Transport and Telecommunication, 2021, volume 22, no. 3, 266–277
Transport and Telecommunication Institute, Lomonosova 1, Riga, LV-1019, Latvia
DOI 10.2478/ttj-2021-0020

ANALYSIS AND PREDICTION OF VEHICLES SPEED IN FREE-FLOW TRAFFIC

Andrzej Maczyński¹, Krzysztof Brzozowski², Artur Rygula³

^{1,2,3}University of Bielsko-Biala
Bielsko-Biala, Poland, 43-309 Willowa 2
¹amaczynski@ath.eu

Speed is a crucial factor in the frequency and severity of road accidents. Light and heavy vehicles speed in free-flow traffic at six locations on Poland's national road network was analyzed. The results were used to formulate two models predicting the mean speed in free-flow traffic for both light and heavy vehicles. The first one is a multiple linear regression model, the second is based on an artificial neural network with a radial type of neuron function. A set of the following input parameters is used: average hourly traffic, the percentage of vehicles in free-flow traffic, geometric parameters of the road section (lane and hard shoulder width), type of day and time (hour). The ANN model was found to be more effective for predicting the mean free-flow speed of vehicles. Assuming a 5% acceptable error of indication, the ANN model predicted the mean free-flow speed correctly in 84% of cases for light and 89% for heavy vehicles.

Keywords: traffic, free-flow, speed, model, neural network

1. Introduction

Speeding is defined as traveling faster than the official set limit or at a speed that is too high for the prevailing road conditions. It is a common phenomenon; more than half of drivers in Poland exceed permitted speed limits (NRSC, 2013). Driving with excessive speed narrows the field of vision, reduces the allowable margin of error, extends braking distance, and shortens the time available for drivers to process information and make decisions, thus increasing the likelihood and impact of crashes (Elvik, 2012). Speeding is the cause of a large proportion of crashes (ITF, 2018). For example, in Poland speeding is the cause of around 31% of fatal crashes, and in the US the figure is 27% (NHTSA, 2018; NRSC, 2017). The relationship between traffic speed and collision rate continues to be a subject for study (Gargoum and El-Basyouny, 2016). In order to reduce the number of crashes and their effects, and bearing in mind the impact of a change in traffic speed on the incidence and severity of crashes, the OECD recommends (OECD, 2018): a reduction in the speed on roads and in speed differences between vehicles; setting speed limits based on safe system rules, i.e., at a level that should allow crash victims to survive with no serious consequences; applying compensatory measures where speed limits are increased (e.g., stricter enforcement of limits); the use of automatic speed control systems. Speed control systems, both spot and automated section, can significantly reduce both traffic speed and the proportion of speeding drivers (De Pauw *et al.*, 2014; Montella *et al.*, 2015; Wilmots *et al.*, 2016; Vadeby and Forsman, 2017).

Due to the importance of speed for road safety, the issue of modeling the speed of vehicles on the roads is widely present in the literature (McFadden *et al.*, 2001; Dinh *et al.*, 2013; Himes *et al.*, 2013; Eluru *et al.*, 2013; Bassani *et al.*, 2014, 2016; Semeida, 2014; Gayah *et al.*, 2018). Speed prediction models are typically used to determine the relationship between road geometric design features and vehicle operating speeds. The visual parameters of the road environment in the context of developed speeds have been highlighted by Yu *et al.* (2019). Brief summary of the literature on vehicle speed modeling can be found in Bhowmik *et al.* (2019). High unsafe speeds occur mainly under free-flow conditions when low-density streams are mostly composed of isolated vehicles (Bassani *et al.*, 2014). Therefore, free-flow speed plays a very important role in the field of traffic research and practice. The mean free-flow speed and its variability across drivers are considered important safety factors (Figueroa and Tarko, 2005).

The free-flow speed depends on many factors and it will be influenced by vehicles characteristic, lane width, weather, posted speed limits, drivers' behavior, road conditions and the environment (Saifizul *et al.*, 2011; Gao *et al.*, 2019; Leong *et al.*, 2020; Silvano *et al.*, 2020). The development of a free-flow speed model based on local traffic conditions, which can accurately estimate free-flow speed without

having to conduct field measurements, is essential for saving time and costs in data collection (Leong *et al.*, 2020). There are a lot of papers talking about the issue. Liu *et al.* (2016) proposed a method of modeling free-flow speed from the viewpoint of hydroplaning. The model enables estimation of free-flow speed changes according to the water depth. Silvano *et al.* (2020) estimated the impacts of road characteristics and posted speed limit changes on the free-flow speed distribution using an extensive dataset of speed observations from urban roads with varying characteristics. The results show that the mean free-flow speed on urban roads is influenced by several road characteristics such as land use, on-street parking and the presence of sidewalks. Yasanthi and Mehran (2020) analyzed the impacts of road-weather events on the free-flow speed of vehicles. Linear and nonlinear regression models were developed with separate models for light and heavy vehicles traveling on the shoulder and median lanes on a highway. The study results show how free-flow speeds of light and heavy vehicles were reduced in adverse road-weather conditions such as wet and snow/icy pavement and during snowfalls. Leong *et al.* (2020) developed and assessed free-flow speed models based on different vehicle classes and road characteristics. Analyses of free-flow speed were conducted based on individual and grouped vehicle classes.

Furthermore, free-flow speed information is the input for macroscopic traffic flow models (see e.g. Frejo *et al.*, 2019; van Erp *et al.*, 2020; Hu *et al.*, 2021). At the same time, microscopic traffic flow simulation models need information about free-flow speed to simulate real traffic flow. These facts will naturally emphasize the importance of research about free-flow speed (He, 2016). It is worth mentioning that information concerns free-flow speed is also used in cases related to the assessment of fuel consumption under real traffic conditions. Chen *et al.* (2017) proposed a mesoscopic fuel consumption estimation model using free-flow speed data as the model parameter and showed that this solution can achieve higher accuracy when compared with similar macroscopic and mesoscopic fuel estimation models. Samaras *et al.* (2019) also deal with the calculation of fuel consumption for different traffic conditions, including free-flow traffic. When there was no measurement data about free-flow speed, a simple model for estimating free-flow speed was formulated and applied in order to calculate the fuel consumption.

Bearing in mind that Poland has significantly lower levels of road safety than the majority of EU countries, the authors were prompted to analyze the mean free-flow speed of vehicles on selected locations at national roads in Poland. In contrast to the existing data for national roads, it was decided to include temporal variability in the analyses and to attempt to estimate hourly averaged free-flow speed for two basic categories of vehicles. The carried out literature analysis did not indicate any previous attempt to formulate a model for the estimation of free-flow speed which would include the division of vehicles into discrete categories and, at the same time, would not be a model with a high degree of data aggregation in the terms of time. In the paper, we analyze prediction accuracy for a simple multiple linear regression model using parameters characterizing traffic conditions, geometric parameters of the road, the type of day and time. Furthermore, using the same set of input parameters, we propose a non-linear model for predicting the mean hourly average free-flow speed based on an artificial neural network.

2. Methods

2.1. Study locations

The analyses presented in this work were carried out for data obtained from selected six High-Speed Weigh-In-Motion (HS-WIM) stations located on national roads in Poland. These stations operate on rural single carriageways with two lanes and belong to the national network of WIM stations. The network is managed by General Directorate for National Roads and Motorways, and includes over one hundred WIM stations on national roads. Table 1 contains basic data about the location of particular stations.

Table 1. Study location

Station	Place	Road	Lane width [m]	Hard shoulder width [m]	Geographical coordinates
1	Sucha	DK94	3.5	2.0	50.546149N 18.234294E
2	Wólka	DK10	3.5	2.0	52.888058N 19.403187E
3	Latkowo	DK15	3.0	0.8	52.830481N 18.312170E
4	Cierpice	DK15	3.5	1.25	52.964796N 18.537019E
5	Miechów	DK7	3.5	2	50.371714N 20.045422E
6	Kurów	DK75	3.0	1.5	49.675509N 20.663465E

2.2. Measurement period and data collection

The considerations presented in the paper are based on data collected during one full month. Analysis of the traffic on the sections of road featuring the selected WIM stations during 2018 indicates that monthly traffic volume is highest in August. Taking this into account, the period chosen for the study was August 2018. The average daily traffic and percentage of the heavy vehicles are shown in Table 2.

Table 2. Average daily traffic (ADT) in [veh./day] and percentage of heavy vehicles (HVs) at each station

Station	1	2	3	4	5	6
ADT	4797	3811	8414	8330	9530	8199
HVs	10%	16%	12%	14%	13%	12%

Traffic data collected from the selected WIM stations included: traffic volume, vehicle classification, vehicle speed, and time headway. For further analysis, vehicles were divided into light vehicles (LVs, permissible gross weight below 3.5 Mg) and heavy vehicles (HVs, permissible gross weight above 3.5 Mg). It was assumed that free-flow conditions included all vehicles, where the time headway to other vehicles was not less than 5 seconds.

The results presented below are discussed in two main parts. First, direct analysis of measurement data shows, in different contexts, the variability of mean speed of free-flow traffic. Second, two models (linear and non-linear) for predicting the mean speed of light and heavy vehicles in free-flow traffic are formulated and verified.

3. Results and discussion

The data collected during the study period was analyzed both in aggregated terms for all stations and in detail, with reference to whether or not it was a workday and the time of day, taking into account hourly schedules recorded at individual stations. The results obtained made it possible to determine statistics on the mean speed of vehicles in free-flow traffic for both heavy and light vehicles.

3.1. Mean light vehicles speed in free-flow conditions

At each study location, violations of the speed limit were recorded. Excess speeds varied, while the proportion of drivers exceeding the speed limit ranged from 17.5% to 60.3%. The highest values were recorded at locations 1 and 2 at night. Mean speeds varied, depending on location, from 80.1 km/h (location 6, day) to 96.3 km/h (location 1, night). At every location, the 85th percentile speed (v_{85}) of light vehicles in free-flow conditions was more than the speed limit. In all locations, the v_{85} speed was higher than the limit by an average of about 10%, except location 6, day, where the difference was 2 km/h: slightly over 2.2%. Table 3 shows descriptive statistics for light vehicles speed in free-flow conditions by study location and time of the day. The following notations were used in the tables: v_s – mean speed, σ_v – standard deviation, v_{85} – 85th percentile speed.

Table 3. Descriptive statistics of light vehicles speed in free-flow conditions

Station	Time of day	v_s [km/h]	σ_v [km/h]	v_{85} [km/h]	Percentage of LVs exceeding speed limit [%]				
					Total	Up to 10 km/h	11 to 20 km/h	21 to 30 km/h	Over 30 km/h
1	Day	94.7	14.9	110	57.6	27.2	16.5	8.3	5.6
	Night	96.3	16.2	113	60.3	25.4	16.8	10.4	7.8
2	Day	91.3	15.2	104	49.6	27.6	12.7	3.9	5.4
	Night	92.4	16.2	109	52.5	27.0	14.1	4.8	6.6
3	Day	84.8	12.7	99	32.2	22.5	6.4	2.4	0.9
	Night	87.3	15.0	103	39.7	23.7	8.9	4.7	2.4
4	Day	87.0	13.0	99	38.6	25.5	8.4	3.4	1.3
	Night	86.0	14.1	99	37.3	23.5	8.4	3.9	1.5
5	Day	85.4	12.7	98	31.2	20.4	7.7	2.2	1.0
	Night	84.6	13.6	98	29.5	17.7	7.6	2.8	1.4
6	Day	80.1	11.9	92	17.5	11.3	4.1	1.4	0.6
	Night	85.4	14.1	99	31.5	17.8	8.3	3.7	1.8

The analysis of the results leads to the observation that there is a strong correlation between overall traffic volumes and average free-flow speeds and the percentage of light vehicles exceeding the speed limit. For stations 1 and 2, which have the lowest daily traffic volumes, the highest average free-flow speeds and a significantly higher percentage of vehicles exceeding the speed limit were observed. The correlation between traffic volume and speed in free-flow traffic is also reflected in the fact that night traffic conditions for light vehicles are generally characterized by higher speeds. On the other hand, comparing the results obtained for stations 3-6, it can be concluded that the speed in free-flow traffic is also influenced by road parameters. E.g. station 5 is characterized by the highest traffic volume, but compared to station 6 it has much more favorable geometrical parameters. This results in significantly higher average speeds of light vehicles and a percentage of speed limit exceedances.

3.2. Mean heavy vehicles speed in free-flow conditions

At each location, heavy vehicles frequently exceeded the speed limit. The proportion of drivers exceeding the limit varied from 68.8% to 95.5%, and the highest values were recorded at night at locations 1 and 2. Moreover, at every location, the mean speed of heavy vehicles in free-flow conditions exceeded the speed limit. The mean speeds ranged from 74 km/h (location 6, day) to 83 km/h (location 1, night). At locations 4 and 6, mean speeds of heavy vehicles at night were higher than in the day. The biggest difference was recorded at location 6. At every location, the 85th percentile speed of heavy vehicles in free-flow conditions was significantly higher than the speed limit. The v_{85} speed exceeded the speed limit by between 18.6% (location 6, day) and 27.1% (location 1, night). On average, v_{85} exceeded the speed limit by just over 20%. Table 4 shows descriptive statistics for the speeds of heavy vehicles in free-flow conditions by location and time of day.

Table 4. Descriptive statistics of heavy vehicles speed in free-flow conditions

Station	Time of day	v_s [km/h]	σ_v [km/h]	v_{85} [km/h]	Percentage of HVs exceeding speed limit [%]				
					Total	Up to 10 km/h	11 to 20 km/h	21 to 30 km/h	Over 30 km/h
1	Day	80.9	7.8	88	91.1	35.4	50.3	5.0	0.4
	Night	83	6.8	89	95.5	27.7	60.5	6.7	0.7
2	Day	79.6	7.7	85	91.4	44.4	42.3	3.4	1.3
	Night	80.4	6.6	85	94.5	40.8	50.1	2.8	0.9
3	Day	77.3	8.4	84	85.3	45.1	36.6	3.2	0.5
	Night	79.5	7.9	87	89.8	36.9	46.9	5.1	1.0
4	Day	75.4	7.5	84	79.4	51.2	26.1	1.8	0.4
	Night	75	7.7	84	76.7	47.9	26.5	1.9	0.4
5	Day	78	8.9	86	85.6	46.5	35.2	3.2	0.7
	Night	77	9	86	80.4	45.1	32.8	2.1	0.4
6	Day	74	8.9	83	68.8	47.4	19.9	1.3	0.3
	Night	77.7	8.7	87	80.6	41.1	37.1	2.1	0.3

The observation results for heavy vehicles also confirm the correlation between traffic volume and average speed in free-flow traffic and the percentage of vehicles exceeding the speed limit. Both these parameters are by far the highest for stations 1 and 2. In general, at night heavy vehicles traffic is at higher speeds. The speed increase is obviously smaller than for light vehicles. This is due to the fact that heavy vehicles are equipped with speed limiters and that they are very often already traveling at maximum speeds during the day. The highest speeds of heavy vehicles, regardless of the time of day, were recorded at station 1. The traffic at the lowest speeds was recorded during the day at station 6 and during the night at station 4.

3.3. Workdays and non-workdays

The size and time distribution of traffic is influenced by the type of day, i.e., whether or not it was a workday. Figure 1 shows the percentage of light and heavy vehicles in free-flow traffic, registered during the day and night at individual stations on workdays and non-workdays.

Figure 1. Percentage of vehicles in free-flow traffic at individual stations a) light vehicles b) heavy vehicles

The results indicate a significantly higher percentage of light vehicles in free-flow traffic at night than during the day. This involves a significant reduction in night-time traffic. It is worth noting that the type of day does not matter much in this case. At station 1 on both workdays and non-workdays 100% of the light vehicles participated in free-flow traffic at night. On workdays the lowest shares were recorded at station 5, 18% during the day and 46% at night, respectively. The largest share during the day (51%) was recorded at station 1 on non-workdays. The type of day has a significant impact on the share of heavy vehicles in free-flow traffic. For each station and both times of the day, these shares are higher on non-workdays than on workdays. The smallest 13% difference in shares was recorded at station 5 during the night-time, the largest 32% at station 1 during the daytime.

3.4. Analysis of measurement data on an hourly basis

The volume and structure of the traffic determine the conditions of free-flow. The results of the analysis of traffic variability on an hourly basis were presented for two stations, 1 and 6, significantly different in terms of road section geometry and average daily traffic. Station 1 has a 3.5 m lane width and a 2.0 m hard shoulder width. For station 6, where higher average daily traffic is observed, these parameters are 3.0 m and 1.5 m, respectively. Moreover, significant differences in the speed of vehicles in free-flow traffic were recorded by the analyzed stations. Station 1 has the highest recorded speeds and station 6 has generally the lowest. Figures 2 and 3 show the hourly distribution of free-flow speed for light and heavy vehicles, respectively for these two locations.

Figure 2. Hourly speed of light vehicles in free-flow traffic recorded at a) station 1 b) station 6 (the top and bottom of each box represent the 75th and 25th percentile values, whiskers show the min. and max., the horizontal line represents the median)

Figure 3. Hourly speed of heavy vehicles in free-flow traffic recorded at a) station 1 b) station 6 (the top and bottom of each box represent the 75th and 25th percentile values, whiskers show the min. and max., the horizontal line represents the median)

The data shown in Figures 2 and 3 indicates that for each hour of the day, the median speed in free-flow traffic is higher for station 1, both for light and heavy vehicles. However, a greater variation in speed by the hour is recorded at station 6. The maximum speed ranges at this station are around 25 km/h (hour 1) for light vehicles and 13.8 km/h (hour 0) for heavy vehicles. In both categories of vehicles at both locations, reduced free-flow speeds in the morning (8:00÷10:00 a.m.) are noticeable.

Figure 4 shows the variability of the percentage of vehicles in free-flow traffic on an hourly basis. Analyzing the results, it is easy to see a clear decrease in free-flow traffic in the morning and afternoon. In extremely unfavorable cases the share of free-flow traffic is slightly below 20% (station 6, 10:00 a.m., 3:00÷6:00 p.m.). On the other hand, station 1 recorded only free-flow situations in late-night hours (hours 0:00÷5:00 a.m.). At station 6, the variability in the share of vehicles in free-flow traffic in individual hours is much greater. In general, however, the recorded share is lower than at station 1 during the same hours. Therefore, it can be concluded that at the section of the road where station 6 is located more frequent vehicle platooning occurred.

Figure 4. Percentage of vehicles in free-flow traffic on an hourly basis at a) station 1 b) station 6 (the top and bottom of each box represent the 75th and 25th percentile values, whiskers show the min. and max., the horizontal line represents the median)

In summary, the higher the proportion of free-flow traffic, the higher traffic speeds are observed. This applies to both light and heavy vehicles. During peak hours when traffic volume is the highest, both traffic speed and the share of free-flow traffic decrease. Consequently, the average speed of vehicles in free-flow traffic is also lower.

3.5. Prediction of the mean speed of vehicles in free-flow traffic

In the first step, the multiple linear regression analysis was used to explain the mean speeds in the free-flow traffic (separately for light and heavy vehicles). The input parameters used for the model were: the recorded average hourly traffic (AHT), the percentage of vehicles in free-flow traffic (PoV) and additional parameters for describing local road geometry and time-related, i.e.: the width of the lane and the width of the hard shoulder, the type of day (work or non-work), and time (hour). The results of the linear regression analysis for the whole set of input parameters are presented in Table 5.

Table 5. Results of the multiple linear regression analysis for the mean speed of vehicles of a given category in the free-flow traffic for an averaging period of 1 hour for the whole set of input parameters

Model	Coeff.	Std. Error	t-value	p-value	95% confidence interval	
					Lower bound	Upper bound
LVs, R ² : 0.47, adjusted R ² : 0.469						
Constant	46.4	1.4	33.1	< 0.001	43.7	49.2
Time (Hour)	-1.8·10 ⁻²	0.01	-1.68	0.09	-0.04	0.003
Type of day	-0.86	0.17	-4.96	< 0.001	-1.21	-0.52
Lane width	9.07	0.42	21.5	< 0.001	8.24	9.89
Shoulder width	1.26	0.22	5.76	< 0.001	0.83	1.69
AHT	-1.4·10 ⁻⁴	7·10 ⁻⁴	-0.2	0.9	-1.5·10 ⁻³	1.2·10 ⁻³
PoV	16.6	0.72	23.1	< 0.001	15.2	18.1
HVs, R ² : 0.252, adjusted R ² : 0.251						
Constant	70.8	1	70.8	< 0.001	68.9	72.8
Time (Hour)	-2·10 ⁻³	7.9·10 ⁻³	-0.29	0.8	-0.02	0.01
Type of day	-1.14	0.12	-9.15	< 0.001	-1.38	-0.9
Lane width	0.41	0.3	1.36	0.17	-0.18	1
Shoulder width	2.2	0.16	14.07	< 0.001	1.89	2.5
AHT	7·10 ⁻⁴	5·10 ⁻⁴	1.48	0.14	-2.3·10 ⁻⁴	1.7·10 ⁻³
PoV	8	0.51	15.5	< 0.001	7	9

Critical t-value = 1.96

The results obtained show that in both models some parameters have a statistically insignificant influence on explained values of mean free-flow speed. Thus backward selection procedure was applied for eliminating insignificant parameters and to determine the proper subset of parameters. Finally, in the case of light vehicles, the model has four parameters: the hourly average percentage of vehicles in free-flow traffic (PoV), type of day and two parameters for describing local road geometry. In the case of heavy vehicles, the model has only three parameters: PoV, type of day and hard shoulder width. The results of the linear regression analysis for the determined subset of significant input parameters are presented in Table 6.

Table 6. Results of the multiple linear regression analysis for the mean speed of vehicles of a given category in the free-flow traffic for an averaging period of 1 hour for the determined subset of significant input parameters

Model	Coeff.	Std. Error	t-value	p-value	95% confidence interval	
					Lower bound	Upper bound
LVs, R ² : 0.47, adjusted R ² : 0.469						
Constant	45.9	1.3	36.7	< 0.001	43.5	48.4
Type of day	-0.85	0.17	-4.9	< 0.001	-1.19	-0.51
Lane width	9.06	0.42	21.5	< 0.001	8.23	9.89
Shoulder width	1.26	0.21	5.89	< 0.001	0.84	1.68
PoV	16.9	0.36	47.2	< 0.001	16.3	17.7
HVs, R ² : 0.251, adjusted R ² : 0.251						
Constant	72.6	0.32	227	< 0.001	71.9	73.2
Type of day	-1.14	0.12	-9.17	< 0.001	-1.38	-0.9
Shoulder width	2.3	0.11	20.6	< 0.001	2.08	2.5
PoV	7.4	0.26	28.7	< 0.001	6.9	7.9

Critical t-value = 1.96

The correlation coefficients obtained for both models show that the multiple linear regression can explain the variability of mean speed in the free-flow traffic only in a limited range. The model for light vehicles has a higher R^2 but the average prediction error of the mean speed in free-flow is almost 40% bigger than the model for heavy vehicles. Descriptive statistics for the error in predicting the mean speed using the multiple linear regression are summarized in Table 7.

Table 7. Descriptive statistics for the prediction error of mean speed in free-flow traffic using multiple linear regression model

	Prediction error [km/h]						
	Min	Max	25 Percentile	Median	75 Percentile	Mean	Standard deviation
LVs	$5 \cdot 10^{-4}$	12.4	1.5	3	5.1	3.6	3
HVs	$5.2 \cdot 10^{-5}$	9.5	1	2.1	3.6	2.6	2.2

In order to better explain the variability of mean speed in free-flow traffic a more complex model should be used. Therefore in the next step, a non-linear model predicting the mean speed in free-flow traffic was formulated, as for the regression model, separately for light and heavy vehicles. Because the dependency between the mean speed in free-flow traffic and our set of independent parameters is generally unknown, it was decided to apply the artificial neural network approach. Neural networks have been used in transport issues for several years (Karlaftis and Vlahogianni, 2011). Other authors have already successfully used MultiLayer Perceptron neural networks with sigmoid activation functions to implement models estimating vehicle speed (McFadden *et al.*, 2001; Dinh *et al.*, 2013; Chen *et al.*, 2013; Semeida, 2014). However, the proposed model is based on another type of network with a radial type of activation function (RBF) with a single neuron on the output layer. RBF-type networks are usually used in tasks where the local character of approximation is more preferable (Brzozowska *et al.*, 2005; Sordyl and Brzozowski, 2018). In RBF-type artificial neural networks, the most popular base function is the Gauss function:

$$\eta_i(\mathbf{x}, \mathbf{t}^{(i)}) = \exp \left[-\frac{\|\mathbf{x} - \mathbf{t}^{(i)}\|^2}{2\sigma_i^2} \right], \quad (1)$$

where: $\mathbf{t}^{(i)}$ – center i (i -th neuron on the hidden layer), $\mathbf{x}$ – vector of N input parameters, σ_i – shape coefficient and:

$$\|\mathbf{x} - \mathbf{t}^{(i)}\| = \left[\sum_{k=1}^N (x_k - t_k^i)^2 \right]^{0.5}. \quad (2)$$

For a single neuron on the exit layer, the predicted mean speed v_a^* is calculated as:

$$v_a^*(\mathbf{x}) = \sum_{i=1}^n w_i \cdot \eta_i(\mathbf{x}, \mathbf{t}^{(i)}), \quad (3)$$

where: w_i – weight coefficients, n – number of centers $\mathbf{t}$.

There is a relationship $n \leq m$ where m is the total number of vectors $\mathbf{x}$ used to train the network (which forms a training set).

Unknown weight coefficients w_i are determined in the learning process of the network. The network was trained using the Gram-Schmidt method (Osowski, 1996). The use of equation (2) to predict the mean speed means a locally oriented approximation in the vicinity of centers with a range defined by the value of the shape coefficient σ_i . The process of training the network for a selected form of the activation function (1) requires an earlier solution of the task of proper selection of the number of centers n and shape coefficients σ_i . In this work, the cumulative version of the iterative k-means algorithm was used to select the initial set of n centers (Osowski, 1996). The shape coefficients were formulated as:

$$\sigma_i = \alpha_i \cdot \max_{0 \leq j \leq m} \|\mathbf{x}^{(j)} - \mathbf{t}^{(i)}\|^2, \quad (4)$$

where α_i is an additional scaling coefficient.

The values of unknown scaling coefficients were obtained by solving an optimization task (Brzozowska *et al.*, 2005). To solve the task, the genetic algorithm was used with the real-value representation of genes in the chromosome (Rothlauf, 2006). The goal function was set as a minimum of the expression:

$$\Omega(\boldsymbol{\alpha}) = \sum_{j=0}^m [v_a(\mathbf{x}^{(j)}) - \sum_{i=1}^n w_i \cdot \eta_i(\mathbf{x}^{(j)}, \mathbf{t}^{(i)}, \alpha_i)]^2, \quad (5)$$

where $v_a(\mathbf{x}^{(j)})$ is the real mean speed in conditions given by the j -th vector of the input parameters.

In order to determine the parameters of the model, a learning process was carried out with randomly selected vectors of input parameters. This set included one third of all mean speeds recorded in the analyzed period at given hours of the day, which corresponds to the number of almost 1500 training vectors. The prediction error was evaluated for the remaining two thirds of the measurement data set. Descriptive statistics for the error in predicting the mean speed v_a^* at a given hour of the day using the proposed model are summarized in Table 8.

Table 8. Descriptive statistics for the prediction error of mean free-flow speed using RBF-ANN model

	Prediction error [km/h]						
	Min	Max	25 Percentile	Median	75 Percentile	Mean	Standard deviation
LVs	$6 \cdot 10^{-4}$	9.6	0.9	1.9	3.4	2.5	2.3
HVs	$4 \cdot 10^{-4}$	6.4	0.6	1.4	2.6	1.9	1.9

Based on the analysis of the prediction errors, the model can be used to predict the mean speed of vehicles in free-flow at a given hour of the day with an average accuracy of 2.5 km/h for light vehicles and just below 2 km/h for heavy vehicles. For both categories, this value is below the average error rate associated with speed measurement in this data set, which is 2.6 km/h for light vehicles and 2.3 km/h for heavy vehicles. A boxplot showing statistics of prediction error for data from individual stations is presented in Figure 5.

Figure 5. Prediction error statistics for data from individual stations a) light vehicles b) heavy vehicles (the top and bottom of each box represent the 75th and 25th percentile values, whiskers show the minimum and maximum, the horizontal line represents the median, the cross represents the mean)

The analysis of prediction errors indicates a comparable median error level for data from individual stations for both categories of vehicle. The greatest difference is in the average prediction error values, but the maximum is 1.3 km/h for the light vehicles and 0.6 km/h for the heavy vehicles.

Figure 6. Scatterplot for predicted and measured mean speeds of vehicles in free-flow traffic a) light vehicles b) heavy vehicles

Scatterplots for the predicted and measured mean speeds of vehicles in free-flow traffic for both categories of vehicles and all stations are shown in Figure 6. Generalizing the obtained prediction results, assuming that a 5% error of indication is acceptable, the formulated model correctly predicted the average free-flow speed for a given hour of the day in 84% of cases for light vehicles and 89% of cases for heavy vehicles, respectively.

4. Conclusions

Vehicle speed is a key factor affecting both the probability of road crashes and the severity of their effects. In these circumstances, improving road safety requires a determined approach to slowing down traffic, achievable by effective enforcement of the applicable limits. The study found that, in Poland, speed limits in free-flow traffic conditions were being exceeded both by light and heavy vehicles. The scale of these violations vary, but they always affect traffic safety. Aggregating the results, it was found that the speed limit of 90 km/h for light vehicles was exceeded on average by 38% of drivers in the day and by 42% at night. The 85th percentile speeds were slightly over 100 km/h by day and 103.5 km/h by night. It can be concluded that light vehicles tend to drive faster at night. For heavy vehicles, aggregated test results show that the speed limit of 70 km/h in free-flow is exceeded by an average of 83.5% of drivers in the daytime and by 86% at night. From a traffic safety point of view, these are very worrying results. The 85th percentile speeds were 85 km/h by day and 86 km/h by night, showing that drivers of heavy vehicles also tend to drive slightly faster at night. This can be confirmed by comparing aggregated mean speeds (77.5 km/h by day and 79 km/h at night). The analysis shows that speed limit violation by heavy vehicles is commonplace.

Looking at the data, it is worth noting that the proportion of drivers exceeding the speed limit in both light and heavy vehicles was higher at the first two locations, where average daily traffic was almost half that of the other four locations and where the hard shoulders were wider. At station 6, the proportion of drivers speeding during the day was the lowest, both in light and heavy vehicles. The carriageways at this location are narrower than at the other stations, which in heavy traffic (i.e., essentially daytime) affects driver behavior. The analysis of the results by time of day indicates a significant increase in the percentage of light vehicles in free-flow traffic at night. On average for all locations, the share was 33% on workdays and 37% on non-workdays. At night, however, it increased to 76% on workdays and 75% on non-workdays, respectively. The proportion of heavy vehicles in daytime free-flow traffic conditions was on average 51% on workdays and 74% on non-workdays for all locations. At night, the shares were 55% and 77%, respectively. Thus, the share of heavy vehicles in free-flow traffic is mainly related to the type of day and to a much lesser extent to the time of day. It is worth noting that the type of day has little impact on the speed at which both light and heavy vehicles travel in free-flow traffic. The hourly results were compared for two locations with significantly different road cross-sections.

The results of these analyses were used to formulate models predicting mean speed in free-flow traffic for light and heavy vehicles. The following input parameters are taken into account: geometric parameters of the road section (lane and hard shoulder width), type of day, time (hour), the average hourly traffic and the share of vehicles in the free-flow traffic. The first proposed model is based on multiple

linear regression which explained the variability of the average free-flow speed only in a limited range. The correlation coefficients are as follows: for heavy vehicles $R^2 = 0.25$ and for light vehicles $R^2 = 0.47$. However, the linear model for light vehicles is characterized approximately by a 40% bigger mean and median error in comparison with the prediction error for heavy vehicles. The values of prediction errors reflect the higher variability of light vehicles speed in free-flow traffic. To increase the prediction accuracy, a non-linear model was proposed based on an artificial neural network with a radial type of activation function. This model uses the same set of input parameters. In comparison to the linear regression model, an improvement in accuracy was obtained. For both light and heavy vehicles, the values characterizing the prediction error (max, 25, 50 and 75 percentile, mean, median) were reduced by at least 20%. The maximum prediction error for light vehicles below 10 km/h and mean error equals 2.5 km/h indicate the high effectiveness of the model. Importantly, from a practical point of view, all input parameters are either fixed or easy to determine e.g., using a single inductive loop. The proposed model could be used by road operators to improve safety, for instance by updating speed limits displayed on variable message signs.

Acknowledgements

The authors would like to thank the General Directorate for National Roads and Motorways for providing access to their measurement data.

References

1. Bassani, M., Catani, L., Cirillo, C., Mutani, G. (2016) Night-time and daytime operating speed distribution in urban arterials. *Transportation Research Part F: Traffic Psychology and Behaviour*, 42, 56-69. DOI:10.1016/j.trf.2016.06.020.
2. Bassani, M., Dalmazzo, D., Marinelli, G., Cirillo, C. (2014) The effects of road geometrics and traffic regulations on driver-preferred speeds in northern Italy. An exploratory analysis. *Transportation Research Part F: Traffic Psychology and Behaviour*, 25, 10-26. DOI:10.1016/j.trf.2014.04.019.
3. Bhowmik, T., Yasmin, S., Eluru, N. (2019) A multilevel generalized ordered probit fractional split model for analyzing vehicle speed. *Analytic Methods in Accident Research*, 21, 13-31. DOI:10.1016/j.amar.2018.12.001.
4. Brzozowska, L., Brzozowski, K., Nowakowski, J. (2005) An application of artificial neural network to diesel engine modelling. In: *Proceedings of the Third Workshop-2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, IDAACS, 2005, 142-146.
5. Chen, D., Chen, L., Liu, J. (2013) Road link traffic speed pattern mining in probe vehicle data via soft computing techniques. *Applied Soft Computing*, 13(9), 3894-3902, DOI:10.1016/j.asoc.2013.04.020.
6. Chen, Y., Zhu, L., Gonder, J., Young, S., Walkowicz, K. (2017) Data-driven fuel consumption estimation: A multivariate adaptive regression spline approach. *Transportation Research Part C: Emerging Technologies*, 83, 134-145. DOI:10.1016/j.trc.2017.08.003.
7. De Pauw, E., Daniels, S., Brijs, T., Hermans, E., Wets, G. (2014) Behavioural effects of fixed speed cameras on motorways: Overall improved speed compliance or kangaroo jumps? *Accident Analysis and Prevention*, 73, 132-140. DOI:10.1016/j.aap.2014.08.019.
8. Dinh, D.D., Kojima, A., Kubota, H. (2013) Modeling operating speeds on residential streets with a 30 km/h speed limit: regression versus neural networks approach. *Journal of the Eastern Asia Society for Transportation Studies*, 10, 1650-1669. DOI:10.11175/easts.10.1650.
9. Eluru, N., Chakour, V., Chamberlain, M., Miranda-Moreno, L.F. (2013) Modeling vehicle operating speed on urban roads in Montreal: A panel mixed ordered probit fractional split model. *Accident Analysis and Prevention*, 59, 125-134. DOI:10.1016/j.aap.2013.05.016.
10. Elvik, R. (2012) Speed Limits, Enforcement, and Health Consequences. *Annual review of public health*, 33, 225-238. DOI:10.1146/annurev-publhealth-031811-124634.
11. Figueroa, A.M., Tarko, A.P. (2005) Speed factors on two-lane rural highways in free-flow conditions. *Journal of the Transportation Research Board*, 1912(1), 49-46. DOI:10.1177/0361198105191200105
12. Frejo, J.R.D., Papamichail, I., Papageorgiou, M., De Schutter, B. (2019) Macroscopic modeling of variable speed limits on freeways. *Transportation Research Part C: Emerging Technologies*, 100, 15-33. DOI:10.1016/j.trc.2019.01.001.
13. Gao, C., Xu, J., Li, Q., Yang, J. (2019) The Effect of Posted Speed Limit on the Dispersion of Traffic Flow Speed. *Sustainability*, 11, 3594. DOI:10.3390/su11133594.

14. Gargoum, S. A., El-Basyouny, K. (2016) Exploring the association between speed and safety: A path analysis approach. *Accident Analysis and Prevention*, 93, 32-40. DOI:10.1016/j.aap.2016.04.029.
15. Gayah, V.V., Donnell, E.T., Yu, Z., Li, L. (2018) Safety and operational impacts of setting speed limits below engineering recommendations. *Accident Analysis and Prevention*, 121, 43-52. DOI:10.1016/j.aap.2018.08.029.
16. He, Sheng-Xue (2016) Will a higher free-flow speed lead us to a less congested freeway? *Transportation Research Part A: Policy and Practice*, 85, 17-38. DOI:10.1016/j.tra.2015.12.003.
17. Himes, S.C., Donnell, E.T., Porter, R.J. (2013) Posted speed limit: To include or not to include in operating speed models. *Transportation Research Part A: Policy and Practice*, 52, 23-33. DOI:10.1016/j.tra.2013.04.003.
18. Hu, Z., Smirnova, M.N., Zhang, Y., Smirnov, N.N., Zhu, Z. (2021) Estimation of travel time through a composite ring road by a viscoelastic traffic flow model. *Mathematics and Computers in Simulation*, 181, 501-521. DOI:10.1016/j.matcom.2020.09.025.
19. International Transport Forum (ITF). 2018. *Speed and Crash Risk*. International Traffic Safety Data and Analysis Group, Paris, OECD.
20. Karlaftis, M.G., Vlahogianni, E.I. (2011) Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies*, 19, 387-399. DOI:10.1016/j.trc.2010.10.004.
21. Leong, L.V., Azai, T.A., Goh, W.C., Mahdi, M.B. (2020) The Development and Assessment of Free-Flow Speed Models under Heterogeneous Traffic in Facilitating Sustainable Inter Urban Multilane Highways. *Sustainability*, 12, 3445. DOI:10.3390/su12083445.
22. Liu, M.-W., Oeda, Y., Sumi, T. (2016) Modeling free-flow speed according to different water depths—From the viewpoint of dynamic hydraulic pressure. *Transportation Research Part D: Transport and Environment*, 47, 13-21. DOI:10.1016/j.trd.2016.04.009.
23. McFadden, J., Yang, W.T., Durrans, S. (2001) Application of artificial neural networks to predict speeds on two-lane rural highways. *Transportation Research Record: Journal of the Transportation Research Board*, 1751, 9–17. DOI:10.3141/1751-02.
24. Ministry of Infrastructure and Development. (2015) *Regulation of the Minister of Infrastructure and Development of 20 October 2015 on the classification of road sections with regard to the rate of fatal accidents and due to the road network safety*, Item 1845. Warsaw: Polish Government Publishing Service (in Polish).
25. Montella, A., Punzo, V., Chiaradonna, S., Mauriello, F., Montanino, M. (2015) Point-to-point speed enforcement systems: Speed limits design criteria and analysis of drivers' compliance. *Transportation Research Part C*, 53, 1-18. DOI:10.1016/j.trc.2015.01.025.
26. National Highway Traffic Safety Administration (NHTSA). (2018) *Traffic Safety Facts data 2016: Speeding*. Washington: National Center for Statistics and Analysis
27. National Road Safety Council (NRSC). (2013) *National Road Safety Programme 2013-2020*. Warsaw: Ministry of Infrastructure and Development (in Polish).
28. National Road Safety Council (NRSC). (2017) *The state of road traffic safety and related activities in 2017*. Warsaw: Ministry of Infrastructure and Development (in Polish).
29. OECD. (2018) *Speed and crash risk*. Paris: ITF/OECD.
30. Osowski, S. (1996) *Neural networks in an algorithmic approach*. Warsaw: WNT (in Polish).
31. Rothlauf, F. (2006) *Representations for Genetic and Evolutionary Algorithms*. Springer-Verlag Berlin Heidelberg.
32. Saifizul, A.A., Yamanaka, H., Karim, M.R. (2011) Empirical analysis of gross vehicle weight and free flow speed and consideration on its relation with differential speed limit. *Accident Analysis & Prevention*, 43(3), 1068-1073. DOI:10.1016/j.aap.2010.12.013.
33. Samaras, C., Tsokolis, D., Toffolo, S., Magra, G., Ntziachristos, L., Samaras, Z. (2019) Enhancing average speed emission models to account for congestion impacts in traffic network link-based simulations. *Transportation Research Part D: Transport and Environment*, 75, 197-210. DOI:10.1016/j.trd.2019.08.029.
34. Semeida, A.M. (2014) Application of artificial neural networks for operating speed prediction at horizontal curves: a case study in Egypt. *Journal of Modern Transportation*, 22(1), 20–29. DOI:10.1007/s40534-014-0033-3.
35. Silvano, A.P., Koutsopoulos, H.N., Farah, H. (2020) Free flow speed estimation: A probabilistic, latent approach. Impact of speed limit changes and road characteristics. *Transportation Research Part A: Policy and Practice*, 138, 283-298. DOI:10.1016/j.tra.2020.05.024.
36. Sordyl, J., Brzozowski, K. (2018) Estimation of a 15-minute equivalent noise level in the crossroad area. In: *Proceedings of the 22nd International Scientific Conference Transport Means, Part I*, Trakai, October 2018. Kaunas: Kaunas University of Technology, pp. 342-346.

37. Vadeby, A., Forsman, Å. (2017) Changes in speed distribution: Applying aggregated safety effect models to individual vehicle speeds. *Accident Analysis and Prevention*, 103, 20-28. DOI:10.1016/j.aap.2017.03.012.
38. van Erp, P.B.C., Knoop, V.L., Hoogendoorn, S.P. (2020) On the value of relative flow data. *Transportation Research Part C: Emerging Technologies*, 113, 2020, 74-90. DOI:10.1016/j.trc.2019.05.001.
39. Wilmots, B., Hermans, E., Brijs, T., Wets, G. (2016) Speed control with and without advanced warning sign on the field: An analysis of the effect on driving speed. *Safety Science*, 85, 23-32. DOI:10.1016/j.ssci.2015.12.014.
40. Yasanthi, R.G.N., Mehran, B. (2020) Modeling free-flow speed variations under adverse road-weather conditions: Case of cold region highways. *Case Studies on Transport Policy*, 8(1), 22-30. DOI:10.1016/j.cstp.2020.01.003.
41. Yu, B., Chen, Y., Bao, S. (2019) Quantifying visual road environment to establish a speeding prediction model: An examination using naturalistic driving data. *Accident Analysis and Prevention*, 129, 289-298. DOI:10.1016/j.aap.2019.05.011.