Transportation Research Part F 25 (2014) 10–26

Contents lists available at ScienceDirect

Transportation Research Part F

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

# The effects of road geometrics and traffic regulations on driver-preferred speeds in northern Italy. An exploratory analysis

Marco Bassani [a,*], Davide Dalmazzo [a,1], Giuseppe Marinelli [a,2], Cinzia Cirillo [b,3]

[a] *Politecnico di Torino, Department of Environment, Land and Infrastructures Engineering (DIATI), 24, corso Duca degli Abruzzi, 10129 Torino, Italy*
[b] *University of Maryland, Department of Civil and Environmental Engineering, 3250 Kim Building, 20742 College Park, MD, United States*

ARTICLE INFO

*Article history:*
Received 14 April 2014
Accepted 24 April 2014

*Keywords:*
Operating speed
Road geometrics
Driving regulations
Driver behaviour
Speed percentile
Random effects model
ANCOVA method
Bayesian information criterion

ABSTRACT

Speeds are affected by several variables such as driver characteristics, vehicle performance, road geometrics, environmental conditions and driving regulations. It is therefore important to study the relationships between speed and such variables in order to facilitate conscious speed management on existing and planned roads, and to induce drivers to select a speed consistent with the posted limit. This relationship is of great interest to those who wish to achieve roadway functionality and improve overall road safety.

A small number of studies have focused on this objective; however, few of them concern urban roads and they are limited to specific road types and recently built-up areas. These studies often refer to the 85th percentile of the speed distribution and are relevant to locations which are homogeneous in terms of geometry, environment, driving regulations and vehicle type.

This paper presents results obtained from a study carried out on urban arterials and collectors characterized by dissimilar geometric features which facilitated the inclusion of a fully representative range of variables. A general model able to predict operating speed for a generic percentile was calibrated using three different strategies: (a) a simple multiple regression analysis in which the variables were selected using the Bayesian Information Criterion (BIC); (b) the analysis of covariance method including random effects on the same set of variables as in (a); and, finally, (c) the analysis of covariance method with random effects and a new selection of variables (again using BIC). The analysis shows a dramatic variation in results depending on the method selected. In particular, when random effects are considered, almost all the variables are found to be statistically significant.



## 1. Introduction

Urban roads are characterized by an elevated accident risk factor, which is mainly due to high traffic volumes, the coexistence of multiple categories of road users moving at different speeds, and the number of daily activities along the roadsides

* Corresponding author. Tel.: +39 011 5645635.
*E-mail addresses:* marco.bassani@polito.it (M. Bassani), davide.dalmazzo@polito.it (D. Dalmazzo), giuseppe.marinelli@polito.it (G. Marinelli), ccirillo@umd.edu (C. Cirillo).
[1] Tel.: +39 011 5645622.
[2] Tel.: +39 011 5645625.
[3] Tel.: +1 301 405 6864; fax +1 301 405 2585.

http://dx.doi.org/10.1016/j.trf.2014.04.019
1369-8478/

(Committee for Guidance on Setting, 1998). In 2010, in Italy, the number of road accidents on urban roads was 150,602 which represented more than 71% of the total crashes, while the number of fatalities was 1744 equal to 45% of the total fatalities recorded on the national road network (Ministero delle Infrastrutture e dei Trasporti, 2011). More than 20% of urban fatalities involved vulnerable road users like pedestrians and cyclists who are prone to accidents with fast passenger cars or light duty vehicles. As a consequence, speed management remains crucial for creating a safe road system and for achieving the EU imposed target of a further 50% reduction in road fatalities in the decade 2011–2020 (European Transport Safety Council, 2012).

There is a general consensus that if speeds on a specific road section decrease, then accidents will be less severe, and therefore fewer crashes will be reported (Global Road Safety Partnership, 2008; Hauer, 2009). Garber and Gadiraju (1989) found that the crash severity is linked to the average speed, while the speed variance has a bearing on the number of crashes. Hauer (2009) stressed that speed data from flow speed observations and crash scene investigations are not estimated with a uniform degree of accuracy, so any results from research into the effects of speed variations are inconclusive and inadequate. Similar conclusions were drawn by Aarts and van Schagen (2006).

Studying the relationship between driver speed and the variables affecting driver behaviour is important for conscious speed management on existing and planned roads and also helpful to those seeking ways to encourage drivers to select a speed consistent with the posted speed limit (PSL) and traffic conditions.

High unsafe speeds occur mainly under free-flow conditions when low-density streams are mostly composed of isolated vehicles. In urban road networks, free-flow speeds are difficult to observe given the level of traffic and the influence of traffic signals and other traffic control devices. According to the Highway Capacity Manual (Transportation Research Board, 2010), in the case of urban roads, free-flow speeds occur at off-peak hours in the central part of urban street segments where traffic control systems do not affect driver speed choice.

The research presented here aims to generate new relationships between geometric variables and speed for urban roads. This has been accomplished by using an approach capable of identifying the most significant variables affecting average speed and by also taking into account the dispersion of the collected data.

The study builds on a previous observational investigation (Bassani & Sacchi, 2012) performed on a limited dataset, which has been enriched with new free-flow speed data. All the observations were conducted within the municipality of Torino (Italy) on urban arterials and collector streets that were selected in order to include a wide range of fully representative variables. Speed measurements were recorded during the times of day in which free-flow traffic conditions prevailed. Different models have been constructed taking into account the hierarchical structure of collected data, thus distinguishing the set to which each speed belongs (lane, section and road).

## 2. Background to operating speed

### 2.1. Driver behaviour on urban roads

In contrast to rural roads, urban roads have more operating functions. In fact, designers have to provide a harmonious and comfortable driving environment in which longitudinal and transversal movements coexist, and where other roadway users, such as cyclists and pedestrians, have to be safely accommodated. However, on too many occasions the operating speeds exceed the PSL (European Transport Safety Council, 2011).

In Fig. 1, each line represents the percentage of drivers that exceeded the PSL during an observational study on selected urban roads in a representative sample of European cities during the decade 2000–2009. Observations reveal that in some countries the percentage of aggressive drivers is constantly high (in Austria, 70% of vehicles exceed the 30 km/h limit in residential zones and 51% exceed the 50 km/h limit), in other cases the percentage is relatively low (roads in Switzerland and

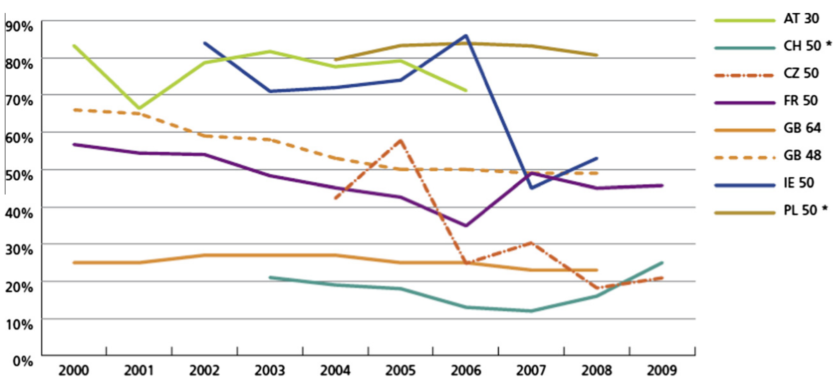



**Fig. 1.** Percentage of cars and vans exceeding the posted speed limits (PSL) on European urban roads in the period 2000–2009 (the labels indicate the country and the PSL of considered roads) (European Transport Safety Council, 2011).

Britain with a PSL equal to 64 km/h), and in others a sizeable decrease has been registered following the introduction of new initiatives on speed control enforcement (roads in Ireland, France, Czech Republic, and Britain with a PSL equal to 48 km/h).

Experience shows that, where drivers are very aggressive, speed enforcement systems should be adopted or, when possible, the geometric characteristics of the road should be re-evaluated and modified (Garrick & Wang, 2005).

In support of this, Mannering (2009) highlighted that speed enforcement plays a large role in determining the speed of uninhibited drivers when travelling above the speed limit, and that other significant variables such as gender, age and ethnicity have a lesser impact.

### 2.2. Operating speeds

It has been widely demonstrated and commonly accepted that in free-flow conditions the driver population can be separated into two groups (Fig. 2). The larger group consists of users that select safe speeds and tend to respect the PSL; this group normally behaves in a manner consistent with road geometry and flow conditions. The second group consists of users that drive fast and aggressively, and frequently exceed the PSL. It has been repeatedly demonstrated (Lamm, Psarianos, & Mailaender, 1999) that, for any specific road section, the speeds of these two groups combined are normally distributed, and that the 85th percentile of speed distribution ($V_{85}$) is commonly considered a good indicator of the threshold between them. This parameter has been widely investigated and a considerable number of authors have linked operating speeds to geometric design variables in $V_{85}$ models (Transportation Research Circular, 2011).

Basing their conclusions on a literature review, Garrick and Wang (2005) sustained that the $V_{85}$ in urban areas is insensitive to longitudinal geometric characteristics, while it is highly sensitive to transversal characteristics of road elements such as lane and shoulder width, margin width, and driveway density. This explains why some have questioned the validity of applying the design speed concept to urban roads.

In literature $V_{85}$ includes the speeds of most conservative drivers and is, therefore, considered an appropriate indicator for operating speeds, and used by designers to make comparisons with the assumed design speed. Any difference between operating and design speeds should be minimized so as to be in harmony with the expectations of careful drivers and to achieve design consistency for new facility designs or roadway redesigns.

According to Lamm et al. (1999), consistency between the PSL, design and operating speed is a prerequisite for safe traffic operations. In particular, road safety literature reports many relationships linking safety to road geometric features (American Association of State Highway & Transportation Officials, 2010), and others where $V_{85}$ is correlated to road geometrics (Transportation Research Circular, 2011).

$V_{85}$, however, cannot represent the entire speed distribution (Fig. 2) and is not useful for the derivation of a possible correlation with accident indicators. This is the reason why some have recently emphasized the need to use the average speed ($V_{50}$) and its standard deviation ($\sigma$) instead of $V_{85}$ (Transportation Research Circular, 2011).

## 3. Background to operating speed models for the urban environment

Operating speed models help road designers with the selection of geometric characteristics of road elements which can positively affect driver behaviour (Garrick & Wang, 2005). An understanding of driver perception of the road environment is helpful when selecting the combination of geometric features and traffic control systems that should encourage road users to adopt the most appropriate speed (National Cooperative Highway Research Program, 2004).

In fact, the speed chosen by drivers is also influenced by factors such as weather conditions, and road lighting efficiency. Bassani and Mutani (2012) demonstrated that driver behaviour is certainly influenced by environmental lighting conditions,

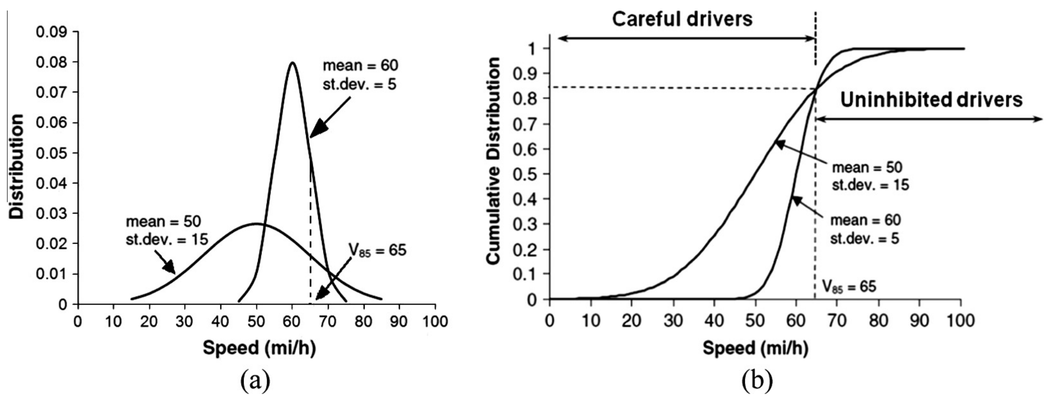



**Fig. 2.** Two examples of free-flow speed distribution different in terms of mean and standard deviation but equal in terms of $V_{85}$, and identification of careful and uninhibited drivers. (a) Probability and (b) cumulative distribution.

driving regulations (e.g., PSL and overtaking rules) and visibility conditions. Misaghi and Hassan (2005) suggested that no single operating speed model should be universally adopted. Model format, independent variables, and regression coefficients differ from one case to the other.

Among other variables, Wang, Dixon, Li, and Hunter (2006) noted that the geometric elements of roads are highly dependent on the PSL. Therefore, they considered it appropriate to investigate the influence of road environment characteristics without including the PSL in speed studies. However, they did not take into account that for some old urban facilities this correlation is very low or non-existent. This is the case, for example, with Italian and European urban roads which, in many cases, were designed and/or built without reference to any design standards.

Survey technique and data collection methodology are also critical elements in speed studies (Transportation Research Circular, 2011). In fact, sample sizes, incidental errors related to the survey methodology adopted (e.g., radar and laser speed gun data need be corrected for the cosine effect), and possible influences on driver behaviour (e.g. due to the presence of the test equipment that might be perceived as a speed enforcement tool) might all affect the quality of the data and the results of the deriving models.

While a large number of operating speed models with different characteristics are available in literature, only a few relate to urban roadways. Among these, the investigation of Fitzpatrick, Shamburger, Krammes, and Fambro (1997) was based on free-flow speed data collected at several sections on eight suburban arterials in Texas with speed limits ranging from 64 to 81 km/h. From their data analysis, the most significant variables were found to be the approach density (frequency of driveways and intersections within the roadway section), and the horizontal radius.

The observations from 27 sites on urban collector streets in central Pennsylvania formed the database considered by Poe and Mason. (1995) who were the first to suggest the inclusion of transversal geometric and environmental factors in speed formulation. The resulting model, valid also for straight sections where the degree of curvature (DC) is zero, was the following:

$$V_{85}\ [\mathrm{km/h}] = 61.7 - 0.23 \cdot \mathrm{DC} - 0.52 \cdot G - 0.82 \cdot \mathrm{HR} - 2.66 \cdot \mathrm{IN} - 1.08 \cdot \mathrm{DR} + 0.15 \cdot \mathrm{LW} (R^2 = 0.67) \tag{1}$$

where $G$ is the absolute value of the roadway gradient, HR is the hazard rating which is a measure of the number and severity of lateral obstructions within 1.5 m of the road, IN is the number of intersections, DR is the number of driveways, and LW is the lane width in m.

The use of linear regression equations has been proposed extensively in literature. The PSL is the most frequently used variable in these kinds of studies; the NCHRP Report 504 (National Cooperative Highway Research Program, 2003) assesses PSL linear effects as a single covariate.

Wang et al. (2006) modified the survey approach; they collected continuous speed data from 200 vehicles equipped with GPS devices on tangents along low speed urban corridors in Atlanta. The authors considered a variety of cross sections and adjoining terrain using both numerical and Boolean parameters. The following equation resulted from their study:

$$\begin{aligned} V_{85}[\mathrm{km/h}] = {} & 50.50 + 10.39 \cdot \mathrm{NL} - 0.08 \cdot \mathrm{RS} - 0.13 \cdot \mathrm{DD} - 0.21 \cdot I + 4.82 \cdot C - 6.82 \cdot \mathrm{SW} - 5.10 \cdot P + 5.30 \cdot \mathrm{LU}_1 \\ & + 5.24 \cdot \mathrm{LU}_2 (R^2 = 0.67) \end{aligned} \tag{2}$$

where NL denotes the number of lanes, RS the density of trees and utility poles (number/km) divided by their average offset from the roadway (m), DD is the density of driveways (driveways/km), I represents the density of $T$-intersections (intersections/km), $C$ indicates the presence of curbs (0 if there is no curb, otherwise 1), SW indicates the presence of sidewalks (0 if there is no sidewalk, otherwise 1), $P$ the presence of parking (0 if there is no on-street parking, otherwise 1), and finally $\mathrm{LU}_1$ and $\mathrm{LU}_2$ are commercial and residential land use indicators respectively (1 for positive case, otherwise 0).

Wang et al. (2006) found that, when using the PSL as a covariate, the intercept and variables such as roadside object, driveway and $T$-intersection densities, and on-street parking become statistically insignificant. The reason for this is that design speed depends on the PSL, so geometric variables are highly correlated to speed limits. The speed limit is, therefore, not an independent parameter and should not be included.

Referring to 67 case studies of four-lane rural and suburban roads in Indiana, Tarko (2009) proposed a model containing a mix of cross-sectional geometric and environmental parameters. The model demonstrates that in free-flow conditions the preferred speed of drivers is the result of a trade-off between the perceived risks of having an accident, the perceived presence of speed enforcement measures, and the subjective cost of time.

Not with standing the availability of operating speed models for different road types and elements, the low coefficient for some of the determinants confirms that the selection of variables is of fundamental importance for the validity of the proposed models. Moreover, the use of a modelling technique other than linear regression analysis for $V_{85}$ is desirable in order to correctly compute the effects of geometric parameters, avoid the assumption of data independence, and limit the use of aggregated data with an overstatement of the coefficient of determination (Transportation Research Circular, 2011).

Although it has been promoted and suggested by many authors (Hassan, 2004; Misaghi, 2003), panel data (PD) models are not yet widely used. The PD analysis is a statistical method employed in social, medical and econometric sciences and deals with data collected over time from the same sample of individuals; following which a regression analysis is performed considering these two dimensions (time and individuals). The use of such an approach in the field of operating speed analysis is complicated by the fact that a number of isolated vehicles have to be tracked through sites, thus multiple speed data need to be collected for individual drivers.

The use of the PD approach on urban roads was first proposed by Tarris, Poe, Mason, and Goulias (1996). In their pioneering work, the authors considered 27 urban collectors in Pennsylvania characterized by different variables in terms of roadway alignment, cross section, roadside elements and land use. In particular, the analysis was limited to curved segments with radii comprised of values between 11 and 230 m. As a consequence of the site characteristics selected, they found that $V_{85}$ depends linearly on the degree of curve (DC), and that a small variation in model coefficients is evident when single or aggregated values for each site are taken into account. The PD methodology used illustrated how speeds are highly dependent on roadway geometry, driver age and time of observation.

The difficulties associated with the use of the PD approach, primarily due to the complications related to the recording over time of operating speeds of individual undisturbed drivers, can be surmounted by employing the modelling technique adopted by Figueroa and Tarko in their 2005 study. They assumed that speeds recorded at a site were normally distributed and proposed a new modelling approach able to calculate any percentile speed as a linear combination of variables affecting both the central tendency and the dispersion of collected data.

## 4. Data collection and methodology

Starting with a literature review and bearing in mind all the limitations and deficiencies of databases and models as highlighted in Misaghi and Hassan (2005), this investigation was organized with an in-field speed survey that resulted in the collection of speed data from each lane of 16 different road sections. The final dataset has the following characteristics:

- cars in free-flow conditions; data corresponding to commercial vehicles were excluded because they were not sufficient in number to form a consistent group of observations;
- cars moving at a uniform speed; cars in accelerating/decelerating conditions and cars influenced by traffic control systems like traffic lights and priority signals at intersections were not considered;
- tangent and curved sections are characterized by very high radii due to the square grid pattern of the road network system of Torino; in only a few cases the roads presented curved segments characterized by radii greater than 250 m which do not impose significant speed variations on contiguous segments.

### 4.1. Survey methodology

In order to ensure the quality and the integrity of the database, field investigations were conducted by means of longitudinal and transversal measurements taken on ordinary lanes and excluding those dedicated to public services (bus and taxi). A number of the measurements were taken by combining the use of a laser speed gun with a high-speed digital video camera; in some cases the camera was used alone with the optic perpendicular to the centreline of the road. Consequently, the database was formed using longitudinal and perpendicular measurement techniques employed at a single point location. At a subsequent stage and through the analysis of video frames, the running speeds of isolated vehicles were calculated and included in the database.

In accordance with (Wang et al., 2006), great attention was paid to the selection of unobtrusive measurement positions, which were located in places which would avoid any psychological effects on drivers, and would limit the cosine effect. Furthermore, measurement periods were not allowed to exceed 15 min, and were repeated at different intervals under similar traffic and weather conditions. Only the speeds of isolated vehicles not conditioned by traffic signals and surrounding vehicles were recorded, thus considering typical free-flow conditions to which any speed-related design parameter usually refers. Sections of 150 m in length were located across the middle point between two successive intersections along tangents, and in the middle section along curves.

Speed data were consolidated only in cases where the time headways and tailways were greater than 5 s as per accepted practice in recent papers (Figueroa & Tarko, 2005; Tarko, 2009), and were rejected in the remaining cases in compliance with best practice suggested in literature (Nie & Hassan., 2007).

Driver behaviour at weekends differs from that on weekdays as a consequence of a change in travel goals (Shinar & Compton, 2004). Observations made by Agarwal (2004) on urban roads reveal that speeds increase only slightly at weekends with respect to weekdays. Assum, Bjørnskau, Fosser, and Sagberg (1999) assert that variations in speed between night-time and daytime is partly explained by the fact that careful drivers, for the most part elderly people and women, try to avoid driving at night-time. More aggressive drivers, a group consisting largely of young males, represent a greater proportion of the total driver population during the hours of darkness. In this research, speed surveys were concentrated during daylight hours with very low traffic volumes. Therefore, the results obtained are relevant for the times of day and days of the week when traffic is light and driver preferences can be observed.

### 4.2. Case studies

As stated previously, the sections were chosen by excluding segments characterized by high variability in features along their length. The road segments (Table 1) form part of the urban road network of the municipality of Torino (Italy). These facilities match the typologies of urban arterials (A) and collector (C) streets indicated by Italian road design standards

**Table 1**
Geometric and operative characteristics of the selected road segments.

| Road # | Section # | Road name Unit → | Type – | NT – | NL – | LW m | A – | R m | $V_{85}$ km/h | PSL km/h | # Data – |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | lungo Stura Lazio | A | 2 | 3 | 3.25 | C | 295 | 96.0 | 70 | 518 |
| 2 | 2 | corso Marche | A | 2 | 2 | 3.50–4.00 | C | 850 | 106.0 | 70 | 439 |
| 3 | 3 | corso Orbassano | A | 1 | 3 | 4.00–4.50 | C | 250 | 96.4 | 70 | 212 |
| 4 | 4 | corso Grosseto | A | 2 | 2 | 3.50 | T | – | 93.0 | 70 | 602 |
| 5 | 5 | corso Massimo d'Azeglio | C | 2 | 2 | 3.50–4.10 | T | – | 74.0 | 50 | 200 |
| | 6 | corso Massimo d'Azeglio | C | 2 | 2<br>3 | 2.80–3.00<br>3.25–4.00 | T | – | 66.7 | 50 | 250 |
| | 7 | corso Massimo d'Azeglio | C | 2 | 3 | 2.80–3.00 | T | – | 65.2 | 50 | 300 |
| 6 | 8 | via Borsellino | C | 1 | 2 | 4.42–4.62 | T | – | 62.2 | 50 | 100 |
| | 9 | via Falcone | C | 1 | 1 | 5.00–5.90 | T | – | 50.0 | 30 | 300 |
| | 10 | via Principi d'Acaja | C | 1 | 1 | 3.85–4.00 | T | – | 46.0 | 50 | 100 |
| 7 | 11 | via Filadelfia | C | 1 | 1 | 4.19–4.35 | T | – | 60.0 | 50 | 100 |
| | 12 | via Filadelfia | C | 1 | 1 | 3.98–4.48 | T | – | 55.0 | 40 | 100 |
| | 13 | via Filadelfia | C | 1 | 1 | 4.30–4.40 | T | – | 54.0 | 50 | 300 |
| 8 | 14 | corso Unità d'Italia | A | 2 | 3 | 3.00 | T | – | 90.8 | 50/70 | 918 |
| | 15 | corso Unità d'Italia | A | 2 | 3 | 3.00 | C | 2100 | 81.4 | 70 | 300 |
| | 16 | corso Unità d'Italia | A | 2 | 3 | 3.00 | T | – | 86.2 | 50/70 | 600 |
| | Total | | – | 25 | 76 | 2.80–5.90 | – | – | – | 30/70 | 5339 |

(Ministero delle Infrastrutture e dei Trasporti, 2001). In the Italian road network system, the arterials support heavy traffic volumes moving at a relatively high speed, and represent the terminals in the urban areas of the freeway system that, when present, encircles cities. Collector streets represent the most important network for internal movements, and provide access to adjacent built-up areas and to the local street network; in this case, road functions are divided between mobility and access with one sometimes prevailing over the other.

Of the total 16 road sections (Table 1), seven are arterials and nine are collectors. In this latter category, there is a further subdivision between divided and undivided travelled ways. Fig. 3 shows the schematic representation of these two types of sections.

The road sections included in the database are listed in Table 1, in which the corresponding information regarding road, geometry and operating characteristics are also reported.

The symbols used to denote the geometric characteristics of the roads in Table 1 and Fig. 3 are as follows:

- NT is the number of travelled ways,
- NL is the number of lanes per direction,
- LW is the lane width,
- A is the horizontal alignment of the segments (T = tangent, C = curve), and
- R is the radius of the centreline.

In some cases, a variation in the LW between lanes of the same travelled way was found. In other cases, the two travelled ways of the same road segment are not symmetrical as is the case with road segments Nos. 3 and 6 (Table 1). Finally, some arterial sections present different PSLs for the two travelled ways, as per segments Nos. 14 and 16.

Other variables presented in Fig. 3 are:

- RSW and LSW, right and left shoulder width respectively (in m),
- MW, median width (in m),
- LP, lane position in the travelled way, assigning an increasing whole number to the external lane and moving toward the centreline, and
- PUB, PKL and S, presence of dedicated bus and taxi lanes, parking lanes and sidewalks near the travelled way.

### 4.3. Variable selection

All the variables that form the database are presented in Table 2, where abbreviations used in this paper are also reported. As can be noted, some variables are numerical continuous (NC), some are numerical discrete (ND), and some are Boolean (B). For the last-named only integer values 0 or 1 were used in the model, with a value of 0 assigned when the road element is absent and 1 when it is present along the considered road segment.

Table 2 includes new variables not listed before. RS and LS indicate the presence of a right or left shoulder respectively in the transversal section, while in the length of 1 km across the same section (500 m before and after the section) the following six variables were also considered:

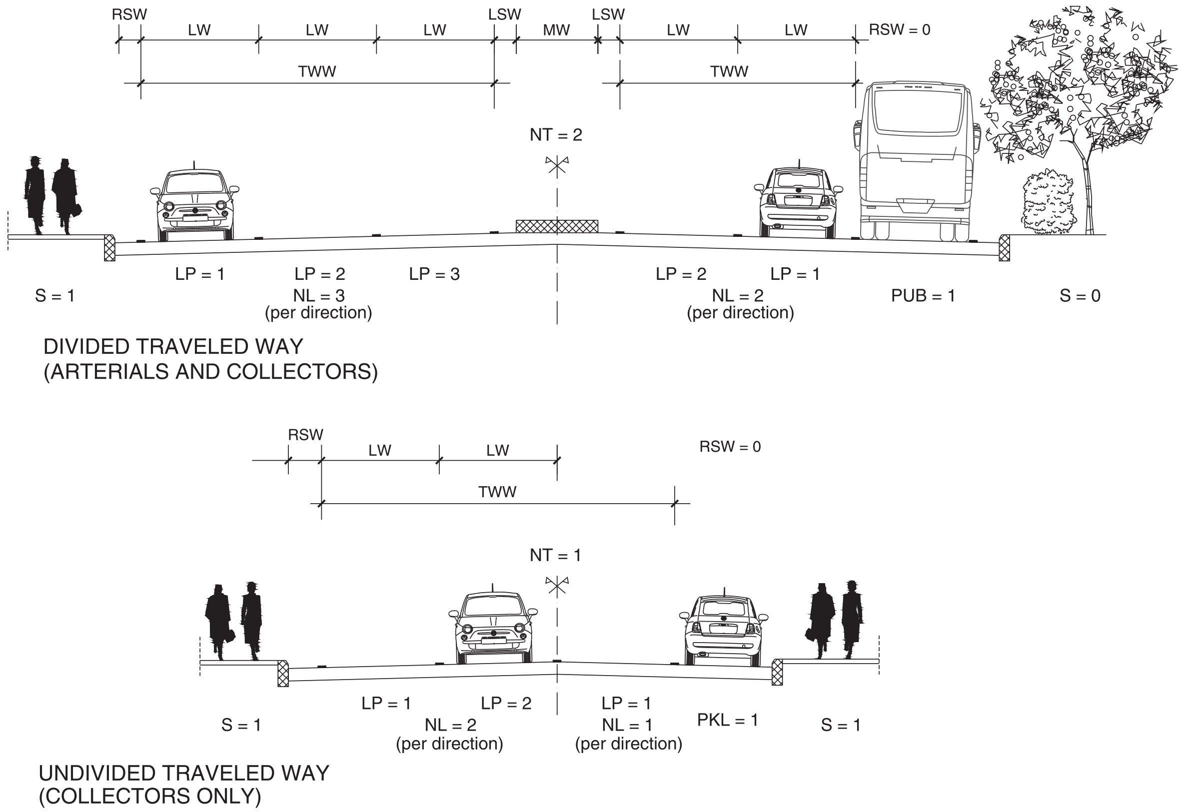



**Fig. 3.** Transversal section parameters for divided and undivided travelled ways.

**Table 2**
Summarized statistics of considered variables.

| # | Variable | Symbol | Type | Unit | min. | max. | $\mu$ | $\sigma$ | Frequency | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Lane position | LP | ND | – | 1.0 | 3.0 | 1.6 | 0.778 | 76 | 100% |
| 2 | Posted speed limit | PSL | ND | km/h | 30.0 | 70.0 | 56.6 | 12.707 | 76 | 100% |
| 3 | No. of travelled ways | NT | ND | – | 1.0 | 2.0 | 1.7 | 0.462 | 76 | 100% |
| 4 | Travelled way width | TWW | NC | m | 7.0 | 17.0 | 10.2 | 2.626 | 76 | 100% |
| 5 | No. of lanes per direction | NL | ND | – | 1.0 | 3.0 | 2.3 | 0.836 | 76 | 100% |
| 6 | Lane width | LW | NC | m | 2.8 | 5.9 | 3.6 | 0.774 | 76 | 100% |
| 7 | Median width | MW | NC | m | 0.0 | 9.1 | 3.9 | 3.753 | 53 | 70% |
| 8 | Right shoulder | RS | B | – | 0 | 1 | – | – | 24 | 32% |
| 9 | Right shoulder width | RSW | NC | m | 0.0 | 3.0 | 0.3 | 0.695 | 24 | 32% |
| 10 | Left shoulder | LS | B | – | 0 | 1 | – | – | 20 | 26% |
| 11 | Left shoulder width | LSW | NC | m | 0.0 | 1.3 | 0.1 | 0.284 | 20 | 26% |
| 12 | Curvature | $1/R$ | NC | $m^{-1}$ | 0.0 | $4.02 \cdot 10^{-3}$ | $5.26 \cdot 10^{-4}$ | $1.17 \cdot 10^{-3}$ | 19 | 25% |
| 13 | Dedicated bus and taxi lane | PUB | B | – | 0 | 1 | – | – | 4 | 5% |
| 14 | Deviation | Dev | B | – | 0 | 1 | – | – | 37 | 49% |
| 15 | Deviation density | DevD | NC | No./km | 0.0 | 8.0 | 2.2 | 2.741 | 37 | 49% |
| 16 | Driveways | D | B | – | 0 | 1 | – | – | 39 | 51% |
| 17 | Driveway density | DD | NC | No./km | 0.0 | 18.2 | 4.1 | 6.206 | 39 | 51% |
| 18 | Intersections | I | B | – | 0 | 1 | – | – | 20 | 26% |
| 19 | Intersection density | ID | NC | No./km | 0.0 | 10.0 | 1.5 | 2.983 | 20 | 26% |
| 20 | Sidewalk | S | B | – | 0 | 1 | – | – | 56 | 74% |
| 21 | Pedestrian crossing | Ped | B | – | 0 | 1 | – | – | 39 | 51% |
| 22 | Pedestrian crossing density | PedD | NC | No./km | 0.0 | 18.2 | 4.3 | 4.789 | 39 | 51% |
| 23 | Parking lanes | PKL | B | – | 0 | 1 | – | – | 16 | 21% |
| 24 | Traffic calming devices | TCD | B | – | 0 | 1 | – | – | 21 | 28% |

- Dev and DevD indicate the presence and the density of deviations respectively (dedicated lanes for leaving or entering the main travelled way),
- D and DD represent the presence and the density of driveways respectively feeding into the travelled way, and
- I and ID denote the presence and the density of intersections respectively with the considered travelled way.

In the same table adjacent to the domain of variation (minimum and maximum values), $\mu$ denotes the average value and $\sigma$ represents the standard deviation of all the numerical variables considered. The frequency expresses the number of times in which the variable has been included in the model with any value other than zero, and it is indicated in both absolute terms (the maximum being equal to 76 i.e. the number of lanes investigated) and in percentage form.

The selection of variables was made while avoiding problems of multi-collinearity which can arise when travelled way width (TWW), NL and LW are simultaneously selected as significant variables. However, it should be noted that NL indicates the number of lanes per direction, so this excludes multi-collinearity for collectors. Moreover, in most of the sections the lane width is not uniformly distributed along the lanes. The contemporary presence of such variables will be discussed later.

In the preliminary phase, the variables were selected by reference to results reported in literature. Of these variables, some were rejected due to the difficulties encountered with their identification, as in the case of land use parameters considered in (Wang et al., 2006).

In fact, the land use around the road segments studied is not homogeneous since commercial activities are present but dispersed throughout residential areas. Moreover, the arterials are located in areas with mixed residential and recreational use in which it is not always clear which type of activity is predominant over the other.

Regarding the variables, LP was considered in an initial phase of our analysis by comparing those lanes which were part of the same travelled way. The results in Fig. 4 relate to the lanes of the same travelled way of road segment No. 1, and demonstrate that there is a significant speed difference between such lanes. Although the Italian Highway Code (Ministero delle Infrastrutture e dei Trasporti, 1992) permits the overtaking manoeuvre on both sides on urban roads, most drivers on these roads still tend to consider the right lane as the "slow" lane, while the left lane is considered the "fast" lane.

This means that most drivers tend to adopt the same behaviour on both urban roads and rural multilane highways. Regarding the latter case, it is worth noting that the driving regulations adopted in continental European countries permit overtaking of slower vehicles on the left side only.

The authors also noted that, in some cases, the difference in speed between lanes was amplified by the presence of a sidewalk near the right lane. Thus, drivers tend to exercise more caution in the first lane (LP = 1) than they do in other positions on the travelled way (LP $\geqslant$ 2). In the example given in Fig. 4, 65.6% of drivers moving along the first lane respect the PSL of 70 km/h, while this percentage decreases to 25.9% in the second lane (LP = 2), and to 7.2% in the third lane (LP = 3).

## 4.4. Data treatment

The speed database consisted of 5339 values observed on 76 lanes (l), 25 single travelled ways, 16 sections (s), and 8 roads (r) (Table 1). The data associated with each lane were, firstly, analyzed in order to verify the homogeneity of the sample and, secondly, to assess the compliance with normal distribution. Only data groups that passed these two tests were subsequently used for the model calibration. The random fluctuation of speeds in the space domain across their average value was assessed by plotting the data on control diagrams in order to detect any undesired trend in the speed data. Fig. 5 shows an example pertaining to the 2nd lane of road segment No. 5. All the groups were found to be homogeneous; hence, all the samples were subjected to a subsequent statistical analysis by means of the $\chi^2$ test.

In 75 of 76 cases the test revealed the samples to be normally distributed:

$$V_{l,i} \sim N(m_l, s_l^2) \tag{3}$$

where $V_{l,i}$ indicates the sample of speed data associated with a generic lane (l). As a result, 75 lanes were considered in the model calibration. Fig. 6 shows the results obtained with the 2nd lane of road segment No. 5.

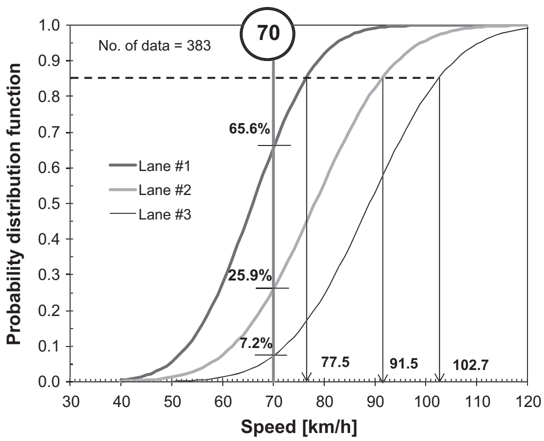



**Fig. 4.** Cumulative distribution of speeds in the three lanes of road segment No. 1.

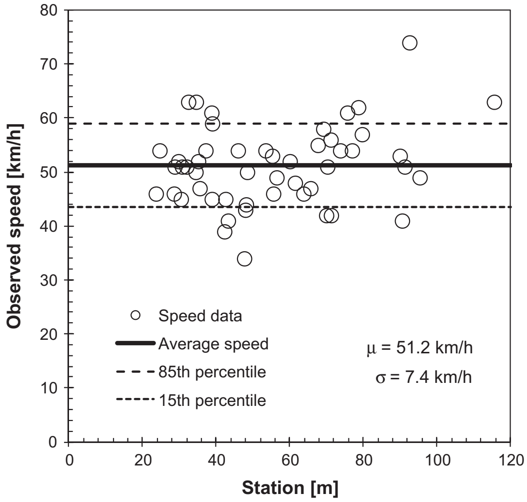
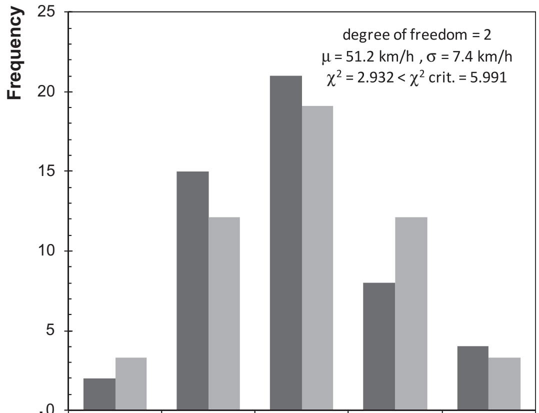



**Fig. 5.** Visual pattern of the homogeneity for road segment No. 5, 2nd lane.




| Speed intervals | < 40 | 40.0 - 47.4 | 47.4 - 54.9 | 54.9 - 62.3 | > 62.3 |
|---|---|---|---|---|---|
| Observed | 2 | 15 | 21 | 8 | 4 |
| Expected | 3.3 | 12.1 | 19.1 | 12.1 | 3.3 |

**Fig. 6.** Synthesis of results of $\chi^2$ test for road segment No. 5, 2nd lane.

### 4.5. Modelling

In this paper, as mentioned in paragraph 3, the speed model proposed by Figueroa and Tarko (2005) was considered for the separate (albeit computed in one step) modelling of both the central tendency of speeds and all the deviations (percentiles) of individual speeds. Furthermore, as a result of the design of this observational study and the characteristics of urban roads, each speed data is a speed percentile for a single lane, on some segments and on some routes, thus the sample dataset consists of a hierarchy of different populations the differences in which may be attributed to this data structure. In accordance with our experimental design, all lanes were observed in randomly selected sections ($a_{l|s}$), which in turn were randomly selected from the total number of sections in each road ($a_{s|r}$), and roads from the entire set of roads in the urban network of Torino ($a_r$).

To illustrate this hierarchy, the following subscripts have been adopted:

- $r$, for the specific road, with $0 \leqslant r \leqslant R$ ($R$ represents the total number of observed roads – equal to 8 in Table 1);
- $s$, for the section within the road, with $0 \leqslant s \leqslant S$ ($S$ represents the number of observed sections – equal to 16 in Table 1);
- $l$, for the lane of a section belonging to a specific road, with $0 \leqslant l \leqslant L$ ($L$ is the number of observed lanes, equal to 75); and
- $i$, for the generic observation.

Consequently, random effects have been included in the statistical model to remove the dependency between any estimation errors of individual observations. The authors included three "a" terms to evaluate the existence of random effects (which measure the random difference between the speed prediction for all roads and the corresponding prediction for a specific road).

The three random effects have been considered to be normally distributed according to the following equations:

$$a_r \sim N(0, \sigma_r^2) \tag{4}$$

$$a_{s|r} \sim N(0, \sigma_s^2) \tag{5}$$

$$a_{l|s} \sim N(0, \sigma_l^2) \tag{6}$$

The dependent variable ($V_{rsl,i}$) is then derived from a random effect (RE) model as follows:

$$V_{rsl,i} = \underbrace{\underbrace{\beta_o}_{\text{intercept}} + \underbrace{\sum_{k=1}^{K} \beta_k^C \cdot X_{ki}}_{\text{central tendency term}} + \underbrace{\sum_{j=1}^{J} \beta_j^D \cdot (Z_p \cdot X_{ji})}_{\text{dispersion term}}}_{\text{model}} + \underbrace{a_r + a_{s|r} + a_{l|sr}}_{\substack{\text{normally distributed}\\ \text{random effects}}} + \underbrace{\varepsilon_{rsl,i}}_{\substack{\text{normally distributed}\\ \text{error}}} \tag{7}$$

in which $\varepsilon_{rsl,i}$ is the error/bias associated with each measurement, $\beta_0$ is the general model intercept, $\beta_k^C$ and $\beta_j^D$ are calibration parameters for the variables affecting the estimated mean ($X_k$), and the estimated standard deviation ($X_j$) respectively, and $Z_p$ is the standardized normal variable. If the three random effects ($a_r$, $a_{s|r}$, $a_{l|sr}$) are excluded from Eq. (7), the model becomes a fixed effect (FE) model in line with the most common approach followed in literature. In Eq. (7) the following additional subscripts have been adopted:

- $p$, for the selected percentile;
- $k$, for the number of significant variables affecting the central tendency ($X_k$), with $1 \leqslant k \leqslant K$, where $K = 24$ as reported in Table 2; and
- $j$, for the number of significant variables affecting the deviation from the mean ($X_j$), with $1 \leqslant j \leqslant J$, where $J = 25$ (variables listed in Table 2 with $Z_p$).

For the calibration, each observation has been associated with the corresponding percentile $p$ ($Z_p$) derived from the lane speed distribution identified by its mean and standard deviation; in particular, $Z_j = 0$ when $p = 50\%$, and $Z_j = 1.036$ when $p = 85\%$.

## 5. Model calibration and results

The data analysis was carried out using R-software version 3.0.2 (R Development Core Team., 2005). Results were obtained via the implementation, in dedicated code, of the Bayesian Information Criterion (BIC) (Schwarz, 1978), which identifies the most significant variables from those selected as possible covariates. The model with the lowest BIC function value ($f_{\text{BIC}}$) calculated according to Eq. (8) has to be preferred:

$$f_{\text{BIC}} = -2 \cdot \widehat{L} + k \cdot \ln(n) \tag{8}$$

where $\hat{L}$ is the maximized value of the log-Likelihood function, n the number of observations, and $P$ is the number of parameters included in the model:

$$P = 1 + K + J \tag{9}$$

where the values 1, $K$ and $J$ denote the size of $\beta_0$, $\beta_k^C$ and $\beta_j^D$ respectively. According to Schwarz (1978), the variables that contribute to the minimization of the BIC function are the most significant and should be selected in the model.

To highlight the contribution made by random effects, three different calibration strategies (CS) have been pursued:

(a) CS#1. This is a FE model where a simple multiple regression analysis is performed by including all the variables selected according to the BIC criterion. CS#1 does not consider the contribution of possible random effects (RE);
(b) CS#2. This is a RE model where an Analysis of Covariance (ANCOVA) is performed using the Restricted Maximum Likelihood (REML) algorithm (McCulloch et al., 2008). The set of variables calibrated is the same set selected in CS#1 (in this case the final value of the BIC function has been derived for comparison); and
(c) CS#3. The ANCOVA method and the REML algorithm are used to calibrate the RE model, but the variables are selected according to the RE model specification using, once again, the BIC criterion.

To run the REML algorithm, the “lme4” package was implemented in the R code (R Development Core Team, 2005). The results obtained from the application of these three strategies are summarized in Table 3. Table 4 reports the synthesis of the statistical performance measurements for the regression analysis, and the quantification of random effects according to Eqs. (4)–(6) for CS#2 and #3.

## 6. Discussion

The operating speed models derived with the three different strategies are characterized by good coefficients of determination and a low level of standard errors. From the results reported in Tables 3 and 4 it appears that:

(a) the use of the RE model (Eq. (7)) leads to a better fit between observed and predicted speed values (Fig. 7);
(b) in the case of CS#1, the use of the BIC criterion results in the same set of significant variables as that derived using the *p*-value (in Table 3 the right shoulder variable - RS - is the only one that exhibits a *p*-value greater than that corresponding to the 95% confidence interval;
(c) the RE model reduces the number of significant variables from 33 (in CS#1 and #2) to 28 in (CS#3);
(d) all the 24 variables listed in Table 2 are significant for all the calibration strategies (#1, #2, and #3);
(e) the number of significant variables affecting the central tendency of the model are 23 for CS#1 and #2, but only 3 for CS#3;
(f) the number of significant variables affecting the dispersion of the data are 10 for CS#1 and #2, and 25 for CS#3.

**Table 3**
Model coefficients and significant variables.

| # | Variable | Calibration strategy #1 | | | | Calibration strategy #2 | | | Calibration strategy #3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Std. error | *t*-Value | Pr (>\|t\|) | Estimate | Std. error | *t*-Value | Estimate | Std.error | *t*-Value |
| – | Intercept | 89.63 | 11.49 | 7.80 | $<2 \cdot 10^{-16}$ | 60.27 | 114.10 | 0.53 | 26.03 | 13.52 | 1.93 |
| 1 | LP | 6.20 | 0.10 | 59.26 | $<2 \cdot 10^{-16}$ | 6.23 | 1.06 | 5.90 | 6.05 | 1.02 | 5.96 |
| 2 | PSL | 0.05 | 0.02 | 2.18 | 0.029 | 0.04 | 0.26 | 0.14 | – | – | – |
| 3 | NT | 41.39 | 2.56 | 16.18 | $<2 \cdot 10^{-16}$ | 24.50 | 32.00 | 0.77 | 16.81 | 7.75 | 2.17 |
| 4 | TWW | −0.58 | 0.29 | −1.99 | 0.047 | −0.03 | 3.04 | −0.01 | – | – | – |
| 5 | NL | −6.42 | 1.19 | −5.42 | $6.31 \cdot 10^{-8}$ | −4.21 | 12.60 | −0.33 | – | – | – |
| 6 | LW | −7.53 | 0.55 | −13.73 | $<2 \cdot 10^{-16}$ | −4.71 | 5.75 | −0.82 | – | – | – |
| 7 | MW | −7.32 | 0.46 | −15.75 | $<2 \cdot 10^{-16}$ | −2.94 | 5.51 | −0.53 | – | – | – |
| 8 | RS | −3.16 | 1.80 | −1.76 | 0.078 | −20.11 | 20.86 | −0.96 | – | – | – |
| 9 | RSW | −7.80 | 1.03 | −7.57 | $4.53 \cdot 10^{-14}$ | 5.29 | 13.08 | 0.40 | – | – | – |
| 10 | LS | −38.48 | 3.25 | −11.84 | $<2 \cdot 10^{-16}$ | −7.86 | 37.64 | −0.21 | – | – | – |
| 11 | LSW | 19.82 | 2.49 | 7.97 | $1.97 \cdot 10^{-15}$ | 1.78 | 27.50 | 0.06 | – | – | – |
| 12 | $1/R$ | −2553.00 | 807.60 | −3.16 | 0.001 | 3230.00 | 8814.00 | 0.37 | 5312.00 | 2432.00 | 2.18 |
| 13 | PUB | 11.95 | 0.77 | 15.52 | $<2 \cdot 10^{-16}$ | 5.24 | 9.70 | 0.54 | – | – | – |
| 14 | Dev | 6.64 | 0.50 | 13.28 | $<2 \cdot 10^{-16}$ | 6.02 | 4.77 | 1.26 | – | – | – |
| 15 | DevD | – | – | – | – | – | – | – | – | – | – |
| 16 | D | 5.17 | 0.72 | 7.20 | $6.98 \cdot 10^{-13}$ | 3.97 | 8.83 | 0.45 | – | – | – |
| 17 | DD | −0.93 | 0.05 | −17.42 | $<2 \cdot 10^{-16}$ | −0.89 | 0.62 | −1.44 | – | – | – |
| 18 | I | −1.32 | 0.38 | −3.47 | 0.0005 | −1.59 | 3.39 | −0.47 | – | – | – |
| 19 | ID | −0.72 | 0.08 | −9.06 | $<2 \cdot 10^{-16}$ | −0.47 | 0.98 | −0.48 | – | – | – |
| 20 | S | −8.84 | 1.39 | −6.35 | $2.34 \cdot 10^{-10}$ | −3.96 | 14.60 | −0.27 | – | – | – |
| 21 | Ped | −27.72 | 2.35 | −11.81 | $<2 \cdot 10^{-16}$ | −4.46 | 28.93 | −0.15 | – | – | – |
| 22 | PedD | −0.64 | 0.10 | −6.51 | $8.47 \cdot 10^{-11}$ | −0.44 | 1.15 | −0.39 | – | – | – |
| 23 | PKL | 7.65 | 0.61 | 12.46 | $<2 \cdot 10^{-16}$ | 1.70 | 7.66 | 0.22 | – | – | – |
| 24 | TCD | 5.98 | 0.38 | 15.75 | $<2 \cdot 10^{-16}$ | 2.60 | 4.88 | 0.53 | – | – | – |
| 25 | $Z_p$ | – | – | – | – | – | – | – | 16.46 | 2.45 | 6.71 |
| 26 | $Z_p \cdot$ LP | 1.09 | 0.09 | 12.15 | $<2 \cdot 10^{-16}$ | 0.84 | 0.02 | 35.52 | 0.74 | 0.02 | 33.23 |
| 27 | $Z_p \cdot$ PSL | 0.10 | 0.01 | 18.04 | $<2 \cdot 10^{-16}$ | 0.11 | 0.00 | 76.69 | 0.01 | 0.00 | 1.39 |
| 28 | $Z_p \cdot$ NT | – | – | – | – | – | – | – | 12.07 | 0.55 | 22.06 |
| 29 | $Z_p \cdot$ TWW | −0.12 | 0.04 | −3.11 | 0.001 | −0.11 | 0.01 | −10.88 | −0.19 | 0.06 | −3.11 |
| 30 | $Z_p \cdot$ NL | – | – | – | – | – | – | – | 0.18 | 0.25 | 0.73 |
| 31 | $Z_p \cdot$ LW | 0.98 | 0.13 | 7.44 | $1.13 \cdot 10^{-13}$ | 0.91 | 0.03 | 26.26 | −1.08 | 0.12 | −9.24 |
| 32 | $Z_p \cdot$ MW | – | – | – | – | – | – | – | −2.85 | 0.10 | −28.59 |
| 33 | $Z_p \cdot$ RS | – | – | – | – | – | – | – | 6.14 | 0.39 | 15.94 |
| 34 | $Z_p \cdot$ RSW | 0.44 | 0.10 | 4.54 | $5.86 \cdot 10$–$^{6}$ | 0.35 | 0.03 | 13.75 | −5.19 | 0.22 | −23.60 |
| 35 | $Z_p \cdot$ LS | – | – | – | – | – | – | – | −19.65 | 0.69 | −28.47 |
| 36 | $Z_p \cdot$ LSW | – | – | – | – | – | – | – | 13.54 | 0.53 | 25.62 |
| 37 | $Z_p \cdot 1/R$ | – | – | – | – | – | – | – | −3305.00 | 172.40 | −19.17 |
| 38 | $Z_p \cdot$ PUB | 2.21 | 0.40 | 5.51 | $3.68 \cdot 10^{-8}$ | 2.23 | 0.11 | 21.24 | 4.36 | 0.17 | 25.73 |
| 39 | $Z_p \cdot$ Dev | 2.65 | 0.29 | 9.11 | $<2 \cdot 10^{-16}$ | 2.19 | 0.08 | 28.66 | 1.87 | 0.13 | 14.42 |
| 40 | $Z_p \cdot$ DevD | −0.32 | 0.05 | −6.14 | $8.72 \cdot 10^{-10}$ | −0.26 | 0.01 | −18.41 | −0.11 | 0.01 | −7.65 |
| 41 | $Z_p \cdot$ D | – | – | – | – | – | – | – | 1.97 | 0.15 | 12.85 |
| 42 | $Z_p \cdot$ DD | – | – | – | – | – | – | – | −0.09 | 0.01 | −7.65 |
| 43 | $Z_p \cdot$ I | – | – | – | – | – | – | – | −0.79 | 0.08 | −9.69 |
| 44 | $Z_p \cdot$ ID | – | – | – | – | – | – | – | 0.08 | 0.02 | 4.41 |
| 45 | $Z_p \cdot$ S | – | – | – | – | – | – | – | −3.31 | 0.30 | −11.12 |
| 46 | $Z_p \cdot$ Ped | – | – | – | – | – | – | – | −11.64 | 0.50 | −23.16 |
| 47 | $Z_p \cdot$ PedD | – | – | – | – | – | – | – | −0.15 | 0.02 | −6.93 |
| 48 | $Z_p \cdot$ PKL | −0.82 | 0.26 | −3.14 | 0.001 | −0.91 | 0.07 | −13.03 | 1.19 | 0.14 | 8.78 |
| 49 | $Z_p \cdot$ TCD | −0.41 | 0.20 | −2.06 | 0.039 | 0.25 | 0.05 | 4.69 | 1.52 | 0.08 | 18.22 |

**Table 4**
Synthesis of statistical analysis.

| | | | | |
|---|---|---|---|---|
| *Calibration strategy #1* | | | | |
| Residual standard error: 4.949 on 5305 DoF | | | | |
| Multiple $R^2$ corrected: 0.9338 | | | | |
| *F*-statistic: 2266 on 33 and 5305 DoF, *p*-value: <2.2 × $10^{-16}$ | | | | |
| Residuals: | | | | |
| Min | 1Q | Median | 3Q | Max |
| −15.821 | −2.598 | 0.255 | 2.792 | 20.365 |
| | | Variance | | Std. dev. |
| Residual | | 6.436 | | 2.537 |
| *Calibration strategy #2* | | | | |
| BIC function at convergence: 18339.96 | | | | |
| Multiple $R^2$ corrected: 0.9956 | | | | |
| Random effects: | | | | |
| Group name | | Variance | | Std. dev. |
| Lane:(Section:Road) (Intercept) | | 33.936 | | 5.825 |
| Section:Road Intercept) | | 29.110 | | 5.395 |
| Road Intercept) | | 0.000 | | 0.000 |
| Residual | | 1.658 | | 1.288 |
| *Calibration strategy #3* | | | | |
| BIC function at convergence: 16351.56 | | | | |
| Multiple $R^2$ corrected: 0.9970 | | | | |
| Random effects: | | | | |
| Group name | | Variance | | Std. dev. |
| Lane:(Section:Road) Intercept) | | 32.749 | | 5.723 |
| Section:Road (Intercept) | | 21.037 | | 4.587 |
| Road (Intercept) | | 92.985 | | 9.643 |
| Residual | | 1.109 | | 1.053 |

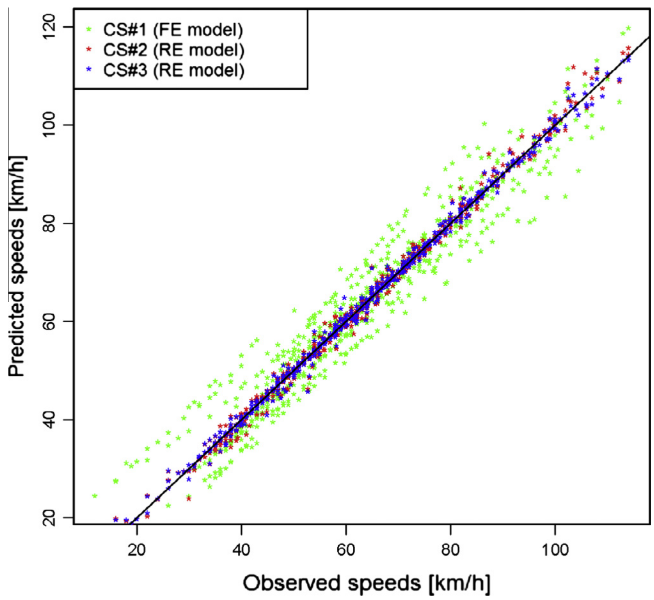



**Fig. 7.** Comparison of observed and modelled speed data on a randomly selected sample of 750 data from the complete database (5339 data).

Fig. 7 shows the comparison between the observed and predicted speeds for a reduced number of randomly selected observations (10 per lane). In the plot it is noticeable that RE models (CS#3, blue points or CS#2, red points), produce much better results than the fixed effect model (CS#, green points).

It should be pointed out that the inclusion of random effects has been effective for a subset of elements. Fig. 8 reports the improvements that occur when the prediction of speeds is made according to CS#3 instead of CS#1, for the data corresponding to four specific lanes. A similar improvement is evident in Fig. 9 where the comparison between observed and predicted values has been illustrated by selecting the speed data for section #10 and road #1 (Fig. 9A and B respectively).

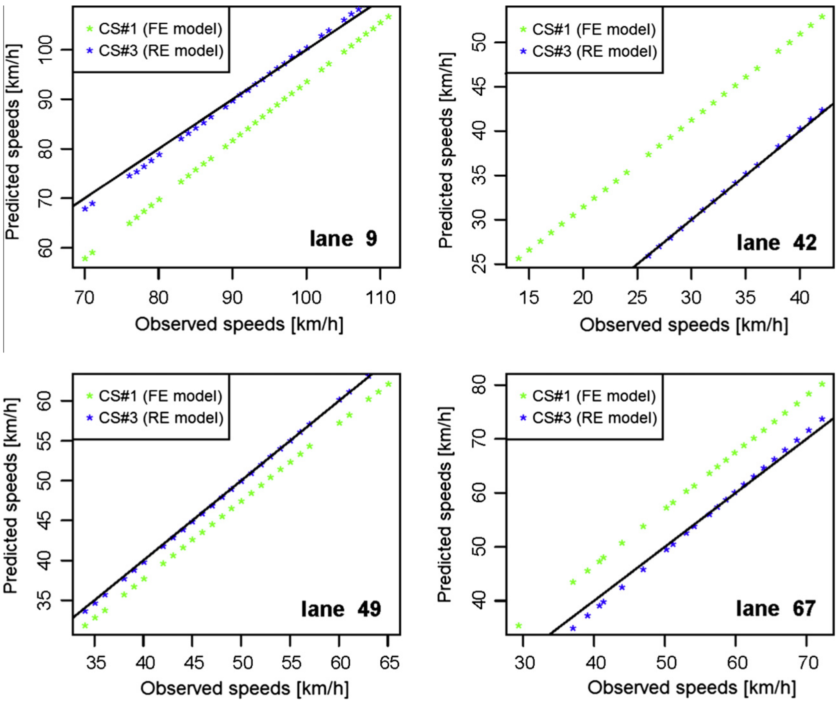



**Fig. 8.** Comparison of observed and modelled speed data on the 4 lanes where the random effects have produced the greatest benefits in terms of speed prediction.

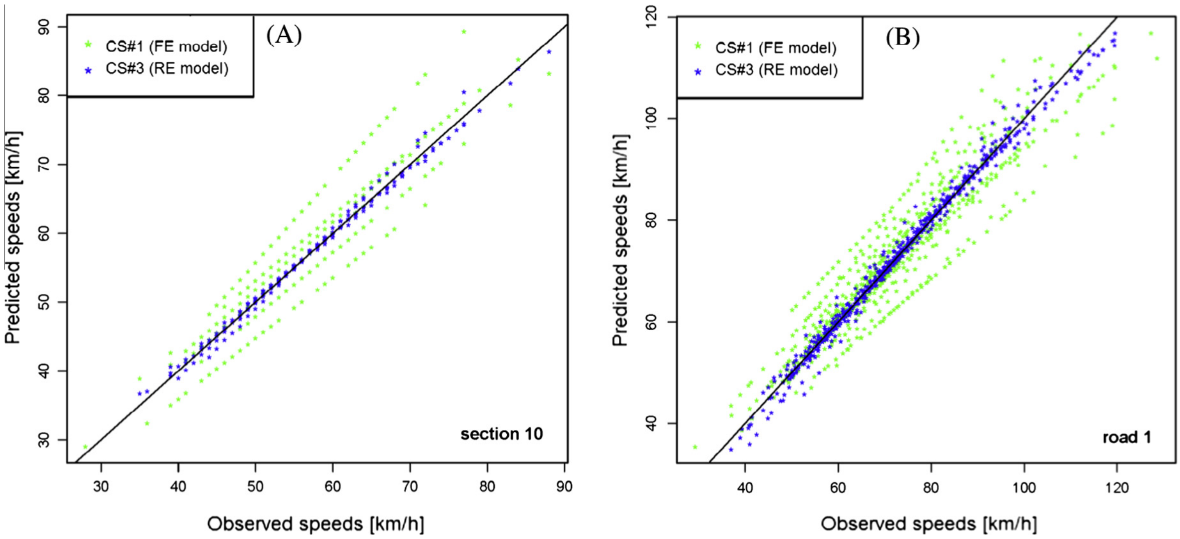



**Fig. 9.** Comparison of observed and modelled speed data on the section (A) and the road (B) where the random effects have produced the greatest benefits in terms of speed prediction.

Similarly Fig. 10 reports the case of 4 lanes, Fig. 11 shows the case of one section and one road, where the random effects produced only a marginal improvement in the model predictions.

To assess the contribution of the variables in each calibration strategy, we report in Fig. 12 the value ranges obtained when multiplying the coefficients $\beta_0$, $\beta_k^C$ and $\beta_j^D$ (Table 3) by the minimum and maximum values of the variables listed in Table 2 (in the case of the normal standardized variable $Z_p$ the extreme values adopted are ±2). From an analysis of Fig. 12 it can be derived that:

(a) the contribution of the normal standardized variable ($Z_p$) is captured by the RE model calibrated according to the strategy CS#3; moreover, $Z_p$ is significant only in this case;

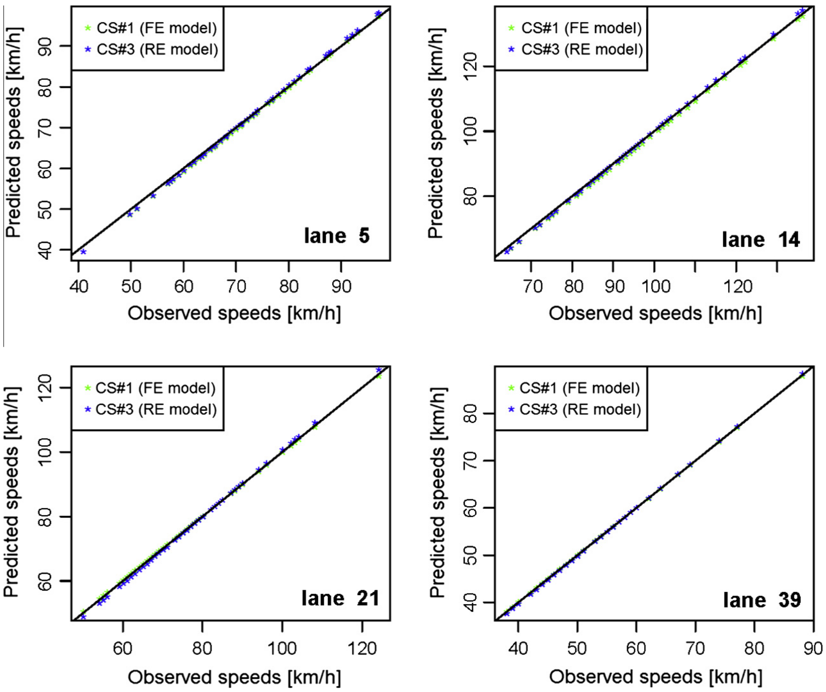



**Fig. 10.** Comparison of observed and modelled speed data on the 4 lanes where the random effects have produced the least benefit in terms of speed prediction.

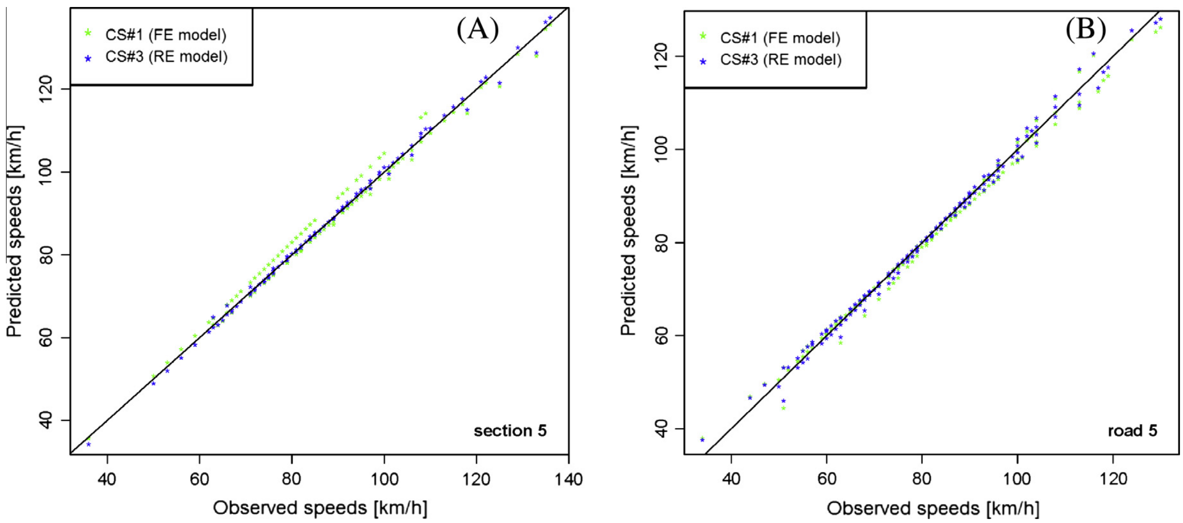



**Fig. 11.** Comparison of observed and modelled speed data on the section (A) and the road (B) where the random effects have produced the least benefit in terms of speed prediction.

(b) in the CS#3 case, the weight of the dispersion term acquires a prominent role to the detriment of the central tendency term (in CS#3 all the geometric and operational variables listed in Table 2 are significant when combined with the normalized standard variable);
(c) the $\beta_k^C$ coefficient associated with LP does not change significantly with a change in the calibration strategy, passing from 6.23 (CS#1) to 6.05 (CS#3);
(d) the number of travelled ways (NT) significantly affects the central tendency of the FE model, while it contributes both to the central tendency and dispersion in the RE model calibrated according to CS#3;
(e) PSL is not included in the RE model calibrated according to CS#3, while its effect is negligible in the FE model;

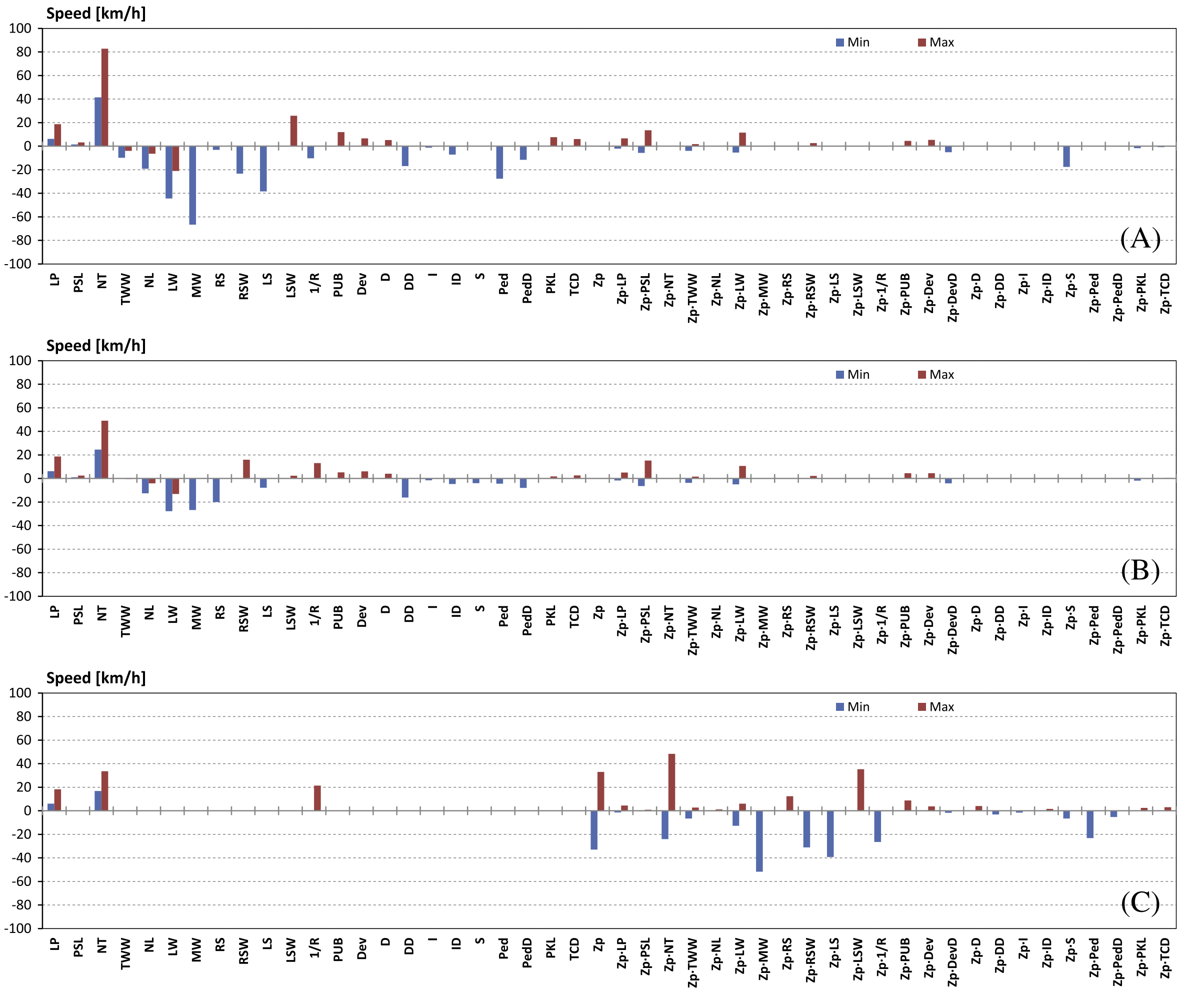



**Fig. 12.** Influence range of considered variables for CS#1 (A), CS#2 (B), and CS#3 (C).

(f) the curvature coefficient ($1/R$) is negative in the FE model (CS#1), while it becomes positive in the RE model (CS#2 and CS#3); in CS#3 it affects the central tendency term ($1/R$) with a positive coefficient and the dispersion term ($Z_p \cdot 1/R$) with a negative coefficient, thus mitigating the effects of the curvature in the final result;
(g) the density of deviations (DevD) does not affect the central tendency term in both the FE and RE models.

By limiting the analysis to the results obtained with the RE model (Fig. 12C), the following additional comments may be included:

(a) the presence and the width of both left (LS and LSW) and right (RS and RSW) shoulders contribute to speed dispersion, even though their combined effects tend to reduce the effects of individual variables;
(b) the presence of a dedicated bus and taxi lane (Pub) close to the travelled way increases the dispersion of speeds;
(c) the presence of a sidewalk (S) and pedestrian crossing (Ped and PedD) makes a significant contribution to the reduction in speed dispersion;
(d) traffic calming devices (TCD) and parking lanes (PKL) make a small contribution to speed dispersion.

## 7. Conclusions

The main objective of this study was to identify the variables influencing central tendency and dispersion of operating speeds on urban arterials and collector roads. The survey was performed on these two road typologies and on a sample set of roads selected from the road network of Torino (Italy). The data has a hierarchical structure, each speed data is a speed percentile for a single lane, on some segments and on some routes. This data collection effort led to the creation of a robust database which was, firstly, validated from a statistical point of view, and, secondly, processed in order to derive predictive equations capable of explaining the effects of geometric and operational variables on the speeds recorded.

As indicated in literature, operating speed models present great variability even when relating to the same road typology with similar geometric characteristics. This can only be explained by differences in driver behaviour, road geometry, environmental conditions, driving regulations and vehicle type between locations. Hence, every model should only be used where it has been calibrated.

According to previous studies, multiple-linear regression analysis enables analysts to select a restricted set of variables which have a strong correlation with observed speeds. This study proposes the use of the random effect model on panel data in order to take into account the hierarchical nature of the data collected and to remove the dependency existing between the estimation errors of individual observations.

The results obtained demonstrate that in an urban environment, greater importance should be attached to the transversal geometric characteristics and less to longitudinal ones; thus confirming the findings in Garrick and Wang (2005).

From among the parameters affecting speeds, the random effect model calibrated according to CS#3 identifies the lane position (LP) and the number of travelled ways (NT) as the two most influential variables on the mean value of the observed speeds. All the longitudinal and transversal geometric variables considered as covariates affect the dispersion term of the calibrated model, thus confirming and quantifying the operational effects of road geometrics. In the calibrated random effect speed model the effects of the posted speed limit are negligible.

In conclusion, initiatives aimed at regulating traffic speed are essential when seeking a reduction in the incidence of crashes and deaths. The enforcement of speed restrictions with speed cameras and other devices is only one of the numerous measures that are effective on transversal sections where such systems operate. The results derived from this study suggest the possibility of acting on driver perception and behaviour via the adoption of an appropriate mix and size of road elements. In contrast to speed cameras, the effects of such measures may be effectively distributed along the road contributing to the formation of a more consistent and legible urban road environment.

## Acknowledgements

The research work included in this paper has been made possible thanks to the funding support of the Compagnia di Sanpaolo and the Politecnico di Torino with the call titled: ‘‘Bando per il Finanziamento di Progetti di Internazionalizzazione della Ricerca’’, approved with the Rectoral Decree n. 208 of the 24th of May, 2013. The Department of Civil and Environmental Engineering of the University of Maryland (College Park, US) is also acknowledged for hosting Dr. Bassani as a visiting professor during the final part of this research.

## References

Aarts, L., & van Schagen, I. (2006). Driving speed and the risk of road crashes: A review. *Accident Analysis and Prevention, 38*, 215–224.

Agarwal, A. (2004). A comparison of weekend and weekday travel behavior characteristics in urban areas. *Graduate school theses and dissertations*. University of South Florida, US.

American Association of State Highway and Transportation Officials. (2010). *Highway safety manual*. Washington, DC, US. ISBN: 978-1-56051-477-0.

Assum, T., Bjørnskau, T., Fosser, S., & Sagberg, F. (1999). Risk compensation. The case of road lighting. *Accident Analysis and Prevention, 31*, 545–553.

Bassani, M., & Mutani, G. (2012). Effects of environmental lighting conditions on operating speeds on urban arterials. *Transportation Research Record: Journal of the Transportation Research Board, 2298*, 78–87.

Bassani, M., Sacchi, E. (2012). Calibration to local conditions of geometry-based operating speed models for urban arterials and collectors. In *Procedia – Social and behavioral sciences, proceedings of the 5th international SIIV congress*, Vol. 53 (pp. 822–833).

Committee for Guidance on Setting and Enforcing Speed Limits (1998). *Managing speed: Review of current practice for setting and enforcing speed limits*. Special Report No. 254, Transportation Research Board, National Research Council, National Academy Press, ISBN 0-309-06502-X.

European Transport Safety Council (2011). *Road safety target outcome: 100,000 Fewer deaths since 2001*. 5th Road Safety PIN Report, Brussels.

European Transport Safety Council (2012). *A challenging start towards the EU 2020 road safety target*. 6th Road Safety PIN Report, Brussels.

Figueroa, A. M., & Tarko, A. P. (2005). Speed factors on two-lane rural highways in free-flow conditions. *Transportation Research Record: Journal of the Transportation Research Board, 1912*, 39–46.

Fitzpatrick, K., Shamburger, C. B., Krammes, R. A., & Fambro, D. B. (1997). Operating speed on suburban arterial curves. *Transportation Research Record: Journal of the Transportation Research Board, 1579*, 55–60.

Garber, N. J., & Gadiraju, R. (1989). Factors affecting speed variance and its influence on accidents. *Transportation Research Record: Journal of the Transportation Research Board, 1213*, 64–71.

Garrick, N. W., & Wang, J. (2005). New concepts for context-based design of streets and highways. *Transportation Research Record: Journal of the Transportation Research Board, 1912*, 57–64.

Global Road Safety Partnership (2008). *Speed management: A road safety manual for decision-makers and practitioners*. Geneva, Switzerland. ISBN: 978-2-940395-04-0.

Hassan, Y. (2004). Highway design consistency: refining the state of knowledge and practice. *Transportation Research Record: Journal of the Transportation Research Board, 1881*, 63–71.

Hauer, E. (2009). Speed and safety. *Transportation Research Record: Journal of the Transportation Research Board, 2103*, 10–17.

Lamm, R., Psarianos, B., & Mailaender, T. (1999). *Highway design and traffic safety engineering handbook*. New York, NY: McGraw Hill Inc.

Mannering, F. (2009). An empirical analysis of driver perceptions of the relationship between speed limits and safety. *Transportation Research Part F, 12*, 99–106.

McCulloch, C. E., Searle, S. R., & Neuhaus, J. M. (2008). Generalized, Linear, and Mixed Models (2nd). ed. John Wiley & Sons.

Ministero delle Infrastrutture e dei Trasporti (1992). *Nuovo Codice della Strada*. (in Italian) Decreto Legislativo 30 Aprile 1992 n.285, Rome, Italy.

Ministero delle Infrastrutture e dei Trasporti (2001). *Norme Funzionali e Geometriche per la Costruzione delle Strade*. (in Italian) Decreto Ministeriale n.6792, Rome, Italy.

Ministero delle Infrastrutture e dei Trasporti (2011). Conto Nazionale delle Infrastrutture e dei Trasporti. (in Italian) Istituto Poligrafico e Zecca dello Stato S.p.A., Rome, Italy.

Misaghi, P. (2003). *Modeling operating speed and speed differential for design consistency evaluation*. M.S. thesis, Carleton University, Ottawa, Canada.

Misaghi, P., & Hassan, Y. (2005). Modeling operating speed and speed differential on two-lane rural roads. *Journal of Transportation Engineering, 131*, 408–417.

National Cooperative Highway Research Program (2003). *Design speed, operating speed, and posted speed practices*. Report No. 504, Transportation Research Board of the National Academies, Washington, DC, ISBN: 0-309-08767-8.

National Cooperative Highway Research Program (2004). *Road safety audits. A synthesis of highway practice*. Synthesis No. 336, Transportation Research Board of the National Academies, Washington, DC, ISBN: 0-309-07015-5.

Nie, B., Hassan, Y., (2007). Modeling driver speed behavior on horizontal curves of different road classifications. *Presented at 86th annual meeting of the transportation research board, transportation research board of the national academies, 07-0782*.

Transportation Research Board (2010). *Highway capacity manual*. National Research Council, Washington, DC, US. ISBN: 978-0-309-16078 (-79, -80).

Transportation Research Circular (2011). *Modeling operating speed*. E-C151 Synthesis Report, Transportation Research Board of the National Academies, Washington, DC.

Poe, C. M., Mason J. M., Jr. (1995). Geometric design guidelines to achieve desired operating speed on urban streets. In *Proceedings of the 65th ITE annual meeting, institute of transportation engineers* (pp. 70–74).

R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. (www.R-project.org).

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Shinar, D., & Compton, R. (2004). Aggressive driving: an observational study of driver, vehicle, and situational variables. *Accident Analysis and Prevention, 36*, 429–437.

Tarko, A. P. (2009). Modeling drivers' speed selection as a trade-off behavior. *Accident Analysis and Prevention, 41*, 608–616.

Tarris, J. P., Poe, C. M., Mason, J. M., Jr., & Goulias, K. G. (1996). Predicting operating speeds on low-speed urban streets: regression and panel analysis approaches. *Transportation Research Record: Journal of the Transportation Research Board, 1523*, 46–54.

Wang, J., Dixon, K. K., Li, H., & Hunter, M. (2006). Operating speed model for low-speed urban tangent streets based on in-vehicle global positioning system data. *Transportation Research Record: Journal of the Transportation Research Board, 1961*, 24–33.