user-friendly, easy to interpret, can be assessed straightforwardly using basic statistical methods, and have the capability to integrate and analyse numerous variables [41,42]. Furthermore, MLR (or SLR) is the most widely adopted technique for speed prediction (see Table 1, column 'Model applied'). Conversely, other authors have applied more complex models, such as ordered probit (OP) or mixed effect models (MEM) [11,12,14,26,35]. The OP is like the logistic regression model, but it predicts an ordinal variable with three or more ordered categories. Conversely, MEM is a statistical technique that analyses data with a hierarchical or correlated structure, considering both random and fixed variations in the data.

According to previous studies, several variables have a significant effect on explaining speed. These variables can be clustered according to the following main groups: horizontal and vertical alignment (e.g., degree of curvature – DC, horizontal curve radius – R, segment width, and longitudinal slope), cross section geometry (e.g., carriageway width, lane width, median width, pedestrian crossing presence, bus stop density, number of lanes), roadside configuration (e.g., presence of a sidewalk, bike line, presence of curb or on-street parking) and other variables (e.g., posted speed limits, density of access points, density of trees and poles, land crossed, traffic composition).

As for horizontal and vertical alignment, a greater degree of curvature decreased V_{85} [26,30,29], as opposed to R [11,25]. In addition, a greater length of the investigated element increases speed [12–14,37,34,38]. Conversely, a greater longitudinal slope decreases operating speed [30,35].

As for cross section geometry, the carriageway width increases V_{85} [14,24,38]. Unlike Bassani et al. [11], someone has observed that as the number of lanes increases, speed increases as well [13,14,28,35]. As for lane width, its effect is controversial. In fact, it increases the speed in Poe and Mason [30], Wang et al. [28], Bassani and Sacchi [24], Martinelli et al. [13]; conversely, it decreases the speed in Bassani et al. [11]; while it has coefficients with alternating signs in Poe and Mason's [26] model, analysing the velocity at four distinct points on a curve. Moreover, the effect of the median width is significant [12,24]. Finally, unlike Hamad and Sacchi [38], the increase in the density of pedestrian crossings decreases the speed [11–13].

As for roadside configuration, the presence of a sidewalk has a negative effect on vehicle speed because it reduces speed [12,13,28, 38], while it has little positive effect in Dihn and Kubot [14] and Eluru et al. [35]. The variation in results may be attributed to the consideration of local roads at low speeds in the latter two studies as opposed to the former studies, which examined various urban roads. As for bike roads, their presence increases the speed [13,35,38]. The presence of access points or intersections reduced speed, as turns are challenging manoeuvres that necessitate a reduction in speed [12,13,28,30,39,34,38]. Moreover, on-street parking negatively influences speed; it is actually beneficial in reducing speed [11,13,28,34,35,38].

As for other variables, posted speed limits (PSL) have a positive effect on speed: the higher speed limits result in higher actual speeds [11,12,27,24,31,36,35]. As for traffic, Al-Bahr et al. [39] showed that speed decreases with an increase in the Annual Average Daily Traffic (AADT). Moreover, the presence of trees creates a longitudinal visual obstacle, slowing down vehicular speed [12,28,38]. However, the influence of the presence of poles is uncertain [12,28,38], as well as the effect of TCMs on speed [11,39,34]. Finally, the effect of land use on speed produced contrasting results. Wang et al. [28] showed that commercial and residential areas positively influence speed, as opposed to Martinelli et al. [13], whereas close proximity to a school decreases speed [13,38].

Interestingly, in more recent studies, the authors considered more variables related to the surrounding road environment (e.g., land use, density of intersections, presence of trees/poles) rather than the geometric characteristics of the road (e.g., degree of curvature, lane width).

As for the fitting performance, most of the studies show a good degree of fit according to the parameters considered (e.g., R^2 around 0.80).

Undoubtedly, all these studies have contributed to the estimation of speed[1] and provided valuable results for research and practice. However, some gaps still persist.

First, most of the studies considered a limited number of locations for surveys. If a range of [61 ÷ 384][2] number of survey locations is considered, only a relatively low number of studies fall into this range [12,14,25,39,34,38]. This is a potential drawback, as the models may either be overly precise and, thus, possibly overfitted or have difficulty capturing the variability of the independent variables [11,26]. Moreover, a small sample leads to restricted representation, sampling bias, lack of diversity, and reduced statistical robustness. It is also worth noting that while a considerable body of research has focused on modelling the V_{85} in rural settings, the existing literature pertaining to urban contexts may benefit from further development [12,13]. Therefore, this study specifically explores speed in urban areas investigating a large sample of spot speed.

Second, many studies have developed different models depending on the road type (e.g. arterials, collectors, and suburban roads) relying on a typical US-based classification [12,24,35]. Since each country has its own unique road network structure with varying characteristics, it is preferable to evaluate the road network by considering the road classification according to the regulations of each state (in Italy, the main references are MIT, 1992, LL.PP., 1995 and MIT, 2001).

Third, several authors have argued for further studies to assess the effect of new variables on speed prediction and to calibrate models on different networks [13,24,32,38]. This lack was mainly observed for (i) roadside characteristics because no studies

[1] In what follows, the term 'speed' coincides with V_{85}, except when expressly indicated.

[2] Inferential statistics would have indicated a representative statistical sample size of [61 ÷ 384] survey locations, adopting a 95 % confidence level, 20 % standard deviation for the response variable (expressed as a percentage of the population mean) and 2 % ÷ 5 % error range.

considered the presence of some elements (e.g., generic obstacles along the road, pavement condition); (ii) roadside markings (e.g., vertical, and horizontal signs and PSL signal); and (iii) TCM, because the presence of these elements have been considered only in a fragmented manner. As for the latter, Obaidet and Mohammad [34], Bassani et al. [11] and Al-Bahr et al. [39] studied the influence of TCM on speed. They developed different models in which TCMs were predictors. More specifically, Obaidet and Mohammad [34] adopted the number of bumps along segments. Bassani et al. [11] considered a generic Boolean variable TCM, while Al-Bahr et al. [39] considered the density of TCM. However, the results could be enhanced for understanding speed variations in the presence of these measures. In fact, although speed decreases as the number of bumps increases, Obaidet and Mohammad [34] have not reported its significance. Moreover, they did not consider other possible TCMs either. Bassani et al. [11] found a small but insignificant contribution to speed dispersion, while Al-Bahr et al. [39] discovered that TCMs reduce speed but only under certain carriageway conditions. Furthermore, neither author reported on what kind of TCMs they considered. Therefore, from this viewpoint, it may be appropriate to evaluate the influence of TCMs in a comprehensive manner by considering different typologies of measures (e.g., chicanes, speed bumps, life-saving islands) within a speed prediction model. Although predictive modelling has not produced relevant results, TCMs still work as more in-depth studies show ([15,43,44]; Vlakved et al., 2022; [6,7]). These studies focused on how speed varies before and after TCMs, thus evaluating its efficiency. For instance, Berloco et al. [7] found that speed decreases before and after Berlin's speed cushions. Moreover, they show that speed decreases more during the morning than in the evening and when the length of the cushion increases. However, the inclusion of TMCs in the modelling of speed could provide rooms for a more inclusive analysis.

This paper aims to cover all the former gaps.

3. Data and methods

This section provides details on the research context, data collection, data cleansing, and data analysis methods.

3.1. Research context

The research was developed in the urban area of Brescia, an important city in Italy located in the eastern part of the Lombardy region. Brescia has a population of 195,906 inhabitants, an area of 90.35 km^2 and a density of about 2168.54 inhabitants/km^2 [45]. Brescia is also the capital of the homonymous province, with more than 1.2 million inhabitants and many important industrial, commercial, and social areas. These aspects contribute to the daily vehicle traffic [46]. However, high speeds and unsafe road infrastructure increase the risk of crashes, especially in urban areas where there are vulnerable users [47]. In 2022, in the province of Brescia, more than 67 % of all crashes occurred in urban areas [2]. From this viewpoint, the development of a model for estimating operating speed could guide road planning. Acting on the road variables that have the greatest impact on speed could be a solution to reduce speed, increase the safety of the road environment, and protect pedestrians and cyclists (e.g., Ref. [6]).

3.2. Methodological framework

An overall methodological framework is conceived to estimate the speed. It defines the steps for the application and incorporates them into the methodology. Fig. 1 provides an illustration of this framework, which is summarised in what follows.

3.2.1. (Step a) - road selection and segmentation

(Step a) selects roads and performs related segmentation. Specifically, roads should represent different conditions in terms of the geometry of the infrastructure and road environment to include a variety of them for the modelling purpose. Therefore, according to this vision, the Mobility Office of the municipality of Brescia selected the most varied critical roads, especially those presenting geometric characteristics which do not comply with Italian Regulations [48,49]. Specifically, a total of 43 roads were surveyed: 30

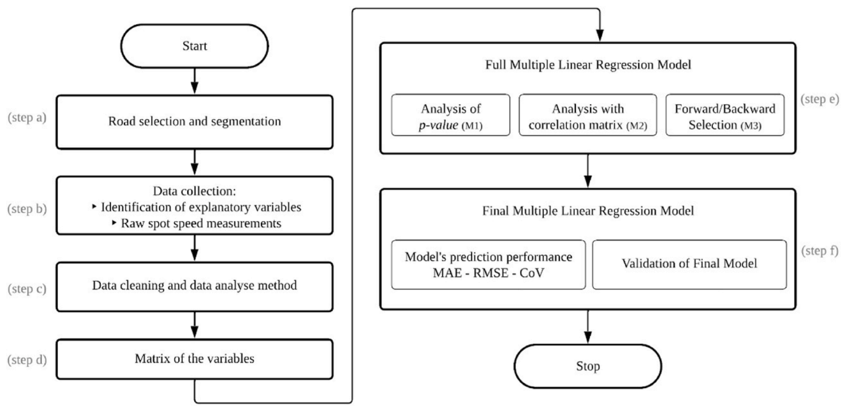

Fig. 1. Flow chart of the methodological framework adopted in this study.

were adopted for calibration, and 13 were preserved for validation tasks. They correspond to a splitting ratio of about 70 % and 30 %, which is usually accepted for modelling [50]. The last value is also higher than the recommended one in Russo et al. [51] and Martinelli et al. [32], who considered a minimum of 10 % of new roads out of the total number used for the calibration.

Four kinds of urban roads were considered, according to MIT [52] and LL.PP. [53].

- '*Strada urbana interquatiere'* (class E*), which enables the penetration of traffic into a specific territorial area, but also the distribution of traffic to the secondary urban road network.
- '*Strada urbana di quartiere*' (class E), which enables traffic to penetrate a specific territorial area.
- '*Strada urbana interzonale*' (class F*), which enables traffic access to specific areas and penetration of local urban roads.
- '*Strada urbana locale*' (class F), which enables traffic access to specific areas.

Next, each road was divided into homogeneous segments according to the following definition: a segment is a portion of the road between two intersections with roads that differ by one level in the class of functionality or have the same class of functionality according to MIT [52] and LL.PP [53]. Selected roads resulted in a total of 130 segments. When traffic was allowed to travel in both directions, the segments were analysed accordingly. Segments were surveyed from April 2023 to June 2023 through a primary analysis of aerial and satellite images and subsequent on-site verification. A total of 240 survey locations (i.e., the section where the spot speed measurement was taken) spread out on about 38 km of the urban road network were identified and surveyed. The selected roads are illustrated in Fig. 2 and presented in Table 2, which is self-explanatory.

3.2.2. (Step b) - data collection

Estimating V_{85} requires the collection of multiple variables concerning road characteristics and spot speed measurements of each segment. Thus, (step b) gathers these variables from various sources.

Spot speed measurements were collected by a laser traffic scanner. This device works by emitting two laser beams perpendicular to the road axis. When a vehicle passes in front of the instrument, the vehicle reflects the laser beams, and the device records its spot speed. The device was placed approximately 1m above the road surface in the midpoint of the selected segment, hidden to avoid any change in driver behaviour due to its presence. The device records the following: date [dd/mm/yyyy] and time [hh:mm:ss]), spot speed [km/h], vehicle length [m], and direction of travel (i.e., ascending or descending) for each passing vehicle. Surveys lasted at least 30 min to gather a representative data sample [32]. Moreover, surveys were conducted during off-peak periods, ensuring that drivers could reach free-flow driving conditions (i.e., dry roads, no congestion, daylight hours and good weather conditions). For these reasons, surveys were carried out on weekdays during daylight hours (08:30 a.m.-05:30 p.m.), excluding 12:30 p.m.-02:30 p.m. and days with severe weather [54–56].

As for the road characteristics, some explanatory variables were identified as possible factors influencing the operating speed (e.g., number of lanes, presence of sidewalk, bus stop density), mining from those that are shown to be significant in the group reported in Table 1 (column 'Significant group of variables'). In addition, further variables were considered in the analysis carried out on the field. The quantification of these factors was made for each segment by digital cartography, on-field measurement, and direct measurement. Table 3 shows the predictors and related statistical parameters (i.e., minimum, maximum, mean, and standard deviation values for each variable).

Although most of the factors are easy to understand and measure, some are concisely described for ease. The distance to a successive intersection (*Int*) was measured from the point where the laser scanner was located (about the midpoint of the homogeneous segment) to the next intersection. The right crossbar width (*Bdx*) and left crossbar width (*Bsx*) were measured as represented in Figs. 3 and 4 for a divided and single carriageway, respectively. The ratio between the total passing flow and road capacity (Q/C) was computed by the ratio between the number of vehicles (excluding bicycles) and the capacity of a specific road according to functional classification [52,53]. The binary variables were identified through digital mapping whereby the presence (1) or absence (0) of each specific element was recorded. Subsequently, field surveys were carried out to confirm the previously obtained values (0 or 1) and to assess the remaining binary variables that could not be easily quantified digitally.

3.2.3. (Step c) - data cleaning and data analysis method

Data should be prepared before modelling. Thus, (step c) makes some preprocessing to clean data; next, some data analysis and modelling are applied to estimate V_{85}. Specifically, spot speed data were downloaded from the laser scanner device, which gathered a total of 48,147 raw spot speed records. However, some cleaning was applied to the raw spot speed samples before computing the operating speed of each homogeneous segment. Because cars represent the most impactful quota of traffic in urban areas, only car speed records were considered. Moreover, these data refer to cars moving in the same travel direction at least 5s after the previous vehicle (headway $\geq$ 5s) as recommended to identify free-flow traffic conditions ([25]; [57]; [13,24,32]). Therefore, by excluding speed records that did not respect these conditions, data cleaning returned 14,651 spot car speed records.

Next, spot car speed records were analysed to compute the observed operating speed (i.e., $V_{85,n}$) for each direction of each homogeneous segment. The procedure is described below. Let:

- N be the set of all surveyed locations and $n \in N$ an individual location.
- $V(n)$ be the set of spot speed records at the location $n \in N$ and $v \in V(n)$ an individual spot speed record.
- Z be the set of ordered speed classes and $z \in Z$ an individual class.

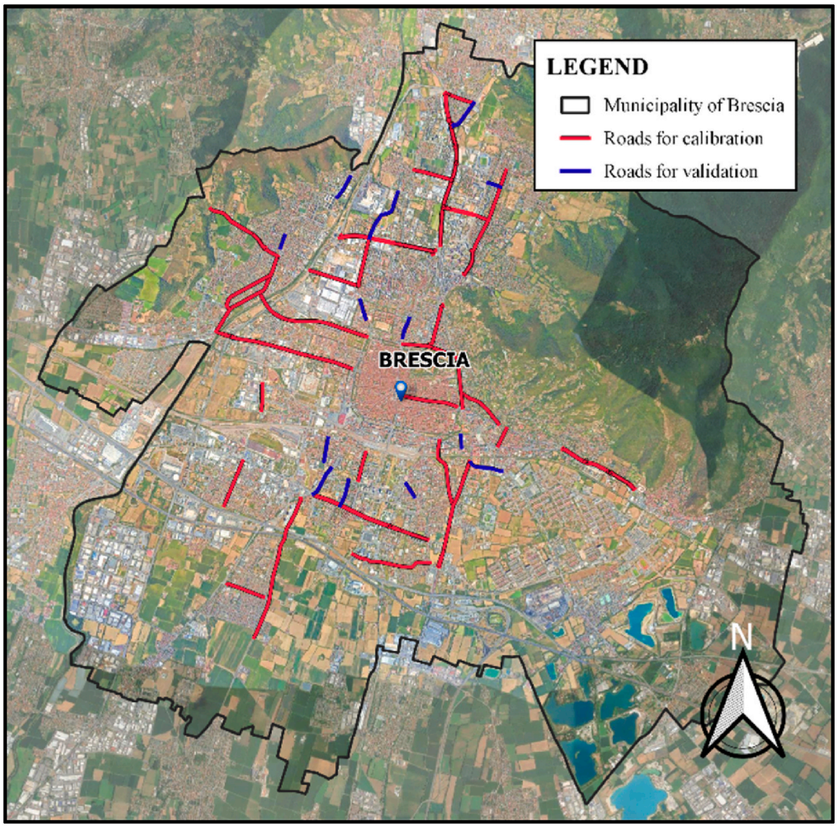

Fig. 2. Surveyed roads considered for the study.

First, the absolute frequency of the observed instantaneous speed (denoted as $F_{z,n}$) is determined as follows:

$$F_{z,n} = \underset{v \in V(n)}{\text{count if}}(v \in z); \quad \forall z \in Z; \ \forall n \in N \tag{1}$$

Second, the relative frequency of the observed instantaneous speed (denoted as $f_{z,n}$) is computed as follows:

$$f_{z,n} = \frac{F_{z,n}}{|V(n)|}; \ \ \forall z \in Z; \ \forall n \in N \tag{2}$$

Third, the relative cumulative frequency (denoted as $\bar{f}_{z,n}$) is calculated as follows:

$$\begin{cases} \bar{f}_{1,n} = f_{1,n} \\ \bar{f}_{z,n} = \bar{f}_{z-1,n} + f_{z,n}; \ \forall z \in Z: \ z \geq 2 \end{cases}; \quad \forall n \in N \tag{3}$$

Fourth, the relative cumulative frequency is plotted on a cartesian plane to depict the cumulative distribution function. The spot speeds are shown along the X-axis, while the Y-axis represents the percentile value of the distribution. The speed value corresponding to the 85th percentile represents the observed operating speed, $V_{85,n}$. Through a graphic analysis of the distribution, the operating speed can be estimated as follows: (1) move vertically along the Y-axis to the 85th percentile; (2) move horizontally until intercepting the curve of the relative cumulative frequencies; (3) from this intersection, move vertically until intercepting the X-axis: the new intersection returns the $V_{85,n}$. An example of it is provided in Fig. 5.

3.2.4. (Step d) - matrix of variables

The road characteristics and $V_{85,n}$ obtained in the previous steps are summarised within the matrix of variables (step d). Specifically, in this matrix, each row represents a survey location, and each column contains the road characteristics at hand as well as the $V_{85,n}$. Each entry of the matrix reports the measured attribute at hand and on the field-measured $V_{85,n}$. Next, this matrix will be used for modelling.

3.2.5. (Step e) - full multiple linear regression model

$V_{85,n}$ constitutes a random continuous variable that follows a normal distribution. Therefore, in (step e), a multiple linear regression model was selected for its inference. It is simple to apply, understand and interpret, and to evaluate with basic statistics (e.g.,

Table 2
Urban road characteristics and survey sections.

Road name	Type of dataset	Road class	Road length	Number of carriageways	Number of lanes per carriageway	Posted Speed Limit[3]	Number of homogeneous segments[4]	Number of survey sections
		[–]	[km]	[#]	[#]	[km/h]	[#]	[#]
Via Ambaraga	C	F*	0.810	1	1	30	4	4[3]
Via Antonio Schivardi	C	E; F*	0.943	1; 2	1,2	50	3	6
Via Attilio Franchi	C	E	0.780	1	1	50	1	2
Via Branze	C	E; F*	0.685	1; 2	1	50	3	6
Viale Caduti del Lavoro	C	F*	0.922	1	1	50	3	6
Via Campane	C	F	0.583	1	1	30	3	6
Via Chiusure	C	E	1.555	1	1	50	5	9
Via Collebeato	V	E*	0.368	1	1	50	1	2
Via Conicchio	V	E*	0.561	1	1	50	1	2
Via Corfù	C	F*	0.446	1	1	50	1	2
Via Corsica	V	E	0.484	1	1	50	1	2
Via Costantino Quaranta	C	F	0.292	1	1	50	3	6
Viale della Bornata	C	E*	1.273	2	2	50	3	6
Via della Chiesa	V	E	0.196	1	1	50	1	2
Via delle Scuole	V	F*	0.885	1	1	50; 30	2	3
Viale della Ziziola	C	F*; F	1.244	1	1	50	3	6
Via divisione tridentina	V	F	0.233	1	1	50	1	2
Viale Duca degli Abruzzi	C	E*	1.690	1; 2	1; 2	50	6	12
Via Famiglia Boccacci	V	E	0.193	1	1	50	1	2
Via F. Carini – Via B. Maggi	C	F	1.060	1	1	50; 30	8	15
Via Filippo Turati	C	E*	0.825	1	1	50	2	4
Via Flero	C	E	2.261	1	1	50	6	12
Via Foro Boario	V	E*	0.497	1	1	50	1	2
Via Galileo Galilei	C	F	0.690	1	1	50	5	10
Via Genova	C	F*; F	0.743	1	1	30	3	5
Via Giuseppe Zola	C	F*	0.397	1	1	30	2	4
Via Lamarmora	C	E*	1.865	2	2; 1	50	6	12
Via Lunga	C	F*	0.399	1	1	50	2	4
Via Milano	C	E*; E	2.229	1	1; 2	50; 30	8	16
Via Nona	C	F*	0.593	1	1	50	3	5
Via Oberdan	C	E*	1.470	2	1	50	3	6
Via Pietro del Monte	V	E	0.296	1	1	50	1	2
Via Pietro Marone	V	F	0.186	1	1	50	1	1
Via Pusterla	C	E*	0.817	1	1; 2	50	2	3
Viale Rebuffone – Via P. Boifava	C	F*	0.760	1	1	30	2	2
Via Rodi	V	F*	0.409	1	1	50	1	2
Via S. Bartolomeo	C	E	0.733	1	1	50	4	8
Via S. Eusacchio	V	E	0.298	1	1	30	1	2

(*continued on next page*)

[3] More than one Posted Speed Limit way be set for some roads owing to different road characteristics.
[4] The number of measured sections is double the number of homogeneous segments because normally a segment can be travelled in both directions. If this is not the case, it means that the segment is only travelled in one direction (one-way).

Table 2 (*continued*)

Road name	Type of dataset	Road class	Road length	Number of carriageways	Number of lanes per carriageway	Posted Speed Limit[3]	Number of homogeneous segments[4]	Number of survey sections
		[−]	[km]	[#]	[#]	[km/h]	[#]	[#]
Via Torricella di Sopra	C	E*; E	1.282	1; 2	1; 2	50	5	10
Via Toscana	V	F*	0.363	1	1	50	1	2
Via Trento	V	E*	0.313	2	2	50	1	1
Via Trieste	C	F*	0.841	1	1	30	5	5
Via Triumplina	C	E*	2.401	2	2	50	6	12
Via Volturno	C	E*	1.890	1	1; 2	50; 30	6	11
Total			38.38				130	240

Abbreviations given reading by row.
C = Road for the calibration dataset; F* = '*Strada urbana interzonale*'; E = '*Strada urbana di quartiere*'; F = '*Strada urbana locale*'; V = Road for the validation dataset; E* = '*Strada urbana interquatiere*'.

Ref. [42]). Let.

- $\tilde{V}_{85,n}$ be the predicted operating speed at $n \in N$.
- K be the set of predictors and $k \in K$ a generic predictor.
- $X_{k,n}$ be the value of the predictor $k \in K$ at location $n \in N$.
- b_k be the value of the regression coefficient associated with the predictor $k \in K$.
- C be the constant of the regression (it represents the intercept of the hyperplane).

The MLR model for estimating the operating speed is formalized as follows:

$$\widetilde{V}_{85,n} = \sum_{k \in K} b_k X_{k,n} + C; \; \forall n \in N \tag{4}$$

The coefficient b_k indicates the relative importance of each predictor in forecasting $\widetilde{V}_{85,n}$. The higher the coefficient in the module, the greater the effect that a marginal variation in the predictor has on $\widetilde{V}_{85,n}$. The sign of b_k is also relevant: a positive value implies that if the predictor at hand, $X_{k,n}$, increases, $\widetilde{V}_{85,n}$ increases as well and vice versa. To determine the coefficient b_k and the constant C of MLR, the ordinary least squares method was utilised by elaborating the records of the calibration dataset. Thus, a full MLR model was set up by employing all matrix's predictors.

Next, for further enhancements, it was still fitted, selecting the best factors that influence the $\tilde{V}_{85,n.}$, trying to achieve a trade-off by a good overall performance and a low number of predictors. Hence, according to the 'principle of parsimony,' it is suggested to include in the model only the minimum number of predictors for explaining $\tilde{V}_{85,n}$. Several methods can be used for this step. In the current study, three refinement techniques were adopted: (M1) analysis of the *p-value* for each variable; (M2) analysis of the correlation matrix; and (M3) application of the forward selection or backward elimination. In approach (M1), predictors with *p-value* greater than 0.1 (i.e., a significance of less than 90 %) were excluded from the set of variables making up the full model. In approach (M2), the correlation matrix was considered to detect if some correlations exist among the factors. A specific predictor is discarded when, compared to another variable, it has a strong correlation (in this case, < -0.7; >0.7) and, at the same time, had a weaker correlation with $V_{85,n}$. In approach (M3), forward selection and backward elimination were applied. Forward selection consists of inserting one variable at a time into the empty model, each time selecting the explanatory variable that best explains the variability of the response variable. Backward elimination is opposed: all variables in the complete model are considered, and through an iterative process, the variable that least explains the variability is removed each time.

3.2.6. *(Step f) - final multiple linear regression model*

The best-fit model is identified in (step f). This is done by comparing the adjusted coefficients of determination (denoted as R^2_{adj}) of the three models. This index indicates the goodness of fit to the observed data, and simultaneously, it enables the comparison of different models in terms of the number of observed data and predictors. Besides R^2_{adj}, R^2, Test-F and its significance, and the residuals analysis are used. This last analysis is based on (i) the predicted zero-sum of residuals (i.e., the random disturbance term across all predictors and observations), (ii) the absence of heteroskedasticity, and (iii) the normal distribution of residuals.

Specifically, let $\varepsilon_n = V_{85,n} - \widetilde{V}_{85,n}$ be the residual at location $n \in N$, and the predicted sum is calculated as follows:

$$\sum_{n \in N} \varepsilon_n \cong 0 \tag{5}$$

The prediction performance of the final model was further evaluated by computing the mean absolute deviation (MAE) and the root mean squared error (RMSE). These are negatively oriented scores: the lower the values, the better the model. In addition, the

Table 3
Summary characteristics for each factor (and sub-factors).

Response Variable	Symbol	Unit of measure	Data type	Data source	Mean	Minimum	Maximum	Standard Deviation
The 85th percentile of the distribution of the free-flow operating speed	*V_85m*	[km/h]	Continuous	DM	50.13	21	75	10.25
Explanatory Variable	**Symbol**	**Unit of measure**	**Data type**	**Data source**	**Mean**	**Minimum**	**Maximum**	**Standard Deviation**
Road Functional Class								
'Strada urbana interquartiere'	*E**	[−]	Categorical	DC	0.3	0	1	0.46
'Strada urbana quartiere'	*E*	[−]	Categorical	DC	0.3	0	1	0.46
'Strada urbana interzonale'	*F**	[−]	Categorical	DC	0.21	0	1	0.41
+*'Strada urbana locale'*+	*F*	[−]	Categorical	DC	0.2	0	1	0.4
Horizontal and vertical alignment								
Homogeneous segments length	*L*	[m]	Continuous	DC	284.5	52	780	161.7
Slope	*i*	[%]	Continuous	DC	−0.04	−4.6	5.1	1.18
Distance to successive intersection	*Int*	[m]	Continuous	DM	123	2.9	798.4	111.4
Cross section								
Number of carriageways	*Ncar*	[n°]	Discrete	DC	1.23	1	2	0.42
Number of lanes	*Nc*	[n°]	Discrete	DC	1.21	1	4	0.44
Lane width	*Lc*	[m]	Continuous	DC	3.26	−0.3	6.9	0.7
Divided median	*Cd*	[−]	Binary	DC	0.19	0	1	0.39
Painted or raised median	*Cc*	[−]	Binary	DC	0.13	0	1	0.33
Right crossbar width	*Bdx*	[m]	Continuous	DC	3.99	0.25	18	2.49
Left crossbar width	*Bsx*	[m]	Continuous	DC	3.36	0	18	2.21
Bus stop density	*Autk*	[stop/km]	Continuous	DC	2.34	0	12.99	2.73
Pedestrian crossing density	*Datt*	[crossing/km]	Continuous	DC	10.81	0	39.47	7.09
Roadside								
Presence of trees	*Al*	[−]	Binary	DC	0.46	0	1	0.5
Presence of other obstacles (e.g., presence of wall, road posts)	*Ost*	[−]	Binary	DC	0.69	0	1	0.46
Average distance to obstacles	*Dt*	[m]	Continuous	OFM	1.34	0	6.5	1.08
Presence of on-street parking	*Cp*	[−]	Binary	DC	0.36	0	1	0.48
Access point density	*Dacc*	[access/km]	Continuous	DC	25.37	0	89.34	21.87
Density of intersections	*Dint*	[intersection/km]	Continuous	DC	9.25	0	57.69	7.69
Presence sidewalk	*M*	[−]	Binary	DC	0.66	0	1	0.47
Presence of a protected cycle route	*Cpr*	[−]	Binary	DC	0.11	0	1	0.31
Presence of an unprotected cycle route	*Cnpr*	[−]	Binary	DC	0.15	0	1	0.36
Presence of a bus/taxi fast lane	*Caut*	[−]	Binary	DC	0.01	0	1	0.07
Presence of curb	*Cu*	[−]	Binary	DC	0.76	0	1	0.43
Presence of a guardrail	*G*	[−]	Binary	DC	0.06	0	1	0.23
Number traffic lights	*S*	[#]	Discrete	DC	0.5	0	4	0.81
Type of road pavement								
Bituminous conglomerate	*CB*	[−]	Binary	DC	0.97	0	1	0.17
Stone paving	*PL*	[−]	Binary	DC	0.03	0	1	0.17
Condition of road pavement								
Good	*PCb*	[−]	Categorical	OFM	0.82	0	1	0.39
+*Medium*+	*PCm*	[−]	Categorical	OFM	0.18	0	1	0.39
Road Marking and Sign								
Visible road markings	*O*	[−]	Binary	OFM	0.87	0	1	0.33
Presence of vertical signs	*V*	[−]	Binary	OFM	0.39	0	1	0.49
Presence of posted speed limit sign	*Ppsl*	[−]	Binary	OFM	0.14	0	1	0.34
Posted speed limit value	*PSL*	[km/h]	Continuous	OFM	45.5	30	50	8.37
Traffic								
Ratio between the total passing flow and road capacity	*Q/C*	[−]	Continuous	DM	0.39	0.03	1.79	0.23
Land crossed								
Commercial or office	*As*	[−]	Binary	OFM	0.61	0	1	0.49
Residential	*Ar*	[−]	Binary	OFM	0.65	0	1	0.48
Industrial	*Ap*	[−]	Binary	OFM	0.1	0	1	0.3
School	*Scu*	[−]	Binary	OFM	0.09	0	1	0.28
Traffic Calming Measures								
Artificial bumps density	*DD*	[bumps/km]	Continuous	DM	0.59	0	19.61	2.43
Presence of chicane	*Ch*	[−]	Binary	OFM	0.04	0	1	0.18
Raised crossing	*RI*	[−]	Binary	OFM	0.29	0	1	0.45
Traffic island	*Is*	[−]	Binary	OFM	0.19	0	1	0.39

Symbols "[−]" means that the variable is without a unit of measure; [#] means per number.
DC = Digital Cartography; DM = Direct Measure; OFM= On the field measurement.
+ Control variables of categorical variables edited in italics +.

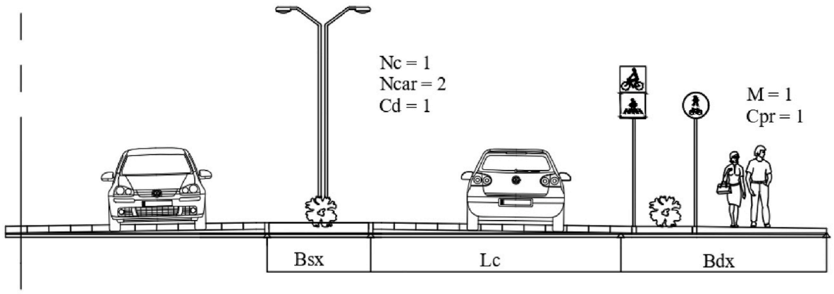

Fig. 3. Example of divided carriageway.

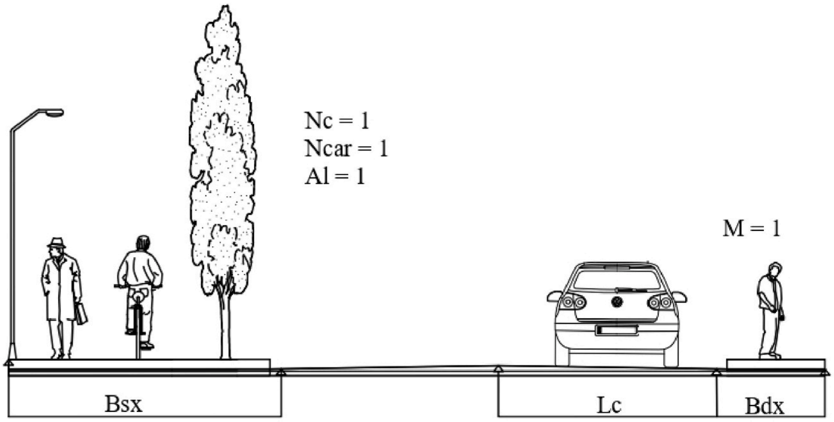

Fig. 4. Example of a single carriageway.

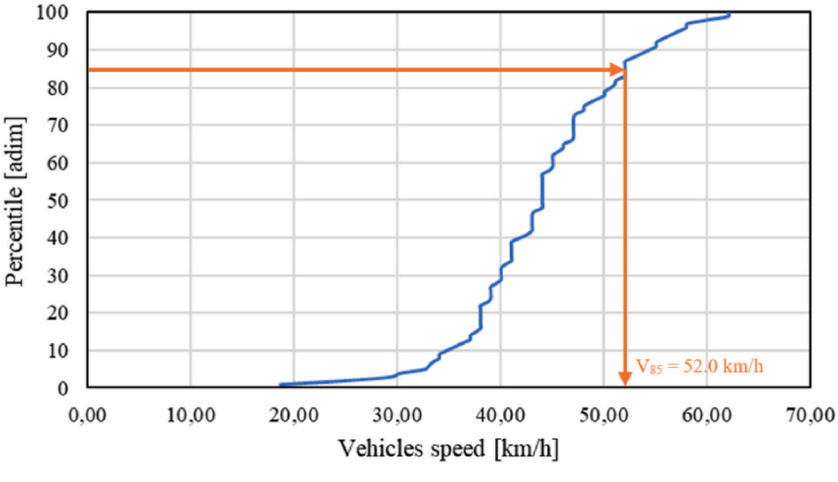

Fig. 5. *Example of V_{85} computation.*

coefficient of variation (CoV) was computed. These statistical indicators were estimated as follows:

$$MAE = \frac{1}{|N|}\sum_{n\in N}|\varepsilon_n| \tag{6}$$

$$RMSE = \sqrt{\frac{\sum_{n\in N}\varepsilon_n^2}{|N|}} \tag{7}$$

$$CoV = \frac{RMSE}{\frac{\sum_{n \in N} V_{85,n}}{|N|}} \tag{8}$$

Finally, to validate the final model, $\tilde{V}_{85,n}$ was compared with the observed $V_{85,n}$ derived from a sample of spot speed records measured along the homogeneous segments of the validation dataset. While $\tilde{V}_{85,n}$ is estimated by eqn. (4), data validation is carried out by analysing the residuals between speeds within specific acceptability ranges. All procedures were implemented using both the software GenStat and MS Excel on a standard personal computer.

4. Results and discussion

According to (step a) and (step b), thirty roads were divided into 116 homogenous segments, namely 215 survey locations (see Table 1, column 'Type of dataset', code 'C'), and described through several variables. Simultaneously, spot speed measurements were taken for each direction of segments where possible.[5] Then, data were cleaned and analysed according to (step c), completing the variables matrix with 200 records of $V_{85,n}$ as shown in (step d).[6] $V_{85,n}$ was utilised as the response variable, and it was estimated by eqns. (1)–(3). Next, eqn. (4) was applied to the matrix to develop a full model for the urban road network, as reported in (step e). Three distinct techniques were applied separately to provide a trained model.

Approach (M1) involved selecting the most significant variables from the final model (*p-value* <0.1). As for approach (M2), a correlation matrix was developed and applied (Appendix 1), while, as for approach (M3), the forward selection was considered, which yielded the same results as backward elimination. To choose the best-fit model, the R^2_{adj} values of the three refined models were calculated and reported in Table 4, and then they were compared according to (step f). Table 4 shows the number of observations (i.e., the total number of records in the variables matrix used for modelling, excluding records that have a large, standardised residual in their respective fitting phases), the two coefficients of determination, the Standard Error, the *F*-test and its significance, and the number of predictors resulting after the fitting phase.

As Table 4 shows, M3 is the model that best fits the data because it is the most accurate (highest R^2_{adj} and lowest standard error) and quite compact. Therefore, it is analysed in detail in what follows.

Table 5 shows some regression statistics for the Full Model and the Final Model (i.e., M3). Specifically, both the Full Model and the Final Model acceptably fit data, as a large *F* coupled with a small p-value indicates a "good" goodness-of-fit. As expected, the final model is more accurate, for each parameter reported in Regression Statistics column of Table 5. Indeed, it is characterised by a smaller number of variables than the full model (i.e., 21 versus 44), and more of these variables are significant at 0.001 level (i.e., 9 versus 4). As for R^2, the final model explains 87.4 % of operating speed variance by the selected predictors, while R^2_{adj} explains 85.8 % of operating speed variance. Moreover, R^2 and R^2_{adj} are two of the best compared to the results of previous works (Table 6).

As for the Standard Error (Table 5), a satisfactory value is obtained, whereby a smaller value indicates a better precision of the best fit line. Similarly, the F-test demonstrates highly significant results for the chosen model (*p-value* <0.001) and therefore, both models have a good fit, but the final is considered.

Table 7 shows both the full model, set up by all variables, and the final model (i.e., M3). Both models are also provided as equations in Appendix 2 to facilitate understanding. In Table 7, the coefficients (*Estimate*) and significance (*p-value*) of the variables are shown. Specifically, it is worth noting that some of the variables in Table 7 are very significant at 0.001 (they are reported in bold), and others are significant up to 0.1 (they are reported in italics).[7] Therefore, these results demonstrate a relevant significant regression effect. Moreover, Table 7 shows that some of the predictors are not significant.

The analysis of variables with *p-value* > 0.1 (i.e., not significant) in the final model is discussed. Specifically, the final model considers the functional classification of roads [52,53] based on four specific variables (*E**, *E*, *F** and *F*). The non-significance of these variables may be due to the absence of a specific regulation that assigns a class to existing roads built before the year 2001 in Italy. Therefore, since many roads were built before that year, this fact may have resulted in a flawed classification of a road, which provided this unexpected finding. For instance, in Brescia, there are many roads that could be classified differently if a single function is considered. However, it is interesting to note that as the functional class decreases, the speed decreases (from +1.13 to −0.54). This result agrees with experience: for urban roads with a penetration function in an area (i.e., higher hierarchical levels of the network), speeds increase. Conversely, for urban roads with an access function in an area (i.e., lower hierarchical levels of the network), speeds decrease. The pavement condition is not significant. The reason for this outcome could depend on the subjectivity of the measure. It was determined during surveys in a qualitative manner rather than quantitatively through objective measurements, such as the International Roughness Index (IRI) parameter. Unlike Eluru et al. [35], the signs of the coefficients bear little relation to experience: optimal asphalt conditions lead to lower speeds. The presence of a priority lane for buses/taxis (*Caut*) is insignificant. This result might

[5] Speed measurements were not carried out for each direction of travel of each homogeneous segment. This is either due to the lack of security in carrying out the measurement or to the effective visibility of the laser instrument, which adversely affects the gait of the vehicles.

[6] After completing the survey, we calculate the true sampling error: it was estimated as 2.8 %, which is in the admitted range [2% ÷ 5 %].

[7] The estimation of eqn. (4) includes variables that were significant at least at 0.10 confidence levels. This is because albeit <0.001 is even higher than the standard 0.05, it can be argued that enabling variables up to 0.10 is appropriate to get a more complete understanding of the results [19, 22].

Table 4
Characteristics of refined models resulting from approaches M1, M2 and M3.

		P-Value resulting model (M1)	Correlation matrix resulting model (M2)	Forward selection resulting model (M3)
Number of observations	**[n]**	196	191	191
R^2	**[%]**	78.5	88.9	87.4
R^2_{adj}	**[%]**	77.3	85.7	85.8
Standard Error	**[%]**	4.84	3.79	3.70
F test		61.20	28.20	55.80
Significance F		<0.001	<0.001	<0.001
Number of predictors		11+constant	42+constant	21+constant

Table 5
Regression statistics for operating speed regression models.

Regression Statistics	Full Model			Final Model		
	DF	SS	MS	DF	SS	MS
Regression	44	17029.	387.01	21	16051.	764.32
Residual	155	3875.	25.00	169	2315.	13.70
Total	199	20904.	105.04	190	18365.	96.66
R	81.5			87.4		
R^2_{adj}	76.2			85.8		
Standard Error	5.00			3.70		
Observations	200			191		
***F* test**	15.48			55.80		
Significance F	<0.001			<0.001		

Table 6
Comparison of coefficients of determination for existing models on urban roads.'

Author, year	R^2	R^2_{adj}
[30]	0.67	
[29]	0.63/0.82/0.80	
[25]	0.71/0.83/0.72	
[26]	0.96	
[27]	0.75/0.54	0.71/0.53
[31]	0.86/0.41/0.14	
[28]	0.67	
[24]		0.97/0.96/0.96
[34]		0.72/0.97
[14]	0.56/0.58	0.55/0.57
[11]		0.93/0.99/0.99
[12]	0.78/0.84/0.77	
[36]	0.74	
[37]		0.94
[39]	0.62 ÷ 0.97	
Our study *Full/Final Model*	*0.82/0.87*	*0.76/0.86*

Some authors provided more models. Results are separated by '/'

depend on the few observations reported in the matrix of variables (step d): only 3 segments have this lane. Thus, more work is required to show its effect, if any, on the operational speed. Finally, TCMs do not even appear in the final model. The literature revealed the possibility of introducing different moderating factors in the speed estimation models (e.g., Ref. [11,39,34]); however, the fitting process shown in (step f) excluded these variables from the whole set. This non-significant result may be due to a combination of two conditions. The speed measurements were carried out in the middle of the homogeneous segment, while the variables associated with the TCMs refer to punctual elements (except the chicane), which are not very widespread in the road network and could be located at different points along the homogeneous segment. For a point element, the speed reduction effect should be felt in the vicinity of the moderation element ([15,43,44]; Vlakved et al., 2022; [6,7]). For these reasons, TCMs should be further studied in future research, preferably by selecting more homogeneous segments characterised by the presence of TCMs.

The physical meaning and sign of the significant variables (i.e, *p-value* ≤ 0.1) is discussed next.

As for horizontal and vertical alignment, the results show that a 1m increase in the length of the homogeneous segment (*L*) slightly increases V_{85}, while keeping all other variables constant at their means and confirming previous research ([14,34], for tangent segment; [12,13,37]). This result is realistic because the longer the road segment, the more a driver will tend to go at a faster speed. Similar reasoning can be made for the distance to the next intersection (*Int*): the greater the distance, the greater the V_{85} of vehicles.

Table 7
Results for operating speed regression models.

Explanatory Variable	Symbol	Full Model				Final Model			
		Estimate	p-value	Lower 95 %	Upper 95 %	Estimate	p-value	Lower 95 %	Upper 95 %
Constant	*C*	25.9	0.146	−8.79	60.59	37.03	**<.001**	31.72	42.34
Road Functional Class									
'Strada urbana interquartiere'	*E**	−0.56	0.812	−5.15	4.03	1.13	0.419	−1.59	3.85
'Strada urbana quartiere'	*E*	0.56	0.764	−3.09	4.21	0.72	0.540	−1.57	3.01
'Strada urbana interzonale'	*F**	−0.23	0.884	−3.37	2.91	−0.31	0.773	−2.39	1.77
Horizontal and vertical alignment									
Homogeneous segment length	*L*	0.01600	**<.001**	0.01	0.02	0.01299	**<.001**	0.01	0.02
Slope	*i*	−0.029	0.929	−0.67	0.62				
Distance to successive intersection	*Int*	0.00538	0.194	0.00	0.01	0.00649	*0.031*	0.00	0.01
Cross section									
Number of carriageways	*Ncar*	−3.01	0.186	−7.44	1.42				
Number of lanes	*Nc*	2.49	0.130	−0.72	5.70	2.983	**<.001**	1.25	4.72
Lane width	*Lc*	−0.060	0.930	−1.41	1.29				
Divided median	*Cd*	0.60	0.814	−4.34	5.54				
Painted or raised median	*Cc*	0.74	0.708	−3.12	4.60				
Right crossbar width	*Bdx*	−0.052	0.796	−0.45	0.34				
Left crossbar width	*Bsx*	−0.321	0.106	−0.71	0.07	−0.275	*0.044*	−0.54	−0.01
Bus stop density	*Autk*	0.186	0.285	−0.15	0.53				
Pedestrian crossing density	*Datt*	−0.1445	0.046	−0.29	0.00	−0.238	**<.001**	−0.33	−0.14
Roadside									
Presence of trees	*Al*	3.184	**<.001**	3.18	3.19	1.969	*0.003*	0.70	3.24
Presence of other obstacles	*Ost*	−2.86	0.009	−4.98	−0.74	−3.12	**<.001**	−4.42	−1.82
Average distance to obstacles	*Dt*	0.140	0.770	−0.80	1.08				
Presence of on-street parking	*Cp*	−3.75	0.004	−6.26	−1.24	−4.861	**<.001**	−6.22	−3.50
Access point density	*Dacc*	−0.0379	0.129	−0.09	0.01				
Density of intersections	*Dint*	−0.3796	**<.001**	−0.51	−0.25	−0.1868	**<.001**	−0.28	−0.09
Presence sidewalk	*M*	0.64	0.570	−1.57	2.85				
Presence of protected cycle route	*Cpr*	3.11	0.035	−1.62	4.26				
Presence of unprotected cycle route	*Cnpr*	1.32	0.380	0.25	5.97				
Presence of a bus/taxi fast lane	*Caut*	−6.23	0.243	−16.64	4.18	−6.19	0.107	−13.68	1.30
Presence of curb	*Cu*	0.94	0.468	−1.59	3.47				
Guardrail	*G*	4.19	0.043	0.17	8.21	2.44	*0.078*	−0.25	5.13
Number of traffic light	*S*	−0.406	0.520	−1.64	0.83				
Type of road pavement									
Bituminous conglomerate	*CB*	22.1	0.192	−10.83	55.03	9.29	**<.001**	4.80	13.78
Stone paving	*PL*	15.8	0.363	−18.11	49.71				
Condition of road pavement									
Good	*PCb*	−1.61	0.198	−4.06	0.84	−0.44	0.596	−2.06	1.18
Road Marking and Sign									
Visible road markings	*O*	3.28	0.086	−0.44	7.00	3.78	*0.002*	1.47	6.09
Presence of vertical signs	*V*	0.168	0.854	−1.62	1.95				
Presence of posted speed limit sign	*Ppsl*	2.10	0.115	−0.49	4.69				
Posted speed limit value	*PSL*	0.1442	0.060	0.00	0.29	0.1014	*0.023*	0.02	0.19
Traffic									
Ratio between the total passing flow and road capacity	*Q/C*	−0.73	0.783	−5.90	4.44				
Land crossed									
Commercial or office	*As*	−3.043	**<.001**	−4.17	0.45	−2.478	**<.001**	−3.73	−1.22
Residential	*Ar*	−1.86	0.116	−4.17	0.45	−3.394	**<.001**	−4.93	−1.86
Industrial	*Ap*	−3.17	0.042	−6.19	−0.15	−2.8	*0.009*	−4.86	−0.74
School	*Scu*	−2.03	0.193	−5.07	1.01				
Traffic Calming									
Artificial bumps density	*DD*	−0.173	0.347	−0.53	0.19				
Presence of chicane	*Ch*	−5.16	0.076	−10.82	0.50				
Raised crossing	*RI*	2.02	0.069	−0.14	4.18				
Traffic island	*Is*	−0.59	0.613	−2.88	1.70				

As for cross section, the findings indicate that lane width (*LW*) has no impact on V_{85}. Although many models in the literature have considered the LW and highlighted its significance [12,13,24,30,34], in this case it is likely that other variables exert a more dominant influence, thereby shadowing the significance of lane width. Consistent with the literature, V_{85} increases as the number of lanes (*NL*) increases [13,14,28]. An expected finding is that a 1 crossing/km increase in the density of pedestrian crossings (*Datt*) decreases V_{85}, while keeping all other variables constant at their means, e.g., Thiessen et al. [12]; Hamad and Sacchi [38]; Martinelli et al. [13]. A 1 m increase in the left crossbar width (*Bsx*) decreases V_{85} while keeping all other predictors constant at their means. This is a novel result because this predictor has not been investigated in previous studies.

As for roadside, the contribution made by the presence of trees (*Al*) deserves attention. Unlike Wang et al. [28] and Thiessen et al. [12], who showed that they decrease speed, this research shows an opposite effect on V_{85}, i.e., their presence increases speed. This result might be justified because trees create a visual barrier that could impede the driver's perception of any hazard that would require a reduction in speed. The same reasoning might be applied to the presence of the guardrail (*G*). Although it is a typical rural feature, it helps channel the traffic, and thus, its presence could increase the speed because drivers may be considered protected against the accidental entry of other vehicles or pedestrians. In contrast, the presence of on-street parking (*P*) is beneficial in reducing the V_{85}, as it has already been shown in previous studies [11,13,28,35]: cars slow down to park and to prevent possible crashes with parked vehicles. The presence of other obstacles different from trees (e.g., presence of walls, road posts) along the segment (*Ost*) is beneficial to reducing the speed and is a novel result. Moreover, a 1 intersection/km increase in the density of intersections (*Dint*) decreases V_{85} along the homogeneous segment, while keeping all other variables constant at their means and confirms previous studies [11,28,30].

As for the type of road pavement, the final model only includes roads made of asphalt (CB). An asphalt pavement favours driving speed because it offers more grip, is usually smooth, has few bumps, and generates less vibration. As a result, the driver will have a faster gait. No previous studies take this feature into account.

As for road markings and signs, the presence of highly visible horizontal signage can encourage higher speeds because it provides drivers with clear and immediate information about the road and traffic conditions. This increases drivers' confidence in road safety and reduces the need to slow down to search for directions or interpret unclear signals. Additionally, well-defined signage helps clearly delineate lanes and stopping areas, reducing the risk of conflicts and increasing traffic flow, thus allowing for higher speeds. Moreover, a 1 km/h increase in the speed limit *PSL* slightly increases $V_{85,}$ while keeping all other predictors constant at their means: previous studies were confirmed [11,12,27,24,31,36].

Finally, as for land crossed, all predictors are useful in reducing V_{85}. If the section is in a residential area, the V_{85} is reduced the most, by to 3.4 km/h. In addition, if the road crosses several zones at the same time (e.g., commercial and residential areas), the speed reduction is the sum of the two entries. Compared to the literature, the same results were obtained by Martinelli et al. [13] but in contrast to Wang et al. [28].

Next, some statistical information is reported to detail the developed model. It includes a comparison between the observed and estimated V_{85}, an analysis of residuals, and an analysis of the model's capacity to make predictions.

Fig. 6 shows the $\left(V_{85,n}, \widetilde{V}_{85,n}\right)$ points are spread near the first quadrant bisector. This condition represents the better scenario, which is close to the ideal situation. For the residual analysis, applying eqn. (5) yields a result of -0.04, which is very close to zero.

The absence of heteroskedasticity is demonstrated in Fig. 7. The residuals did not show any scheme, but they form a cloud of points. That means that residuals exhibit a variation of the same amount with higher or lower values of explanatory variables.

Finally, Fig. 8 shows that the frequency distribution of the residual classes (grey dots) overlaps the normal distribution (blue line). Thus, the model's residual dispersal conforms well with the normal distribution curve.

Furthermore, the prediction capacity of the model was assessed by computing the mean absolute deviation (MAE), the root mean squared error (RMSE), and the coefficient of variation (CoV). These statistical parameters were estimated using equations (6)–(8). In this case, the MAE, RMSE and CoV are 2.77 km/h, 3.48 km/h and 0.06, respectively. Therefore, they are contained: the more the values tend to zero, the better the model's prediction performance.

To further validate the newly developed model as a practical tool for estimating and understanding vehicle speeds, the validation procedure was conducted according to (step f). This involved analysing the residuals obtained by comparing the estimated $\widetilde{V}_{85,n}$ with

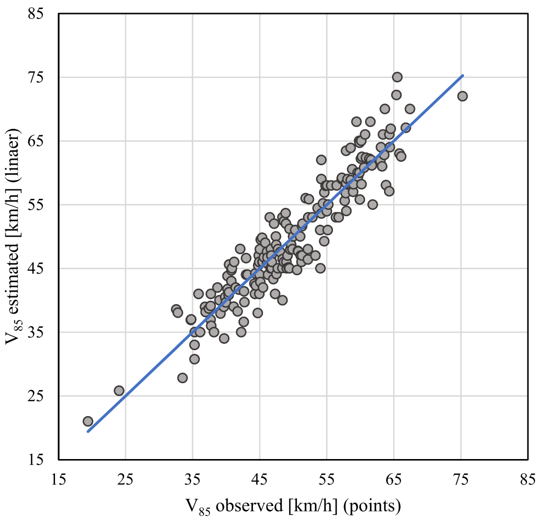

Fig. 6. Scattered plots for performance evaluation of developed model.

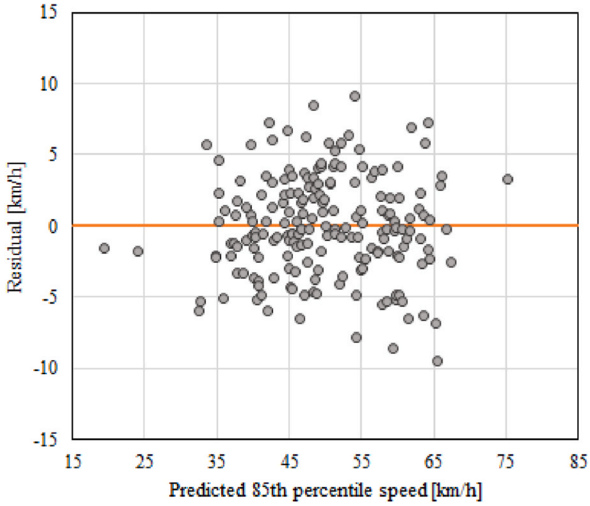

Fig. 7. Residuals distribution.

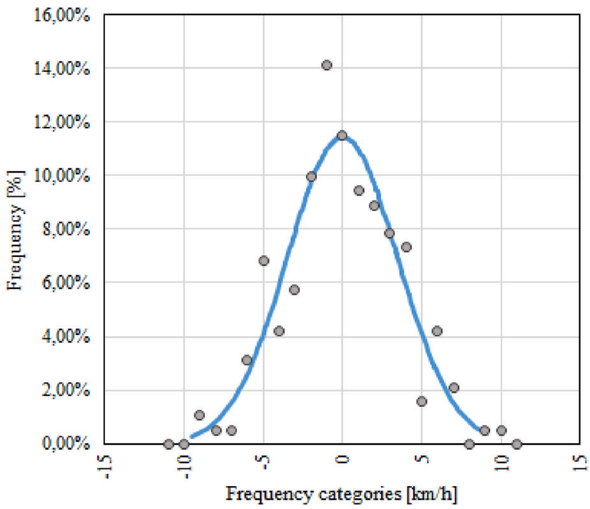

Fig. 8. Frequency distribution of residuals.

the observed V_{85} on additional roads, which were not previously selected for the model calibration. Therefore, several homogeneous segments were extracted from the 13 roads of the validation dataset (see Table 2, column 'Type of dataset', code 'V'). Spot speed measurements and road characteristics were derived as indicated in (step b) and (step c). Spot speeds were used to calculate the observed $V_{85,n}$, while road characteristics were used in the model to derive the estimated $\widetilde{V}_{85,n}$. The two speeds were then compared, and Fig. 9 shows the residuals of V_{85} with three ranges of acceptability based on the empirical rule.

As for Standard Deviation ($\sigma = 4.4$) of residuals, the accepted intervals have as extremes [-σ;+σ], [-2σ;+2σ], and [-3σ;+3σ], that is [−4.4; +4.4], [−8.8; +8.8], and [−13.2; +13.2]. Fig. 9 shows that 16 of the 25 residuals fall within the first standard deviation of the mean [−4.4; +4.4]. Moreover, the results of MAE, RMSE and CoV (3.87 km/h, 4.41 km/h and 0.08, respectively) further confirmed that the developed model is applicable to unsampled road segments within the network at hand.

5. Conclusions and research perspectives

Vehicle speed is a relevant driver to road safety. However, the safety can be compromised by excessive speeds. High speeds significantly increase the probability and severity of crashes, especially those involving vulnerable road users like pedestrians and cyclists, particularly on urban roads. Although speed is a combination of several factors, including inappropriate driver behaviour,

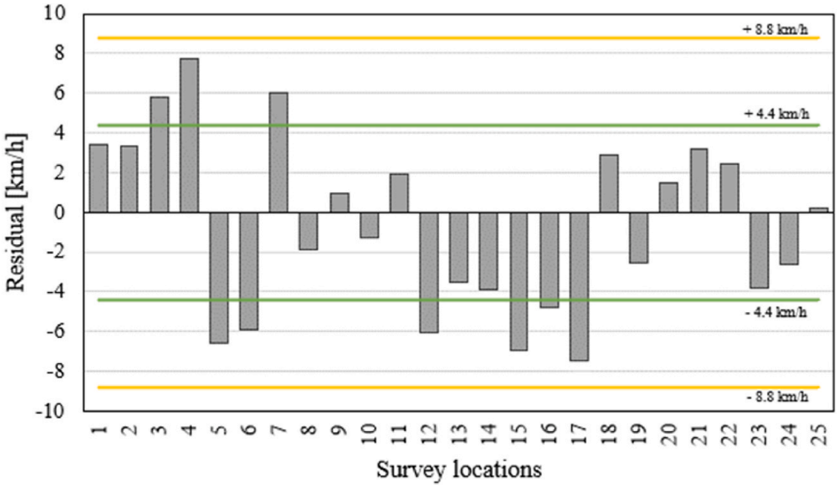

Fig. 9. Validation of final model.

strategic intervention on infrastructures can effectively counteract undesirable speed. For this purpose, it is customary to consider the operating speed (i.e., V_{85}) as benchmark in the management of existing road infrastructure. However, even if the computation of this speed is simple, speed datasets are limited to only the roads sampled. Accordingly, to support design and planning choices by road authorities, the proposition of prediction models can be envisaged where the influence of road characteristics on V_{85} is analysed.

Although the literature provided interesting models for it, this paper contributes to the field in different ways, as follows.

- This study has considered many survey locations to better understand the variability of factors as opposed to much research that has analysed few survey locations. The resulting prediction models provided a good fit of the observed values and returned statistical indicators higher than those returned from the most of literature.
- Unlike other Italian studies, the developed models are applicable to a medium-sized urban area in Northern Italy.
- The study introduced new explanatory variables: left crossbar width, visibility of road markings, presence of obstacles, presence of guardrails, distance to the successive intersection, type of road pavement (all these are significant with *p-value* $\leq$ 0.1) and the functional class of the road (according to MIT [52]), The new variables expanded the results of previous studies because they helped refine the V_{85} estimation.
- This study demonstrated that traffic calming measures did not have a significant effect on operational speed, despite being considered in a comprehensive manner, unlike previous literature.

Some models have been specified, calibrated, and validated using 48,000+ spot speed data on the city of Brescia (Italy). The main results showed that the best model has the advantage of being simple to apply and is a suitable fit for observed operating speed values. Moreover, V_{85} increases when the homogeneous segment length, the distance to successive intersection, increases and the number of lanes increase. In addition, the presence of trees and guardrails increase V_{85}. Moreover, bituminous conglomerate roads are predictors that increase V_{85} as well as visible road markings and posted speed limit value. Conversely, the V_{85} decreases when the left crossbar width, the density of pedestrian crossing, and the density of intersections increase. Moreover, it decreases if there is on-street parking or other obstacles near the lane, and when the road crosses commercial/office, residential areas, or industrial sites. Therefore, these results could benefit public authorities and road management to revamp and enhance infrastructure with the purpose of, e.g. (i) reducing speed in urban areas, (ii) protecting the more vulnerable users, (iii) increasing road urban safety and (iv) replanning public transport networks [58].

Although these important results were obtained, this study could be enhanced. While it enabled the identification of the variables most influential on the V_{85}, the research was constructed and developed according to a straightforward modelling approach and a limited number of observed operating speed values in the urban network, even if larger than most of the previous literature. Moreover, the estimation of V_{85} requires the quantification of a relevant number of variables, which means that data collection is burdensome and often not immediate. Nevertheless, this study can be replicated in other urban contexts. More specifically, the prediction model could be applied directly in a different urban area, or a recalibration of the coefficients could be carried out for a specific context. Moreover, future research could be developed to enhance the results by employing different modelling approaches, elaborating more data, and analysing the influence of new variables on the speed. From this viewpoint, more innovative models based on machine learning could be developed to improve prediction performance. These models could leverage their greater ability to describe the intricate relationships existing among operating speed and its predictors, as has already been demonstrated in other transportation engineering fields [59–61]. In addition, new measurements could be conducted to introduce new variables. In this regard, a more specific in-depth study could investigate the influence of different traffic calming elements on the V_{85}. Though this paper had already tried to examine such additions, the results are not very appreciable, thus making it appropriate to carry out further analysis. New research could compare the speed data obtained from laser equipment, as used in this study, with speed data collected by floating car data. Indeed, the latter refer to real-time data collected from vehicles in motion through GPS tracking systems or onboard sensors. These data provide detailed information on speed, location, and other traffic metrics. They are particularly valuable for obtaining accurate and up-to-date data without the need for installing costly devices on roads, enabling more efficient and dynamic traffic management. These data could be used to analyse and manage urban traffic, monitor road conditions, and improve transportation planning, especially when many spot speed measurements could not be performed. Besides, the overall methodological framework could be enhanced by integrating additional elements that can characterize the drivers' behaviour. For instance, a more in-depth analysis of intersections, differentiating between types (e.g., signalised, unsignalized) could be significant. This would enable further refinement of the initial segmentation setup. Finally, this study focuses exclusively on the spot speed of cars. Consequently, the results are primarily applicable to small vehicles on urban roads, because this study does not consider mixed traffic volumes that can include other types of vehicles. Nevertheless, it is recommended to extend this study to include a broader range of vehicles to achieve a more comprehensive understanding of traffic dynamics.

CRediT authorship contribution statement

Stefano Raccagni: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Roberto Ventura:** Writing – review & editing, Visualization, Supervision, Methodology. **Benedetto Barabino:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Data and code availability

The data that has been used is confidential.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Benedetto Barabino reports financial support was provided by Ministero dell'Istruzione, dell'Università e della Ricerca. Stefano Raccagni reports a relationship with University of Brescia that includes: funding grants. Benedetto Barabino is an Associate Editor of Heliyon Journal. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors express their gratitude to the Mobility Office of the Municipality of Brescia for their invaluable assistance in selecting a wide range of critical roads for inclusion in the modelling, particularly those with geometric characteristics that deviate from Italian regulations.

This paper is part of some research activities developed by the authors within the framework of the "PNRR": SPOKE 7 "CCAM, Connected Networks and Smart Infrastructure" - WP4, CUP D83C22000690001.

Finally, one of the authors who worked on this paper was supported during the writing by the University of Brescia Department of Civil, Environment, Land and Architecture Engineering and Mathematics (DICATAM), which was included within the research grant "Implementazione di un sistema di monitoraggio della sicurezza del traffico stradale per l'analisi e la mitigazione del rischio di incidentalità", CUP: "D73C2200077002".

Appendix

Appendix 1. Correlation matrix for the factors. The pairs of high correlated variables (correlation coefficient greater than 0.70 o lesser than −0.70 are bold edited)

	V85	F1	F2	F3	F4	L	Ncar	Nc	Lc	Q/C	i	M	Cpr	Cnpr	Bdx	Bsx	Cu	*Cc*	Cd	As	Ar	Ap	Scu	Ppsl
V85	1.00																							
F1	0.50	1.00																						
F2	0.10	−0.42	1.00																					
F3	−0.28	−0.33	−0.34	1.00																				
F4	−0.40	−0.32	−0.32	−0.25	1.00																			
L	0.60	0.33	0.09	−0.07	−0.40	1.00																		
Ncar	0.30	0.56	−0.07	−0.28	−0.27	0.20	1.00																	
Nc	0.43	0.65	−0.23	−0.24	−0.23	0.37	0.61	1.00																
Lc	0.10	0.12	0.15	−0.07	−0.24	−0.10	0.08	−0.14	1.00															
Q/C	0.35	0.38	0.18	−0.19	−0.46	0.30	0.26	0.05	0.34	1.00														
i	0.06	−0.05	0.04	0,00	0.02	0.04	−0.03	−0.10	0.08	0.10	1.00													
M	0.18	0.05	0.03	−0.02	−0.07	0.13	−0.11	0.07	0.04	0.03	−0.05	1.00												
Cpr	0.31	0.31	0.03	−0.22	−0.17	0.31	0.34	0.44	−0.07	−0.01	0.03	0.21	1.00											
Cnpr	0.16	−0.02	0.15	−0.18	0.03	0.13	0.07	0.05	−0.07	0.02	−0.01	0.12	−0.10	1.00										
Bdx	0.03	0.04	0.18	−0.06	−0.20	0.06	0.28	0.07	0.03	0.17	0.05	−0.33	0.08	0.10	1.00									
Bsx	−0.10	−0.22	0.19	0.09	−0.07	−0.02	−0.20	−0.22	−0.02	0.1	−0.05	−0.11	−0.14	−0.04	0.07	1.00								
Cu	0.29	0.24	−0.04	0,00	−0.23	0.16	0.17	0.16	0.18	0.14	0.02	0.29	0.20	−0.18	−0.05	−0.03	1.00							
Cc	0.28	0.48	−0.15	−0.19	−0.19	0.17	0.19	0.44	−0.03	0.28	0.01	0.21	0.31	−0.13	−0.1	−0.16	0.18	1.00						
Cd	0.26	0.45	−0.03	−0.25	−0.23	0.18	**.81**	0.39	0.14	0.27	−0.04	−0.17	0.12	0.16	0.35	−0.10	0.12	−0.18	1.00					
As	−0.08	−0.02	0.19	−0.04	−0.17	0.04	0.25	0.12	0.07	0.06	−0.02	−0.15	0.11	0.06	0.18	0.09	0.02	0.06	0.12	1.00				
Ar	−0.50	−0.33	−0.09	0.22	0.26	−0.28	−0.37	−0.33	−0.03	−0.25	−0.04	−0.08	−0.31	−0.18	−0.17	−0.01	−0.09	−0.29	−0.24	−0.21	1.00			
Ap	0.10	−0.14	0.18	0.11	−0.16	0.10	−0.14	0,00	−0.09	−0.04	0,00	0.06	−0.09	−0.01	0.07	0.15	0.03	−0.03	−0.12	−0.24	−0.14	1.00		
Scu	0.02	0.08	−0.04	−0.07	0.03	0.04	0.05	0.06	−0.01	0.06	−0.01	−0.08	0.02	0.01	0.17	0.11	−0.04	−0.01	0.18	0.06	−0.08	−0.1	1.00	
Ppsl	0.05	−0.10	0.03	−0.02	0.10	0.01	−0.11	−0.05	0,00	−0.06	0.16	0.19	0.08	0.14	−0.09	−0.10	−0.02	−0.06	−0.08	0.05	−0.08	−0.13	0.14	1.00
PSL	0.37	0.27	0.20	−0.16	−0.37	0.24	0.24	0.25	0,00	0.19	−0.04	−0.24	0.03	−0.08	0.23	0.17	0.15	0.13	0.2	0.13	−0.17	0.14	−0.01	−0.38
O	0.47	0.22	0.19	−0.02	−0.44	0.25	0.10	0.15	0.31	0.34	0.09	0.14	0.16	0.04	0.04	0.08	0.51	0.15	0.07	0.05	−0.22	0.13	−0.04	0.05
V	0,00	0.07	0.01	−0.09	−0.01	0.20	0,00	0.09	−0.05	0.16	0.05	0.01	0.15	−0.02	−0.01	−0.01	−0.08	0.22	−0.12	0.12	0.03	−0.06	0.05	0.13
Dacc	−0.46	−0.48	−0.08	0.24	0.40	−0.26	−0.43	−0.32	−0.20	−0.41	0.04	0.03	−0.23	−0.15	−0.19	0.01	−0.06	−0.23	−0.41	−0.08	0.46	−0.06	0.01	0.07
Int	0.36	0.05	0.21	−0.17	−0.12	0.25	0.10	0.03	0.02	0.23	0.12	0.02	0.09	0.09	0.01	0.04	0,00	0,00	0.10	0.07	−0.37	0.07	0.03	0.13
Dint	−0.59	−0.22	−0.17	0.15	0.30	−0.41	−0.20	−0.15	−0.14	−0.26	−0.13	0.04	−0.14	−0.09	−0.14	0,00	−0.11	−0.07	−0.21	−0.04	0.39	−0.11	−0.10	−0.07
Al	0.22	0.38	−0.07	−0.18	−0.17	0.04	0.36	0.28	0.16	0.26	−0.04	−0.19	0.04	−0.10	0.25	0.06	0.09	0.08	0.39	0.04	−0.11	0.10	0.15	−0.16
Ost	−0.12	−0.21	0.08	0.11	0.03	0.03	−0.02	−0.03	0.16	−0.03	−0.05	0.18	−0.08	0.10	−0.12	−0.09	−0.10	−0.07	−0.04	0.08	−0.02	−0.03	0.05	0.01
Dt	0.19	0.10	−0.01	−0.06	−0.05	0.13	0.29	0.19	0.08	0.02	−0.03	0.24	0.06	0.19	0.04	−0.10	0.20	0.03	0.27	−0.01	−0.12	−0.10	0.09	−0.07
Cp	−0.44	−0.35	0.03	0.12	0.24	−0.22	−0.19	−0.23	−0.16	−0.19	0.12	−0.50	−0.17	−0.06	0.32	0.16	−0.41	−0.28	−0.12	0.09	0.29	−0.08	0.03	0.01
Caut	0,00	0.11	−0.05	−0.04	−0.03	0,00	−0.04	−0.03	0,00	0.10	0,00	0.05	−0.03	−0.02	−0.03	0.05	0.04	−0.03	−0.03	−0.09	0.05	−0.02	−0.02	−0.03
Autk	0.09	0.06	0.23	0.08	−0.42	0.07	0.22	0.12	0.06	0.25	−0.05	−0.07	−0.03	0.05	0.27	0.19	0.14	−0.06	0.24	0.22	−0.05	0.10	−0.03	−0.13
Datt	−0.46	−0.28	−0.17	0.09	0.43	−0.38	−0.11	−0.18	−0.16	−0.38	0.01	−0.19	−0.16	−0.15	−0.02	0.02	0,00	−0.21	−0.05	−0.04	0.43	−0.11	0.02	−0.03
G	0.30	0.32	−0.11	−0.12	−0.12	0.15	0.18	0.19	0.06	0.14	0,00	0.03	−0.04	0.06	−0.02	−0.09	−0.02	0.04	0.22	0.02	−0.33	−0.08	0.01	0.10
CB	0.35	0.12	−0.01	−0.19	0.07	0.16	−0.04	0.09	0.05	0.05	0.20	−0.13	0.08	−0.15	−0.09	0.07	0.30	0.07	−0.06	−0.09	0.01	0.06	−0.05	−0.04
PL	−0.34	−0.11	0.01	0.20	−0.09	−0.15	0.04	−0.08	−0.06	−0.05	−0.20	0.13	−0.07	0.13	0.08	−0.10	−0.31	−0.07	0.07	0.08	0.01	−0.06	0.05	0.02
PCb	0.11	0.05	−0.03	0.15	−0.16	0.16	0.10	0.04	−0.08	0.07	−0.05	0.08	0.09	0,00	0.03	−0.01	0.38	0.02	0.09	−0.03	−0.02	0.07	−0.09	−0.12
PCm	−0.11	−0.05	0.03	−0.15	0.16	−0.16	−0.10	−0.04	0.08	−0.07	0.05	−0.08	−0.09	0,00	−0.03	0.01	−0.38	−0.02	−0.09	0.03	0.02	−0.07	0.09	0.12
S	0.30	0.45	−0.1	−0.13	−0.26	0.38	0.27	0.52	0.04	0.12	0.02	−0.07	0.16	0.18	0.12	0.01	0.03	0.16	0.28	0.09	−0.17	0,00	0.01	−0.06

(*continued on next page*)

(*continued*)

	V85	F1	F2	F3	F4	L	Ncar	Nc	Lc	Q/C	i	M	Cpr	Cnpr	Bdx	Bsx	Cu	*Cc*	Cd	As	Ar	Ap	Scu	Ppsl
DD	−0.15	−0.16	−0.07	0,00	0.26	−0.17	−0.13	−0.11	0,00	−0.18	0.01	−0.01	−0.10	−0.09	−0.17	−0.07	−0.12	−0.09	−0.12	−0.12	0.18	−0.08	−0.01	0.05
Ch	−0.11	−0.12	0.11	0.10	−0.09	0.17	−0.10	−0.09	−0.04	0.07	0.11	0.08	0.22	−0.07	−0.03	0,00	0.11	−0.07	−0.09	0.04	0.14	−0.06	−0.06	0.32
RI	−0.11	−0.17	−0.03	0.02	0.21	−0.04	0.04	−0.02	−0.03	−0.13	0.02	0.02	0.16	−0.05	0.02	0.02	−0.05	−0.07	−0.02	0.04	0.12	−0.18	0.20	0.07
Is	0.09	0.09	0.08	−0.02	−0.17	0.07	0.23	0.10	0.06	0.09	0.02	−0.12	0.12	0.12	0.13	−0.11	−0.03	0.21	0.10	0.20	−0.19	−0.16	0.04	−0.08

	PSL	O	V	Dacc	Int	Dint	Al	Ost	Dt	Cp	Caut	Autk	Datt	G	CB	PL	PCb	PCm	S	DD	Ch	RI	Is
PSL	1.00																						
O	0.18	1.00																					
V	−0.13	0,00	1.00																				
Dacc	−0.24	−0.28	−0.1	1.00																			
Int	0,00	0.16	0.04	−0.18	1.00																		
Dint	−0.24	−0.29	0.1	0.19	−0.41	1.00																	
Al	0.28	0.11	−0.01	−0.37	−0.05	−0.05	1.00																
Ost	−0.21	−0.14	0.16	0.12	−0.01	−0.08	−0.1	1.00															
Dt	0.11	0.13	0.01	−0.13	0.03	−0.15	−0.01	0.36	1.00														
Cp	−0.12	−0.23	0.04	0.30	−0.08	0.12	0.01	−0.11	−0.38	1.00													
Caut	0.04	0.03	−0.06	−0.06	0,00	0.01	0.08	−0.11	0.01	−0.05	1.00												
Autk	0.24	0.21	−0.15	−0.14	−0.01	−0.15	0.07	−0.04	0.02	−0.01	0.03	1.00											
Datt	−0.11	−0.17	−0.01	0.33	−0.23	0.30	0,00	−0.02	−0.03	0.26	0,00	−0.05	1.00										
G	−0.03	0.09	0.08	−0.18	0.28	−0.16	0.09	0.07	−0.09	−0.09	−0.02	−0.04	−0.21	1.00									
CB	0.34	0.46	0.01	0.07	0.10	−0.25	0.11	−0.12	0.02	0.02	0.01	−0.06	0.10	0.04	1.00								
PL	−0.33	−0.46	−0.02	−0.07	−0.11	0.26	−0.10	0.12	−0.03	−0.01	−0.01	0.06	−0.11	−0.04	**−.99**	1.00							
PCb	0.18	0.20	−0.08	−0.08	−0.1	−0.12	0.01	−0.09	0.15	−0.27	0.03	0.03	0.05	−0.23	−0.01	0.01	1.00						
PCm	−0.18	−0.20	0.08	0.08	0.10	0.12	−0.01	0.09	−0.15	0.27	−0.03	−0.03	−0.05	0.23	0.01	−0.01	−1,00	1.00					
S	0.21	0.20	0.06	−0.30	0.04	−0.15	0.19	−0.03	0.10	−0.06	0.04	0.18	−0.08	0.23	0.10	−0.11	−0.07	0.07	1.00				
DD	−0.21	−0.09	−0.06	0.20	0.12	0.01	−0.16	0.03	0.04	0.11	−0.02	−0.09	0.22	−0.06	0.04	−0.04	−0.12	0.12	−0.11	1.00			
Ch	−0.35	0.07	0.24	0.17	0.05	−0.04	−0.06	0.13	−0.15	0.25	−0.01	0.06	0.04	−0.05	0.04	−0.03	0.02	−0.02	−0.05	−0.05	1,00		
RI	−0.26	−0.07	0.19	0.15	0.04	0,00	−0.10	0.26	0.18	0.12	−0.05	−0.12	0.23	−0.15	0.10	−0.11	0.04	−0.04	−0.13	0.37	0.18	1,00	
Is	0.10	0.08	0.02	−0.18	0,00	−0.12	0.06	0.01	0.19	0.02	−0.03	0.16	−0.02	0,00	−0.06	0.07	−0.11	0.11	0.14	0.17	−0.09	0.12	1,00

Appendix 2. Full Model e Final Model are provided as equations

Full Model

$$V_{85} = 25.9\,C - 0.56\,E^{*} \bullet 0.56\,E \bullet -0.23\,F \bullet 0.01600\,L \bullet -0.029\,i \bullet 0.00538\,Int \bullet -3.01\,Ncar \bullet 2.49\,Nc \bullet -0.060\,Lc \bullet 0.60\,Cd \bullet 0.74\,Cc \bullet -0.052\,Bdx \bullet -0.321\,Bsx \bullet 0.186\,Autk \bullet -0.1445\,Datt \bullet 3184\,Al \bullet -2.86\,Ost \bullet 0.140\,Dt \bullet -3.75\,Cp \bullet -0.0379\,Dacc \bullet -0.3796\,Dint \bullet 0.64\,M \bullet 3.11\,Cpr \bullet 1.32\,Cnpr \bullet -6.23\,Caut \bullet 0.94\,Cu \bullet 4.19\,G \bullet -0.406\,S \bullet 22.1\,CB \bullet 15.8\,PL \bullet -1.61\,PCb \bullet 3.28\,O \bullet 0.168\,V \bullet 2.10\,Ppsl \bullet 0.1442\,PSL \bullet -0.73\,Q/C \bullet -3043\,As \bullet -1.86\,Ar \bullet -3.17\,Ap \bullet -2.03\,Scu \bullet -0.173\,DD \bullet -5.16\,Ch \bullet 2.02\,RI \bullet -0.59\,Is$$

Final Model

$$V_{85} = 37.72\,C + 0.419\,E^{*} \bullet 0.540\,E \bullet 0.773\,F \bullet 0.01299\,L \bullet 0.00649\,Int \bullet 2.983\,Nc \bullet -0.275\,Bsx \bullet -0.238\,Datt \bullet 1.969\,Al \bullet -3.12\,Ost \bullet -4.861\,Cp \bullet -0.1868\,Dint \bullet -6.19\,Caut \bullet 2.44\,G \bullet 9.29\,CB \bullet -0.44\,PCb \bullet 3.78\,O \bullet 0.1014\,PSL \bullet -2.478\,As \bullet -3.394\,Ar \bullet -2.8\,Ap$$

References

[1] S. Khaddar, B.K. Pathivada, V. Perumal, Modeling over speeding behavior of vehicles using a random parameter negative binomial approach: a case study of Mumbai, India, Transp. Res. Interdiscip. Perspect. 18 (2023) 100790.
[2] ACI, Automobile Club d'Italia, Report Incidenti Stradali 2022, 2023.
[3] C. Lantieri, R. Lamperti, A. Simone, M. Costa, V. Vignali, C. Sangiorgi, G. Dondi, Gateway design assessment in the transition from high to low-speed areas, Transport. Res. F Traffic Psychol. Behav. 34 (2015) 41–53.
[4] H. Kim, D. Jung, Estimation of optimal speed limits for urban roads using traffic information big data, Appl. Sci. 11 (12) (2021) 5710.
[5] J.S. Schaefer, M.A. Figliozzi, A. Unnikrishnan, Evaluation of posted speed limits reductions on urban roads with a high percentage of cyclists, Transport. Res. Rec. 2676 (6) (2022) 685–695.
[6] A. Badiger, Kuldeep, M. Archana, V. Anjaneyappa, Effectiveness of speed calming measures along arterial roads, Lecture Notes in Civil Engineering 261 (2023) 691–702.
[7] N. Berloco, S. Coropulis, G. Garofalo, P. Intini, V. Ranieri, Analysis of the factors influencing speed cushion effectiveness in the urban context: a case study experiment in the city of bari, Italy, Sustainability 15 (8) (2023) 6352.
[8] O. Mora, M. Figueroa, D. Borrero, J. Abello, A. Mattos-Rodriguez, T. Naranjo, C. Urrego, Operational assessment of speed Control measures on capacity of urban roads in metropolitan area of barranquilla, Colombia, Civil Engineering and Architecture 11 (1) (2023) 22–35.
[9] A. Montella, A. Calvi, F. D'Amico, C. Ferrante, F. Galante, F. Mauriello, M. Rella Riccardi, A. Scarano, A methodology for setting credible speed limits based on numerical analyses and driving simulator experiments, Transport. Res. F Traffic Psychol. Behav. 100 (2024) 289–307.
[10] ETCS – European Transport Safety Council, Reducing Speeding in Europe, PIN Flash report 36 (2019).
[11] M. Bassani, D. Dalmazzo, G. Marinelli, C. Cirillo, The effects of road geometrics and traffic regulations on driver-preferred speeds in northern Italy. An exploratory analysis, Transport. Res. F Traffic Psychol. Behav. 25 (2014) 10–26.
[12] A. Thiessen, K. El-Basyouny, S. Gargoum, Operating speed models for tangent segments on urban roads, Transport. Res. Rec. 2618 (1) (2017) 91–99.
[13] V. Martinelli, R. Ventura, M. Bonera, B. Barabino, G. Maternini, Effects of urban road environment on vehicular speed. evidence from Brescia (Italy), Transport. Res. Procedia 60 (2022) 592–599.
[14] D.D. Dinh, H. Kubota, Profile-speed data-based models to estimate operating speeds for urban residential streets with a 30km/h speed limit, IATSS Res. 36 (2) (2013) 115–122.
[15] R. Ewing, E. Dumbaugh, The built environment and traffic safety: a review of empirical evidence, J. Plann. Lit. 23 (4) (2009) 347–367.
[16] T. Esposito, R. Mauro, F. Russo, G. Dell'Acqua, Speed prediction models for sustainable road safety management, Procedia - Social and Behavioral Sciences 20 (2011) 568–576.
[17] S.A. Gargoum, K. El-Basyouny, Exploring the association between speed and safety: a path analysis approach, Accid. Anal. Prev. 93 (2016) 32–40.
[18] M. Bonera, G. Maternini, Methodology for the application of predictive method for road safety analysis in urban areas. The case study of Brescia, Transport. Res. Procedia 45 (2020) 659–667.
[19] B. Barabino, M. Bonera, G. Maternini, A. Olivo, F. Porcu, Bus crash risk evaluation: an adjusted framework and its application in a real network, Accid. Anal. Prev. 159 (2021) 106258.
[20] F. Porcu, A. Olivo, G. Maternini, M. Coni, M. Bonera, B. Barabino, Assessing the Risk of Bus Crashes in Transit Systems, vol. 81, European Transport/Trasporti Europei, 2021.
[21] E.S. Park, K. Fitzpatrick, S. Das, R. Avelar, Exploration of the relationship among roadway characteristics, operating speed, and crashes for city streets using path analysis, Accid. Anal. Prev. 150 (2021) 105896.
[22] B. Barabino, M. Bonera, G. Maternini, F. Porcu, R. Ventura, Refining a crash risk framework for urban bus safety assessment: evidence from Sardinia (Italy), Reliab. Eng. Syst. Saf. 245 (2024) 110003.
[23] M. Bonera, B. Barabino, G. Yannis, G. Maternini, Network-wide road crash risk screening: a new framework, Accid. Anal. Prev. 199 (2024) 107502.
[24] M. Bassani, E. Sacchi, Calibration to local conditions of geometry-based operating speed models for urban arterials and collectors, Procedia - Social and Behavioral Sciences 53 (2012) 821–832.
[25] K. Fitzpatrick, C.B. Shamburger, R.A. Krammes, D.B. Fambro, Operating speed on suburban arterial curves, Transport. Res. Rec. 1579 (1) (1997) 89–96.
[26] C.M. Poe, J.M. Mason, Analysing influence of geometric design on operating speeds along low-speed urban streets: mixed-model approach, Transport. Res. Rec. 1737 (1) (2000) 18–25.
[27] K. Fitzpatrick, P. Carlson, M. Brewer, M. Wooldridge, Design factors that affect driver speed on suburban streets, Transport. Res. Rec. 1751 (1) (2001) 18–25.
[28] J. Wang, K.K. Dixon, H. Li, M. Hunter, Operating-speed model for low-speed urban tangent streets based on in-vehicle global positioning system data, Transport. Res. Rec. 1961 (1) (2006) 24–33.
[29] J.P. Tarris, C.M. Poe, J.M. Mason Jr, K.G. Goulias, Predicting operating speeds on low-speed urban streets: regression and panel analysis approaches, Transport. Res. Rec. 1523 (1996).

[30] C.M. Poe, J.M. Mason, Geometric design guidelines to achieve desired operating speed on urban streets. Proceedings of the 65th ITE Annual Meeting, 1995, pp. 70–74.
[31] K. Fitzpatrick, S.-P. Miaou, M. Brewer, P. Carlson, M.D. Wooldridge, Exploration of the relationships between operating speed and roadway features on tangent sections, J. Transport. Eng. 131 (4) (2005) 261–269.
[32] V. Martinelli, R. Ventura, M. Bonera, B. Barabino, G. Maternini, Estimating operating speed for county road segments – Evidence from Italy, International Journal of Transportation Science and Technology 12 (2) (2023) 560–577.
[33] K.K. Tottadi, C. Varma, A. Mehar, Influence of horizontal curve geometry on operating speeds on four-lane divided highways under heterogeneous traffic conditions, J. Inst. Eng.: Series A 103 (4) (2022) 1135–1145.
[34] M. Obaidat, A. Mohammad, GPS vehicle tracking in urban areas, Civil-Comp Proceedings 100 (2012).
[35] N. Eluru, V. Chakour, M. Chamberlain, L.F. Miranda-Moreno, Modeling vehicle operating speed on urban roads in Montreal: A panel mixed ordered probit fractional split model, Accid. Anal. Prev. 59 (2013) 125–134.
[36] D. Aleksandra, S. Šurdonja, D. Malatestinić, L. Strineka, Operating speed on urban road network – case study in Rijeka, Croatia. Fifth International Conference on Road and Rail Infrastructure, 2018, pp. 1113–1119.
[37] Y. García-Ramírez, L. Paladines, C. Verdesoto, P. Torres, Car Drivers Do Not Choose Their Speed in Urban Environments: Speed Models in Tangent Streets, 2021, pp. 421–428, 2021-April.
[38] H. Hamad, E. Sacchi, Operating speed prediction models for urban streets in residential areas: a case study in the city of saskatoon, in: S. Walbridge, M. Nik-Bakht, K.T.W. Ng, M. Shome, M.S. Alam, A. El Damatty, G. Lovegrove (Eds.), Proceedings of the Canadian Society of Civil Engineering Annual Conference 2021, Springer Nature, Singapore, 2022, pp. 1–14. Cur.
[39] T.M. Al-Bahr, S.A. Hassan, O.C. Puan, N. Mashros, N.S.A. Sukor, Speed-flow-geometric relationship for urban roads network, Appl. Sci. 12 (9) (2022) 4231.
[40] HCM, Highway Capacity Manual, seventh ed., 2022.
[41] A. Maji, A. Tyagi, Speed prediction models for car and sports utility vehicle at locations along four-lane median divided horizontal curves, Journal of Modern Transportation 26 (4) (2018) 278–284.
[42] G. Sil, S. Nama, A. Maji, A.K. Maurya, Effect of horizontal curve geometry on vehicle speed distribution: a four-lane divided highway study, Transportation Letters 12 (10) (2020) 713–722.
[43] H. Gonzalo-Orden, H. Pérez-Acebo, A.L. Unamunzaga, M.R. Arce, Effects of traffic calming measures in different urban areas, Transport. Res. Procedia 33 (2018) 83–90.
[44] A. Solowczuk, Effect of traffic calming measures implemented on the approach to the tempo–30 zone on the degree of speed reduction. IOP *Conference Series: materials Science and Engineering* 603 (2) (2019) 022044.
[45] ISTAT, Istituto Nazionale di Statistica, Bilancio Demografico Anno 2022, Brescia, Istat.it, 2023.
[46] Lombardia -RL. Regione, Piano Regionale per la Mobilità e i Trasporti, 2016.
[47] Polis – Regione Lombardia, Incidenti stradali a ciclisti regione Lombardia, 2020 (In Italian).
[48] MIT – Ministry of Infrastructure and Trasport, Norme funzionali e geometriche per la costruzione delle strade, 2001. DM (6792/2001). SO alla GU n.3. Del 4.1.2002. (In Italian).
[49] DP- Documento di Piano, Comune di Brescia, 2023.
[50] P. Flach, Machine Learning: the Art and Science of Algorithms that Make Sense of Data, Cambridge University Press, 2012.
[51] F. Russo, S.A. Biancardo, M. Busiello, Operating speed as a key factor in studying the driver behaviour in a rural context [Number: 2], Transport 31 (2) (2016) 260–270.
[52] MIT, Nuovo Codice Della Strada, 30-04-1992, p. 285. Ministry of Infrastructure and Trasport (1992) - Decreto Legislativo.
[53] LL.PP. – Ministry of Public Works, Direttive per la redazione, adozione ed attuazione dei piani urbani del traffico, 1995. GU Serie Generale n.146 del 24-06-1995 - Suppl. Ordinario n. 77. (In Italian).
[54] M. Bassani, L. Catani, C. Cirillo, G. Mutani, Night-time and daytime operating speed distribution in urban arterials, Transport. Res. F Traffic Psychol. Behav. 42 (2016) 56–69.
[55] T.J. Kwon, L. Fu, K. Hossain, Effect of Inclement Weather Conditions on Macroscopic Traffic Behavior. *Proceedings*, vol. 3, Annual Conference- Canadian Society for Civil Engineering, 2012.
[56] X. Ni, H. Huang, A. Chen, Y. Liu, H. Xing, Effect of heavy rainstorm and rain-induced waterlogging on traffic flow on urban road sections: integrated experiment and simulation study, J. Transport. Eng. Part A: Systems 147 (10) (2021).
[57] A. Polus, K. Fitzpatrick, D.B. Fambro, Predicting operating speeds on tangent sections of two-lane rural highways, in: Transportation Research Record, vol. 1737, Publisher: SAGE Publications Inc, 2000, pp. 50–57, 1.
[58] B. Barabino, Transit bus route network design: a model and its application in a real network, WIT Trans. Built Environ. 107 (2009) 369–382.
[59] R. Ventura, B. Barabino, G. Maternini, Prediction of the severity of exceeding design traffic loads on highway bridges, Heliyon 10 (1) (2024) e23374.
[60] R. Ventura, G. Maternini, B. Barabino, Traffic hazards on main road's bridges: real-time estimating and managing the overload risk, IEEE Trans. Intell. Transport. Syst. 15 (2024) 12239–12255.
[61] R. Ventura, A. Ghirardi, D. Vetturi, G. Maternini, B. Barabino, Comparing the vibrational behaviour of e-kick scooters and e-bikes: evidence from Italy, International Journal of Transportation Science and Technology (2023).