Prediction Models for Space Mean Speed on Urban Roads

Mariangela Zedda and Francesco Pinna

Abstract The research investigates the relationship between driving behavior and characteristics of the road environment in urban area. This study allows the identification of factors which influence driving speed. The purpose is to develop mathematical models which link driving behavior with infrastructure geometric characteristics. The parameter used to describe driving behavior is space mean speed. This is very important because it considers speed of vehicles traveling a given segment of roadway during a specified period of time and it is calculated using the average travel time and the length for the roadway segment. This speed is used to understand driving behavior in normal traffic flow and in daylight conditions. The data are collected on urban road tangents. These roads have the common characteristic to be single or dual carriageways with two lanes for each direction with a speed limit of 50 km/h. With a multiple linear regression, two models are developed and validated to predict speed. Statistically significant variables include traffic characteristics (flow, number of vehicles entering and leaving traffic stream) and geometric design attributes (lanes width, type of median, tangent length and type of left-lateral obstacle). This study can be useful to both traffic manager and road designers because developed models could implement design guidelines, especially regarding road tangents.

Keywords Urban road · Space mean speed · Multiple linear regression
Model validation · Road design · Road tangent

M. Zedda · F. Pinna (✉)
Department of Civil Environmental Engineering and Architecture,
University of Cagliari, Cagliari, Italy
e-mail: fpinna@unica.it

M. Zedda
e-mail: zedda.mangy@gmail.com

© Springer Nature Singapore Pte Ltd. 2018

N. Chaki et al. (eds.), *Proceedings of International Conference on Computational Intelligence and Data Engineering*, Lecture Notes on Data Engineering and Communications Technologies 9, https://doi.org/10.1007/978-981-10-6319-0_2

1 Introduction

Speed is a parameter which involves two fields closely related: safety and road design. In the first case, it is the element that influences the occurrence and the severity of an accident; in the second case, it is the factor used to design road infrastructure elements. Therefore, speed affects both the probability the accident happens and the injury severity. So, first, speed defines risk level of being involved in an accident: people need time to process information, to decide how to react to situation in order to have a reaction able to avoid the crash. Second, the speed affects severity of accident and injury. In fact, at higher speed, more energy is released colliding with another vehicle, road user, or obstacle. Some road types, such as urban street, have traffic conditions and interaction of multiple urban users more complex than others [1]. This, for example, depends on road infrastructure configuration, absence or presence of pedestrians, cyclists, the type of car, etc. In addition, cause of accidents is due to both road users' lack of respect for traffic code rules, and presence of road infrastructure deficiencies. In fact, roads should induce drivers to adopt speeds which are designed and in most cases this does not happen. This is probably due to the fact that in Italy standard prescribes the use of the design speed in order to size the horizontal alignment elements. The design speed is an ideal speed defined with the aim to design geometric elements, and safety, regularity, and efficiency of vehicles movement depend on these. In Italy, traditional design approach does not consider the real driving behavior in relation to geometric and environmental characteristics of the road infrastructure. So to meet the road safety requirements, it would be appropriate to make the roadway congruent, at project level, with real speeds. For these reasons, the driving behavior is studied in order to develop predictive models that can be used as a tool to improve the Italian standard. To pursue this aim, these models should be validated with data about other urban roads.

2 Literature Review

The focus of the present study is on factors influencing drivers' speed, in normal traffic condition, while traveling on urban street tangents, with special attention on geometric design and urban environment. The models in literature, although provide important results, cannot be considered universally valid. The causes may be differences between one country and another (but also differences between city of the same country) in terms of habits, vehicle fleet, traffic rules, climatic conditions, and orography of the area. For example, the developed models for rural roads cannot be applied to urban areas because drivers assume a different style to travel urban or suburban roads; environment characteristics are often dissimilar; flow, traffic characteristics, and geometric design elements are different. As regards the developed models for urban area, many researchers use speed limits as independent

variable; this cannot be used in the case study because urban roads are distinguished by the same speed limit of 50 km/h, which, moreover, is the only one in Italy for urban areas. Finally, several researchers used instantaneous speed as dependent variable. These speeds are measured at one specific point of roadway and not considered running speed variations adopted by drivers. The behavioral dynamics of drivers in urban areas are highly complex. Perhaps this fact has deterred the construction of mathematical models capable of representing the driving behavior in relation to urban environment characteristics. In Italy, the developed models may concern only the suburban area, while at the international level, there is poor scientific production especially aimed at studies concerning the urban tangents. A number of models are available in literature, but none of these considers urban roads, tangent and space mean speed. In consideration of the above, a general overview of international scientific production, regard to urban roads, is shown below.

Tarris et al. [2] collect vehicle operating speed data along 27 urban collectors in Pennsylvania. Models are developed for individual and aggregate data. In both cases, the variable used to predict operating speed is degree of curve.

Poe and Mason [3] study operating speed on 27 urban collectors in Pennsylvania. The tracking is used to include only free flow passenger cars, which consist of vehicles with time headways of 5 s or more. They introduce a mixed model approach to analyze influence of geometric elements on operating speed. The significant variables are degree of curvature, longitudinal grade, lane width, and roadside characteristics.

Fitzpatrick et al. [4] examine operating speed on 78 urban/suburban sites in Arkansas, Missouri, Tennessee, Oregon, Massachusetts, and Texas. Free flow speed is obtained using time headway of 5 s or more and a tail-way of 3 s or more. The posted limit and access density are found to be statistically significant to predict operating speed.

Wang et al. [5] collect speed data, in tangent sections, using 200 vehicles equipped with GPS. It is established that roadside density, driveway density, intersection density, sidewalk presence, and parking presence are negatively associated with operating speed on urban streets while a number of lanes, curb presence, and commercial and residential land uses are positively associated with operating speed.

Karin F.M. Arosson [6] uses two methods for speed investigation on urban street links. The first approach (micro model) models average travel speed results produced by a microscopic traffic simulation model calibrated by using observed driver behavior. In the second approach (macro model), space mean speed data is modeled for three street link types based on aggregated speed data collected in the field. In particular, urban model shows a relationship between speed and flow, average number of crossing pedestrians and cyclists.

Ali et al. [7] investigate the relationship between free flow speed and geometric variable on 35 urban street segments located in Fairfax County Virginia. The study shows a statistically significant relationship between free flow speed and posted

speed limits, median type, and segment length. Free flow speed is obtained using time headway of 7–8 s or more and a tail-way of 4 or 5 s.

To recap the above discussion, Table 1 summarizes mathematical models.

Table 1 Summary of speed models

Autor (s)	Models	R2
Tarris et al. (1996)	V_{mean} = 53.5 − 0.265DC (aggregate mean speed)	0.82
	V_{mean} = 53.8 − 0.272DC (individual mean speed model)	0.63
Poe e Mason (2000)	V = 57,47 − 0,23 DEGCVR − 3,17 LANWIDN − 1,23 HZRT5 N	–
Fitzpatrick et al. (2003)	V85 = 7,68 + 0,98 PSL	0.90
	V85 = 7,68 + 0,83 PSL-0,05 AD	0.90
Wang et al. (2006)	V85 = 31.564 + (6.491 lane.num) − (0.101 roadside) − (0.051 driveway)-(0.082 intersection) + (3.01 curb) − (4.265 sidewalk)-(3.189 × parking) + (3.312 × land.use1) + (3.273 × land.use2)	0.67
Arosson (2006)	Vobs = 39,8-0,20 Flow-0,24 Ped-5,24 Lanes + 4,73 BicSep-5,54 Park	0.66
Ali et al. (2007)	FFSmean = 39,3 + 8,,6PS45 + 3,7PS40	0.76
	FFS85 = 42,3 + 10,4PS45 + 3,8PS40	0.77
	FFSmean = 37,4 + 6,8PS45 + 2,6PS40 + 13,5SL	0.87
	FFS85 = 37,4 + 8PS45 + 2,1PS40 + 3,6MT + 13SL	0.86
where		
V_{mean}	Mean operating speed	
DC	Degree of curve	
V	Mean speed	
DEGCVR	Degree of curve	
LANWIDN	Lane width	
HZRT5 N	Roadside hazard rating	
V85	85th percentile operating speed	
PSL	Posted speed limit	
AD	Access density	
V85	85th percentile speed	
lane.num	number of lanes	
roadside	density of trees and utility poles divided by their average offset from roadway	
driveway	density of driveways	
intersection	density of T-intersections	
curb	0 if there is no curb; otherwise 1	
sidewalk	0 if there is no sidewalk; otherwise 1	
parking	0 if there is no on-street parking; otherwise 1	

(continued)

Table 1 (continued)

Autor (s)	Models	R2
land.use1	1 if land use is commercial; otherwise 0 (baseline is park and office land use)	
land.use2	1 if land use is residential; otherwise 0 (baseline is park and office land use)	
Vobs	Observed space mean speed	
Flow	Observed average traffic flow in the studied direction of travel expressed in vehicles per 5 min	
Ped	Average number of crossing pedestrians and cyclists (summarized in groups of 5, 15 and 25 people per 15 min and 400 m)	
Lanes	Number of lanes in the studied direction (1 or 2)	
BicSep	Separated bicycle lane (yes = 1; no = 0)	
Park	Roadside parking permitted (yes = 1; no = 0)	
FFSmean	Mean free flow speed	
FFS85	85th percentile free flow speed	
PSi	Posted speed limit (i = 45 mph, 40 mph, 35mph)	
MT	Median type (divided = 1, no median = 0)	
SL	Segment length ratio	

3 Data Collection and Database

The data are collected on seven urban road tangents. These roads have the common characteristic to be single or dual carriageways with two lanes for each direction and to have an urban speed limit of 50 km/h. The preliminary analysis of collected speeds shows that drivers do not respect speed limit: 68.6% of drivers exceeded 50 km/h. Also, operating speed is calculated. The AASHTO defined it as "the speed at which drivers are observed operating their vehicles during free-flow conditions" [8]. The computed value on seven urban roads is 71 km/h.

"Viale Colombo" is not considered in developing models because this road is chosen in order to check the accuracy of speed equations developed in this study. This road has the greatest number of independent variables which are identified on all roads. The principal characteristics of roads are described in Table 2.

The next research step is to identify road section and then to perform monitoring campaign. We start from some observations reported in literature. Fitzpatrick et al. [9] define the section as a portion of a suburban arterial between horizontal curves and/or control devices. The sections selected are at least 200 m from an adjacent horizontal curve and 300 m from adjacent signal or stop sign. In another study, Fitzpatrick et al. [10] establish that to avoid the effect of traffic control devices on vehicle speed, distances between study sections and a signalized intersection should be at least 200 m. Polus et al. [11] identify at least 500 m from the study section and any intersection to eliminate the effect of traffic control devices on vehicle

Table 2 Pattern of road features

Urban road geographic coordinate	Variables					
	Section length (m)	Tangent length (m)	Bus lane (m)	Lane width (m)	Fast lane width (m)	Obstacle right-lateral width (m)
Diaz 39°12′18.75″N 9°7′56.57″E	106	450	3.3	3.0	3	3.25
Lungo Saline 39°12′20.76″N 9° 9′44.11″E	119.2	680	No	3.3	3.3	0.4
Poetto 39°11′59.07″N 9° 8′42.58″E	120.2	450	3.3	3.0	3	3.65
Marconi 39°14′17.92″N 9° 8′40.46″E	135.1	670	No	2.75	2.75	0.3
Elmas 39°14′13.17″N 9° 5′41.78″E	150.0	1000	No	3.75	3.75	0.5
Monastir 39°14′10.75″N 9° 5′49.03″E	143.0	800	No	3.75	3.75	0.3
Colombo 39°12′34.79″N 9° 7′4.40″E	103.0	430	No	3.25	3.25	1.00
	Access points[a]	Travel direction	Type of left-lateral obstacle[b]		Median type[c]	
Diaz 39°12′18.75″N 9° 7′56.57″E	Yes	2	2		1	
Lungo Saline 39°12′20.76″N 9° 9′44.11″E	No	2	2		3	
Poetto 39°11′59.07″N 9° 8′42.58″E	No	2	2		2	
Marconi 39°14′17.92″N 9° 8′40.46″E	Yes	2	1		0	
Elmas 39°14′13.17″N 9° 5′41.78″E	Yes	1	3		–	

(continued)

Table 2 (continued)

Urban road geographic coordinate	Variables					
	Section length (m)	Tangent length (m)	Bus lane (m)	Lane width (m)	Fast lane width (m)	Obstacle right-lateral width (m)
Monastir 39°14′10.75″N 9° 5′49.03″E	Yes	1	0		–	
Colombo 39°12′34.79″N 9° 7′4.40″E	Yes	2	1		1	

[a]*Access points* Elements that allow vehicles to enter and leave traffic stream
[b]*Type of obstacle left-lateral* Type of obstacle next to the fast lane that causes influences on driving behavior. The variable is:
3 if there is not obstacle; *2* if there is median; *1* if lanes in opposite direction are not separate; *0* = presence of parking
[c]*Median type* Three types of central reservation characterize the urban roads in exam. The variable is: *0* if there is not median; *1* if the median, decorated with trees or hedges, allows the U-turn; *2* if the median, decorated with trees or hedges, is continuous; *3* if the median is indicated by road markings

speeds. Selected study segment should be sufficiently distant from traffic control devices so that driver speed on the roadway is not influenced by acceleration and deceleration zones. We use the criteria indicated in Italian Standard [12] to determine stopping visibility distance and zone lengths for vehicles in deceleration or acceleration. Calculated values are less than 200 m and, in view of these considerations, we prefer to use a distance greater than or equal to 200 m between section of the studied tangent and curve, stop sign, traffic lights, and yield signs close of roundabouts. Therefore using this method, study tangent has a length that varies between 100 and 150 m.

The equipment used to detect all information associated to vehicular traffic is the Radar EasyData. It records for each vehicle: date (day/month/year), time (hours/minutes/seconds), spot speed (km/h), vehicle length (meters), and direction of travel. A digital camcorder is used to check if there are any measurement errors. Also, it is used to analyze user behaviors and to identify any external or internal traffic factors able to affect driver behavior. In particular, we use two radars placed at the road section beginning and end. Figure 1 shows the scheme used for the surveys campaign.

The data collection is performed under dry weather conditions. Six hours of data are collected at each road section: four hours of morning traffic between 8.00 and 12.00 a.m. and two hours of afternoon traffic between 4.00 and 6.00 p.m. In particular surveys, campaign is performed in the weekdays at random, excluding

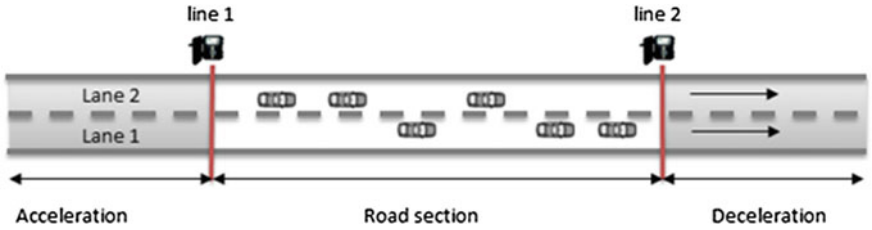


Fig. 1 Schematic representation of the observed road segment with marked lines for vehicle passage time readings

Saturdays, Sundays, public holidays, days of strike by public transport, days with road works and events of various kinds in the streets under study or near the study sections. In the first section (line 1) and in the second section (line 2), the instantaneous speed and the transit times are collected for more than 97,000 vehicles.

The collected data are processed to obtain information about driver behavior and traffic flow characteristics. In particular, database is organized as follows. Vehicles are numerated and for each travel time, vehicle type (car, bus, motorcycles, and heavy vehicles) and numerical code of driver behavior (travel lane, change lane, driving behavior influenced by crossing pedestrian, etc.) are assigned. The anomalies regarding instrument errors and/or behavior drivers that are not the subject of study (influence due to the presence of radar, Traffic Code infractions, etc.) are recorded using numerical code. Once the database is prepared, data are controlled and processed. A first check is to verify that anomalies do not exceed 10% to prevent these can affect the quality of information altering the final result. The space mean speed is calculated using the following equation:

$$\begin{aligned} \bar{v}_{\text{sms}} &= (\text{distance traveled})/(\text{average travel time}) = d/\left(\left(\sum t_i\right)/n\right) \\ &= (n \times d)/\left(\sum t_i\right), \end{aligned} \tag{1}$$

where

v_{sms} space mean speed;
d distance traveled or length of roadway segment;
n number of observations;
t_i travel time of the i-th vehicle.

The space mean speeds are aggregated in 5 min intervals. So, the 5 min intervals are removed if the anomalies exceed 10%.

4 Analysis Method

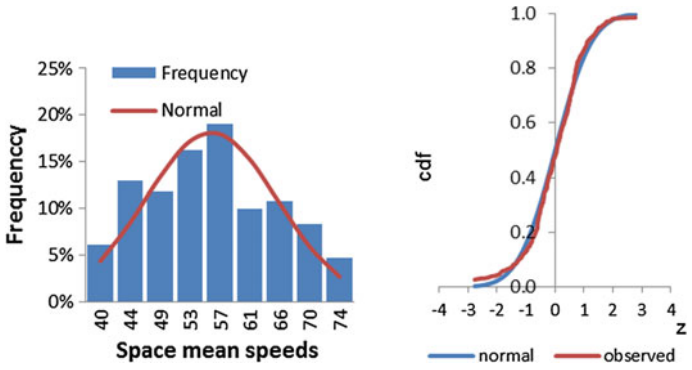
4.1 Data Analysis

The space mean speeds are analyzed with specific statistical tools to obtain more information about the data and to understand if they need further investigation. Then the dependent variable analysis is carried out to test normality assumption. The histogram reveals that vehicle speeds distribution is bell-shaped and approximately symmetric. Kolmogorov–Smirnov test shows that speed data do fit the Gaussian distribution as shown in Fig. 2.

4.2 Regression Models

The process to develop and validate mathematical model is carried out to predict real driver behavior. The method is based on the observation of speeds on urban road tangent. Specifically, a multiple linear regression is used to developed models and stepwise selection method is used to select independent variables. In particular, the procedure begins with no variables in the model and adds a variable in succession according to the criterion of partial F-statistic. At each phase, a variable is added, whose partial F-statistic yields the smallest p-value. Variables are entered as long as the partial F-statistic p-value remains below a 0.05. When the addition of any of the remaining variables yields a partial p-value > 0.05, procedure stops.

This procedure shows that some variables, commonly used in mathematical models, are not statistically significant (p-value > 5%). For example, traffic composition is investigated: vehicles are divided in different categories (motorcycles, cars, trucks, and buses). Probably these variables are not significant because traffic



(a) Histogram of dependent variable (b) Kolmogorov-Smirnov test

Fig. 2 Space mean speed analysis: **a** speed distribution **b** normality test

is homogenous (approximately 95% of vehicles are cars). Also, we consider access density, number of crossing pedestrians and cyclists, presence of sidewalk, curb, and parking. Moreover, posted speed limit does not effect on speed in fact, analyzing collected data, the speed limits are substantially exceeded by drivers: 68.6% of drivers has exceeded 50 km/h.

Six variables are used to develop models:

- flow (F): number of vehicles over road section aggregated in 5 min intervals;
- number of vehicles entering and leaving traffic stream ($F_{in,out}$): these are vehicles that travel partially road section, aggregated in 5 min intervals. This is due to the presence, in urban area, of driveways, parking areas, median which allows the U-turn
- presence of crosswalk (C): 0 no crosswalk, 1 crosswalk;
- tangent length (TL): is not the length of road section but the length of tangent between two intersections in which is located the study section;
- lane widths (LW): is the sum of two lane widths in each direction (travel lane and fast lane);
- the left-lateral obstacle is interpreted in two ways:

  i. type of left-lateral obstacle (O): is the type of obstacle next to the fast lane and causes influences on driving behavior. This variable is used to develop model with collected data on every urban road. The variable is:
  3 if there is no obstacle;
  2 if there is median;
  1 if the lanes in opposite direction are not separate;
  0 = presence of parking;

  ii. type of median (M): the urban roads in exam are characterized by three types of central reservation. This variable is used to develop model with collected data on roads with median. The variable is:
  0 if there is not median;
  1 if the median, decorated with trees or hedges, allows the U-turn;
  2 if the median, decorated with trees or hedges, is continuous;
  3 if the median is indicated by road markings.

Two models are developed as shown in Table 3. The collected data on every urban road are used to develop the first model. The adjusted coefficient of determination (R2) is high, indicating that 81% of the variability is explained. The collected data on roads characterized by two lanes for traffic in each direction are used to develop the second model. This model is important to investigate the type of median. The adjusted R2 value is also high, indicating that 84.5% of the variability is explained.

All variables are significant at the 5% significance level (95% confidence level) for these two models. In other words, p-value is < 0.05 for all independent variables. Variance inflation factor (VIF) is calculated to detect multicollinearity.

Table 3 Multiple linear regression model results

Model 1	Urban road: Diaz, Lungo Saline, Poetto, Marconi, Elmas, Monastir.						
V_{ms} = 14.21 – 0.05 F − 0.64 $F_{in,out}$ – 11.84 C + 0,03 TL + 3,84 O + 4,07LW							
Regression analysis							
Multiple R		0.90					
R2		0.813					
R2 Adjusted		0.810					
Standard Error		4.03					
Observations		362					
ANOVA						Alpha 0.05	
	gdl	SQ	MQ	F		*p*-value	
Regression	6	25200.21	4200.03	258.6		2.76E-126	
Residual	355	5765.21	16.24				
Total	361	30965.42					

	Coeff.	Standard error	Stat t	*p*-value	Lower 95%	Upper 95%	VIF
Intercept	14.21	2.97	4.78	2.5E-06	8.36	20.06	
F	−0.05	0.02	−3.18	1.6E-03	−0.08	−0.02	1.00
$F_{in,out}$	−0.64	0.07	−8.88	3.4E-17	−0.78	−0.50	5.43
C	−11.84	0.76	−15.59	4.0E-42	−13.33	−10.35	3.20
TL	0.03	0.00	11.44	5.1E-26	0.02	0.03	5.91
O	3.84	0.28	13.82	4.5E-35	3.29	4.38	1.14
LW	4.07	0.60	6.83	3.6E-11	2.90	5.25	3.88

(continued)

Table 3 (continued)

Model 2	Urban road: Diaz, Lungo Saline, Poetto, Marconi.					
Model 2	Urban road: Diaz, Lungo Saline, Poetto, Marconi.					
$V_{ms} = 21.80 - 0.60\ F_{in,out} + 4.11\ M + 0.05\ TL$						
Regression analysis						
Multiple R	0.920					
R2	0.846					
R2 Adjusted	0.845					
Standard Error	3.806					
Observations	262					
ANOVA						Alpha 0.05
	gdl		SQ	MQ	F	*p*-value
Regression	3		20583.37	6861.12	473.58	1.4106E-104
Residual	358		3737.87	14.49		
Total	261		24321.23			

	Coeff.	Standard error	Stat t	*p*-value	Lower 95%	Upper 95%	VIF
Intercept	21.80	1.28	16.9	5.8E-44	19.28	24.33	
$F_{in,out}$	−0.60	0.16	−3.7	0.0001	−0.90	−0.29	1.00
M	4.11	0.34	12.2	1.6E-27	3.45	4.77	2.41
TL	0.05	0.00	23.1	6.0E-65	0.05	0.05	1.11

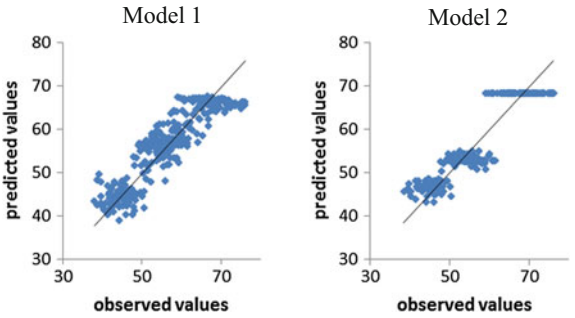


Fig. 3 Scatter plot between observed values and predicted values

All VIF values are less than 10: these indicate that no collinear variables are present in both models.

The observed values are plotted versus predicted values as shown in Fig. 3: most of the points fall close to the 45° line demonstrating a linear relationship between the observed and predicted space mean speeds.

Diagnostic analysis and residual tests are performed to assess adequacy of models. Specifically, linearity, homoscedasticity, independence, and normality are verified [13]. Figure 4 shows that in both models, the following assumptions are satisfied:

- linearity: relationship between the predictors and the outcome variable is linear;
- homoscedasticity (homogeneity of variance): the error variance is constant;
- independence: the errors associated with one observation are not correlated with the errors of any other observation;
- normality: the errors are normally distributed. Kolmogorov–Smirnov test and Quantile–Quantile plot (Q–Q plot) are used to verify the condition of normality.

Leverage values and Cook's distance are calculated to identify the presence of any anomalies observations.

The leverage statistic, h, measures the influence of yi on its predicted value ŷi. The observed value yi is influential if:

$$h_i > (2(k+1))/n, \tag{2}$$

where

h_i is the leverage for the i-th observation;

k is the number of β's in the model (excluding $\beta 0$).

In both models, hi are less than 0.04, Fig. 5 shows that there are not observations which influence regression models.

Cook's distance, D, is another measure of the influence of observations. Cook's distance is defined as:

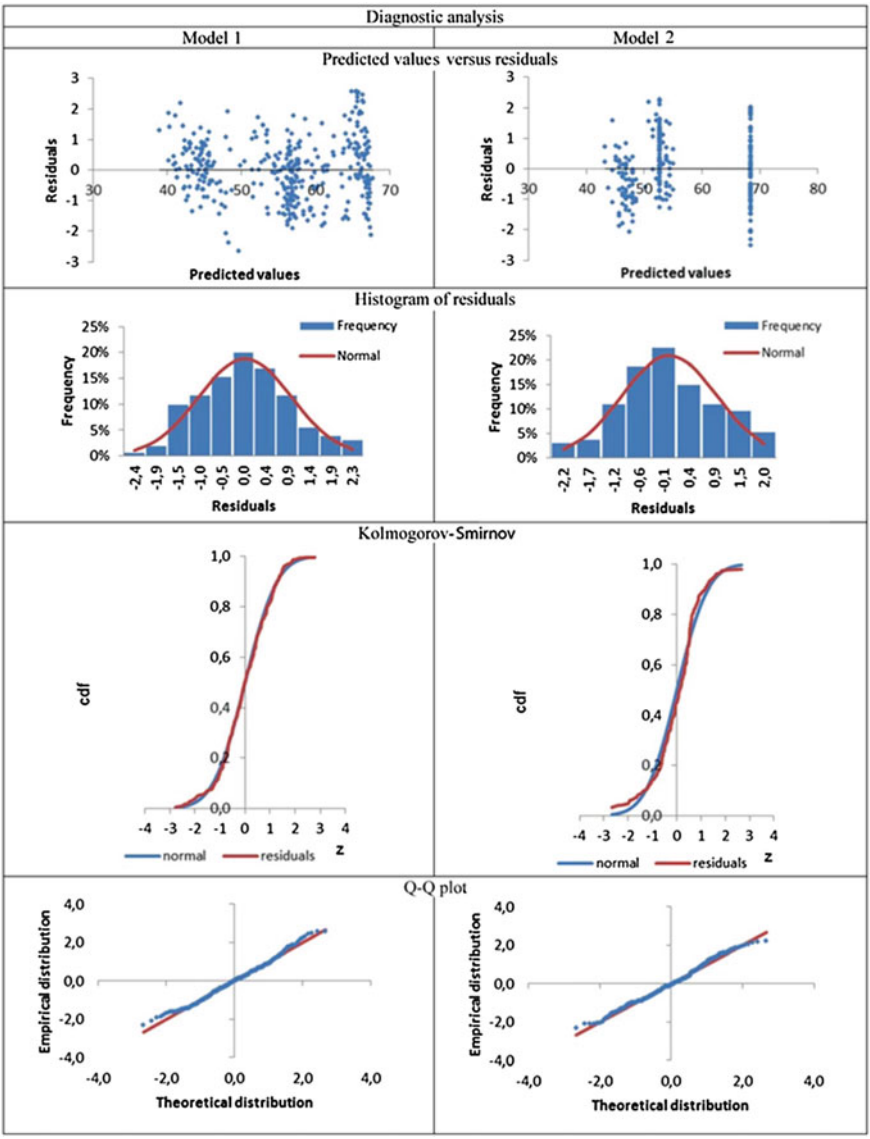


Fig. 4 Diagnostic analysis

$$D_i = \sum_{(j=1)}^{n} (y_j - \hat{y}_{j(i)})^2 / (k+1)\ \text{MSE} \tag{3}$$

y_i is the j-th fitted response value;
$\hat{y}_i$ is the j-th fitted response value, where the fit does not include observation i.;

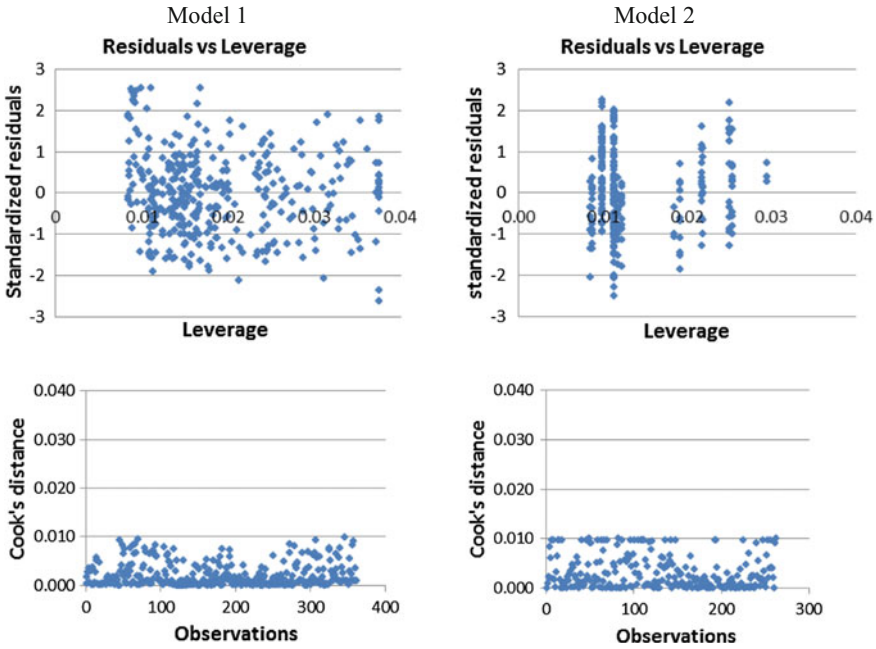


Fig. 5 Leverage values (*up*) and Cook's distance (*down*)

k is the number of coefficients in the regression model;

MSE is mean square error.

Values of Cook's distance that are greater than 1 may be problematic. In the model 1, D is 0.0098 while in model 2, D is 0.0097 as shown in Fig. 5.

In view of all this, Table 4 illustrates the summary of models for space mean speeds. The first model is developed with collected data on every urban road. The second one is developed with data collected on roads characterized by two lanes for traffic in each direction.

5 Validation of Speed Models

Validation is a useful process to test the predictive capacity of model using a new dataset. For this reason, all roads except one, "Viale Colombo", are considered in developing models. This road is chosen for validation because it has the greatest number of independent variables identified on all roads.

In order to check the accuracy of speed equations developed in this study, the models are applied to new data. The observed values are plotted versus predicted values as shown in Fig. 6: most of points fall close to the 45° line demonstrating a good agreement between two sets of speed values.

Table 4 Multiple linear regression model results

Regression equations	R^2_{adj}
V_{ms} = 14.21 – 0.05 F – 0.64 $F_{in,out}$ – 11.84 C + 0.03 TL + 3.84 O + 4.07 LW	0.81
V_{ms} = 21.80 − 0.60 $F_{in,out}$ + 4.11 M + 0.05 TL	0.84
Where	
V_{ms} = space mean speed (aggregated in 5 min intervals) [km/h];	
F = flow (aggregated in 5 min intervals) [v/min];	
$F_{in,\ out}$ = number of vehicles entering and leaving traffic stream (aggregated in 5 min intervals) [v/min];	
C = presence of crosswalk: 0 no crosswalk, 1 crosswalk;	
TL = length of tangent between two intersections in which is located the study section [m];	
O = type of left-lateral obstacle. The variable is: 3 if there is not obstacle; 2 if there is median; 1 if the lanes in opposite direction are not separate; 0 = presence of parking.	
LW = lane widths is the sum of two lane widths in each direction (travel lane and fast lane) [m];	
M = type of median. The variable is: 0 if there is not median; 1 if the median, decorated with trees or hedges, allows the U-turn; 2 if the median, decorated with trees or hedges, is continuous; 3 if the median is indicated by road markings.	

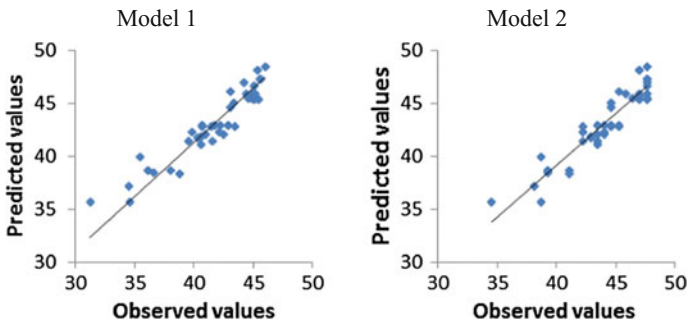


Fig. 6 Scatter plot predicted speeds and observed speeds

The Kolmogorov–Smirnov two-sample test is used to compare predicted speeds to observed speeds. The test confirmed that there is no significant difference between predicted speeds and observed speeds as shown in Table 5.

Table 5 Kolmogorov–Smirnov two-sample test

Model 1		Model 2	
Alpha	0.05	Alpha	0.05
D-stat	0.27	D-stat	0.26
p-value	0.09	p-value	0.08
D-crit	0.29	D-crit	0.29
Observed	41	Observed	41
Predicted	41	Predicted	41
D-stat < D-crit	yes	D-stat < D-crit	yes
p-value > 0.05	yes	*p*-value > 0.05	yes

6 Conclusions

Several studies investigate relationships between driving behavior and roadway characteristics but only a few of these study the tangents of urban roads and consider the space mean speeds.

The proposed study allows a better understanding of driving behavior in urban areas. In particular, this paper addresses the modeling of space mean speeds on tangents. Specifically, multiple linear regression is used and two models are developed to predict speeds. These models include design characteristics that have a significant impact on space mean speeds such as tangent length, lane widths, type of left-lateral obstacle, median type, and presence of crosswalk. Flow and number of vehicles entering and leaving traffic stream are the variables of traffic characteristics.

This is due to the urban area which is characterized by the presence of a greater number of intersections, parking areas, and driveways. Moreover, diagnostic analysis and residual tests are performed to assess adequacy of models and the assumptions are satisfied in both models. Finally, the predictive capacity of models is tested using a new dataset. The test confirms that there is no significant difference between predicted speeds and observed speeds.

Developed models confirm that some variables have more influence on driving speed than other. The research shows that some variables, commonly used in mathematical models, are not statistically significant. For example, we investigate traffic composition: vehicles are divided in different categories (motorcycles, cars, trucks, and buses). Probably these variables are not significant because traffic is homogeneous (approximately 95% of vehicles are cars). Also, we consider access density, number of crossing pedestrians and cyclists, presence of sidewalk, curb, and parking. Moreover, posted speed limit urban roads are distinguished by the same speed limit of 50 km/h, which is the only one in Italy for urban areas that does not effect on speed. In fact, analyzing collected data, the speed limits are substantially exceeded by 68.6% of drivers.

The research contributes to literature on examining vehicle speeds on Italian urban roads. This study can be useful to both traffic manager and road designers because models developed could implement design guidelines, especially regarding

road tangents design. However, it would be appropriate to validate models in other urban roads to determine if drivers have similar behaviors. Furthermore, the results can help roadway designers to better understand expected speeds and, as a result, design the urban roads considering drivers' real driving behavior.

We are currently working on driver behavior study in free flow condition. Particularly, we are establishing time intervals that define isolated vehicle in Italian urban roads. This is important to provide methods to estimate free flow speed as a function of infrastructure geometric characteristics.

References

1. European Road Safety Observatory (2006), 2007. Speeding, Available from Internet: www.erso.eu.
2. Tarris, J.P., Poe C.M, Mason, J.M., Goulias, K.G.: Predicting operating speeds on low-speed urban streets: regression and panel analysis approaches. Transportation Research Record: Journal of Transportation Research Board. 1523, 46–54 (1996).
3. Poe C.M., Mason J.M. Analyzing influence of geometric design on operating speeds along low-speed urban streets: mixed-model approach. In Transportation Research Record: Journal of Transportation Research Board, No 1737, 2000, pp. 18–25.
4. Fitzpatrick, K., and Carlson, P. Selection of Design Speed Values. In Transportation Research Record 81th, Annual Meeting, TRB, Washington, D.C., 2001, pp. 3–11.
5. Wang J., Dixon K.K., Li H., Hunte M.P. Operating-speed model for low-speed urban tangent streets based on in-vehicle global positioning system data. In Transportation Research Record: Journal of Transportation Research Board, No 1961, 2006, pp. 24–33.
6. Aronsson Karin F.M.: Speed characteristics of urban streets based on driver behaviour studies and simulation. Doctoral Thesis in Infrastructure Royal Institute of Technology Stockholm, Sweden, 2006.
7. Ali, A.T., Flannery, A., Venigalla, M.M.: Prediction models for free flow speed on urban streets. In Transportation Research Record 86th Annual Meeting, 2007.
8. American Association of State Highway and Transportation Officials. A Policy on Geometric Design of Highways and Streets. Washington, D.C. (2001).
9. Fitzpatrick, K., Carlson, P., M. Brewer, and Woolridge M. Design Factors That Affect Driver Speed on Suburban Streets. In Transportation Research Record 80th Annual Meeting, Washington, D.C., 2001, pp. 18–25.
10. Fitzpatrick, K., Shamburger, C.B., Krammes, R.A., and Fambro, D.B. Operating Speed on Suburban Arterial Curves. In Transportation Research Record No 1579, TRB, National Research Council, Washington, D.C., 1997, 89–96.
11. Polus, A., Livneh, M., and Craus, J. Effect of Traffic and Geometric Measures on Highway Average Running Speed. In Transportation Research Record No 960, 1984.
12. Ministero delle Infrastrutture e dei Trasporti. Norme Funzionali e Geometriche per la Costruzione delle Strade. Decreto Ministeriale n.6792, November 5, Rome, Italy, 2001.
13. Washington, S.P., Karlaftis, M.G., Mannering, F.L. Statistical and Econometric Methods for Transportation Data Analysis, 2nd ed. Boca Raton, London & New York: CRC Press, ISBN 9781420082869, 2011, pp. 63–143.



http://www.springer.com/978-981-10-6318-3

Proceedings of International Conference on
Computational Intelligence and Data Engineering
ICCIDE 2017
Chaki, N.; Cortesi, A.; Devarakonda, N. (Eds.)
2018, XVI, 372 p. 188 illus., Softcover
ISBN: 978-981-10-6318-3