



MODELS FOR DETERMINING ANNUAL AVERAGE DAILY TRAFFIC ON THE NATIONAL ROADS

M. SPŁAWIŃSKA¹

ABSTRACT: One of the basic parameters which describes road traffic is Annual Average Daily Traffic (AADT). Its accurate determination is possible only on the basis of data from the continuous measurement of traffic. However, such data for most road sections is unavailable, so AADT must be determined on the basis of short periods of random measurements. This article presents different methods of estimating AADT on the basis of daily traffic (VOL), and includes the traditional Factor Approach, developed Regression Models and Artificial Neural Network models. As explanatory variables, quantitative variables (VOL and the share of heavy vehicles) as well as qualitative variables (day of the week, month, level of AADT, the cross-section, road class, nature of the area, spatial linking, region of Poland and the nature of traffic patterns) were used. Based on comparisons of the presented methods, the Factor Approach was identified as the most useful.

keywords: roads, traffic flow variability, Annual Average Daily Traffic (AADT), multiple regression, artificial neural networks

1. INTRODUCTION

One of the most important parameters describing traffic is the Annual Average Daily Traffic (AADT), which represents the average number of vehicles passing through a section of road in 24 consecutive hours over one year. Average daily traffic volume is one of the measurements used for different purposes, such as the design and planning of road solutions, noise calculation or study of road accidents. Its incorrect estimation and forecasting can lead to numerous errors, including those described by A. Olma [1]:

- inaccurate designation of road traffic classes, and thus, incorrect surface design,

¹ Cracow University of Technology, Faculty of Civil Engineering, Ul. Warszawska 24, 31-155 Krakow, Poland,
e-mail: msplaw@pk.edu.pl

- inaccurate determination of reliable traffic volume, which can trigger the adoption of an incorrect number of lanes at entry and exit points at intersections, creation of road sections which are too broad or too narrow, inappropriate geometric arrangements intersections and junctions, improper traffic management at intersections, incorrect design of communication systems and their components,
- overloading of routes designed for less traffic volume than which actually can occur,
- inadequate protection against the impact of roads on the environment (noise, air pollution),
- erroneous results of traffic safety analyses, economic analyses and forecasts of road accidents.

Obtaining an accurate and reliable value of AADT is possible only on the basis of data from continuous automated traffic measurements. However, chiefly due to the cost of collecting and processing the data, there are no such data from most road sections, so AADT must be determined on the basis of short periods of random measurements. In Poland, the indicator method is used, based on 19 independent conversion factors (7 for weekdays and 12 for months). The indicators, representative of the entire network of national roads outside cities with specific traffic profiles (economic, tourist, recreational), are published every five years in a document entitled Road Traffic (Ruch Drogowy) [2]. Based on those indicators, a simplified conversion of daily traffic volumes into AADT can be performed. However, such transposition of short-term measurement results is exposed to a high risk of error associated with the variation of fluctuations in traffic volume for different sections of national roads and other random factors. In literature, other approaches to determining the conversion factors can also be found, as described in [3, 4] and by L. Jin, Ch. Xu, and J. Fricker J. [5]. In the US, the coefficients of weekly variability are determined for each day of the week as well as a collective lower number of days, with varying degrees of aggregation of days with similar daily traffic volume values (5, 4, 2 coefficients). They can be perceived separately or combined with months; in which 84, 60, 48 or 24 coefficients are obtained. In some cases, only the coefficients of weekly variability or only the coefficients of seasonal variability can be used. However, the works on the evaluation of this method by L. Jin, Ch. Xu, J. Fricker J. [5], M. Sławińska [6], S. Granato [7], and R. Luttinen, O. Virpi [8] showed that even with a more flexible approach, where choosing the number and type of coefficients depends on the situation, the method is too general and the results are inaccurate. In foreign literature, other estimation techniques can be found for estimating AADT, for instance, on the basis

of weekly average daily traffic (WADT) described by M. Lewis, and D. Albright [9]. The coefficients required to convert WADT into AADT are given in the Highway Capacity Manual (HCM) [10] and must be approved by the state administration. There are newer methods which enable more accurate determination of AADT (improvement of accuracy by up to 5%) (described by A. Olma [1], L. Jin, Ch. Xu, J. Fricker J. [5], M. Lewis M., D. Albright [9], J. Eom, M. Park, T. Heo, L. Huntsinger [11], M. Sabry, H. Abd-El-Latif, S. Yousef, N. Badra [12, 13], B. Selby, K. Kockelman [14], X. Wang, K. Kockelman [15] H. Wu H., Z. Zhang [16] Y. Zhang, Y. Xie [17]), and include: linear multiple regression, analysis of variance (ANOVA), time series (TS), data mining with Artificial Neural Networks (ANN) and Co-clustering Collaborative Filtering (CCCF), Kriging geostatistical method, or hybrid models (FBFN, ARIMA - MLANN, GA + TDNN). Nevertheless, it is hard to find a comparison of the methods and their resulting conclusions.

The aim of the article is to present the most useful (best ratio of cost input to accurate results) AADT estimation methods developed for Polish conditions. The following methods were analyzed: an indicator method, multiple regressions and a model using artificial neural networks. Explanatory variables used were: quantitative variables: daily traffic (VOL), heavy vehicles (hv) [%]; and qualitative variables: day of the week, month, the level of AADT, cross-section, technical class of the road, spatial features, character of the area, region of Poland, and the nature of traffic patterns.

2. DATABASE

The analysis was performed based on data from continuous measurement stations named Golden River in 2010 (station locations are shown in Figure 1) with respect to the daily traffic volumes of the total number of vehicles in a road section. It should be noted that the most commonly used recorders (Marksman 660) are subject to failure and data derived from them are incomplete and subject to errors. From a 16-hour supplementary manual measurements which are carried out twice a year (May and September) it appears that the automatic classification is performed incorrectly [18]. The resulting data can be used only for the distribution of vehicles for light vehicles and heavy vehicles, which is sufficient to determine the share of heavy vehicles. Because the analyzed models relate to all vehicles it does not affect the accuracy of calculating the AADT. Statutory holidays and those directly related to them were removed from the analysis due

to the difference in values of the daily traffic volumes and the daily average (Table 1). The database was also completed with data defining the characteristics of the roads and their immediate environment, such as their location (region of Poland, summer tourism area), the geometric features (elements of the road section, type of hard shoulder), environment management (land development, facilities in the area), the interaction of neighboring cities (provincial cities or those with a population exceeding 200,000 and peri-urban areas), the type of traffic patterns, the height of the terrain, data on nearby industry and commerce, and possible road works in the years between 2000-2010 (repairs, remodeling, construction on adjacent sections). Due to the large number of variables (12 distinguished road characteristics, traffic and environment) in relation to the measurement stations, the relationship between certain characteristics (ie. the dominant type of building and the impact on cities) and the lack of diversity of environment development and landform (domination of fields, meadows, or forests surrounded by roads and flat terrain) it was decided to aggregate the variable elements. Table 2 shows a fragment of the resultant database (1 indicates the analyzed aspect relates to the given case).

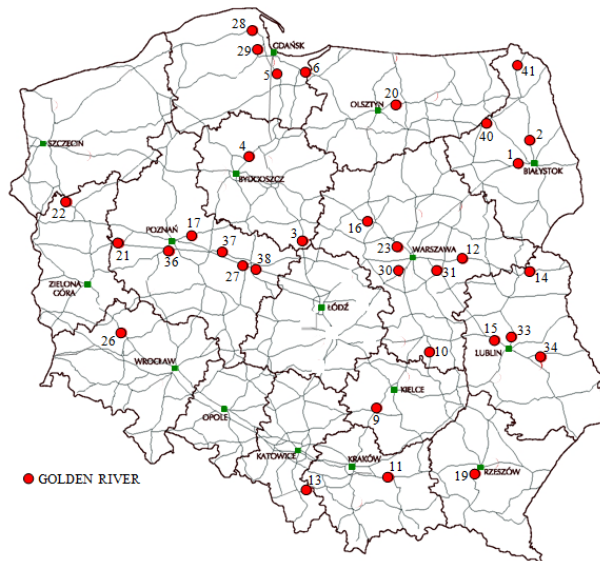


Fig. 1. Map of the locations of the continuous measurement stations (source: based on [18])

Table 1. The list of public holidays and related dates in 2010

Days of public holidays in 2010		the "period" days off work
New Year	01.01.	01.01 - 03.01.
first day of Easter	04.04.	02.04 - 06.04.
second day of Easter	05.04.	
Public Holiday	01.05.	30.04. - 04.05.
Third of May National Day	03.05.	
first day of Whitsun	23.05.	21.05. - 24.05.
Corpus Christi	03.06.	02.06. - 07.06.
National Independence Day	11.11.	10.11. - 15.11.
Christmas Day	25.12.	24.12. - 31.12.
Boxing Day	26.12.	

Table 2. Characteristics of Golden River measuring stations

Number of stations	roads class	cross section		long distance traffic			tourist area		the impact of cities				region of Poland						road works					
		1x2	2x2 (1x4)	yes	no	international traffic	yes	no	< 20 km	> 20 km	suburban	no	Pomerania	Warmia – Masuria – Podlasie	Greater Poland	Mazovia	Silesia	Lesser Poland	yes	no				
33	GP	1		1		1		1			1													
34	GP	1		1		1		1				1					1						1	
11	GP	1		1					1														1	
4	GP	1		1				1				1	1										1	
28	GP	1		1		1	1				1		1										1	
40	GP	1		1		1	1					1		1										
36	A		1	1		1		1	1							1							1	

In accordance with what was earlier established [19, 20], the analysis was carried out during three time periods; throughout the whole year and in the months for which the highest accuracy of AADT estimation was achieved; from May to October and from September to October for each day of the week and for the days with the highest accuracy of AADT estimation; from Tuesday to Thursday (the so-called most favorable period of measurement). In order to determine the model relationship between the characteristics of traffic volume variability over time and AADT, the estimated annual average daily traffic, called AADT_E (average of the values obtained for consecutive days at a preferred period), was determined by four methods in various combinations of identified road features and the environment. In order to evaluate the model used, the estimated values were compared to empirical values and a Mean Absolute Percentage Error (MAPE) was calculated – Eq. (2. 1).

$$(2.1) \quad MAPE = \left| \left(\frac{AADT - AADT_E}{AADT} \right) \times 100\% \right|$$

where: AADT was determined indirectly by AASHTO method – Eq. (2. 2)

$$(2.2) \quad AADT = \frac{1}{7} \sum_{i=1}^7 \left[\frac{1}{12} \sum_{j=1}^{12} \left(\frac{1}{n} \sum_{k=1}^n VOL_{ijk}^d \right) \right] [\text{Veh}/24\text{h}]$$

where: i – day of the week; j – month of the year; k – subsequent day in the month (e.g. the second Monday in January); VOL_{ijk}^d – daily traffic volume i – of the day of the month j , n – the number of i – the days in the month

3. FACTOR APPROACH OF AADT ESTIMATES

The conversion of daily traffic volumes (VOL) into AADT was made in accordance with Eq. (3.1). Adopted as conversion coefficients of seasonal and weekly fluctuations in traffic volume were:

1. coefficients related to the road characteristics and their environment defined by M. Spławińska in the development [20] – FA method,
2. coefficients included in the development of the Road Traffic [2] – FB method,

$$(3.1.) \quad AADT = \frac{VOL}{W_i \times W_j} [\text{Veh}/24\text{h}]$$

where: VOL – daily traffic [P/24h], W_i – day-of-week factor is the ratio of weekday (e.g. Monday) average daily volume (WADT) to the AADT; W_j – monthly factor is the ratio of monthly (e.g. January) average daily traffic (MADT) to the AADT

Analyses show that the FA method gives far more accurate results than the FB method. For the annual period, in all days of the week, for all points, a reduction of the average value of MAPE at 10.4% was achieved. The average MAPE for the FA method is 6.5% and for the FB method, 7.3%. Furthermore, in the case of the FB method, the MAPE values within one station

and between the stations show much greater variation compared to the FA method (mean standard deviation calculated from the IEA increased by about 12%), indicating less stability and the possibility of committing a greater error when selecting adverse periods of measurement. When using the FA method, in the case of days from Tuesday to Thursday, a 6.6% reduction in the average value of MAPE was obtained. For a period covering months V – X and IX – X, a 22.7% reduction of the average value of MAPE and a 23.2% lower value of average standard deviation were achieved. In the case of days from Tuesday to Thursday, a 13.7% reduction in the average value of MAPE for months V – X and a 24.3% reduction for months IX – X were obtained. In addition, analyses show that, as expected, there was a reduction in the average value of MAPE within one method in the case of estimating AADT on the basis of daily traffic volumes from the most advantageous period of traffic measurement.

Similarly, in the case of months V – X, IX – X (all days of the week) and for the year, months V – X and IX – X for the days Tuesday – Thursday, a respective reduction was obtained (in reference to the whole year, all days of the week):

- FA method: 22.3%; 29.6%; 11.6%; 35.1% and 41.8%,
- FB method: 10.2%; 18.1%; 15.2%; 32.6% and 31.2%.

These results confirm that the best periods to carry out traffic measurements to estimate AADT are months V – X (in the FA method, in particular IX – X) and days from Tuesday to Thursday. The AADT estimated on the basis thereof may be flawed in the case of: the FA method – 4.2% (3.8% for the months IX – X), and the FB method – 4.9%.

The AADT estimates based on measurement results from the remaining months (I – IV, XI, XII) from Monday to Sunday may be erroneous, or even doubling to 8.2%.

Figures 2 and 3 show the average values of the MAPE and σ obtained as a result of the AADT estimation according to FA and FB methods in three distinguished periods. High MAPE in station 36 on the A2 highway can be caused by its location in close proximity to the city of Poznań (high impact of urban area).

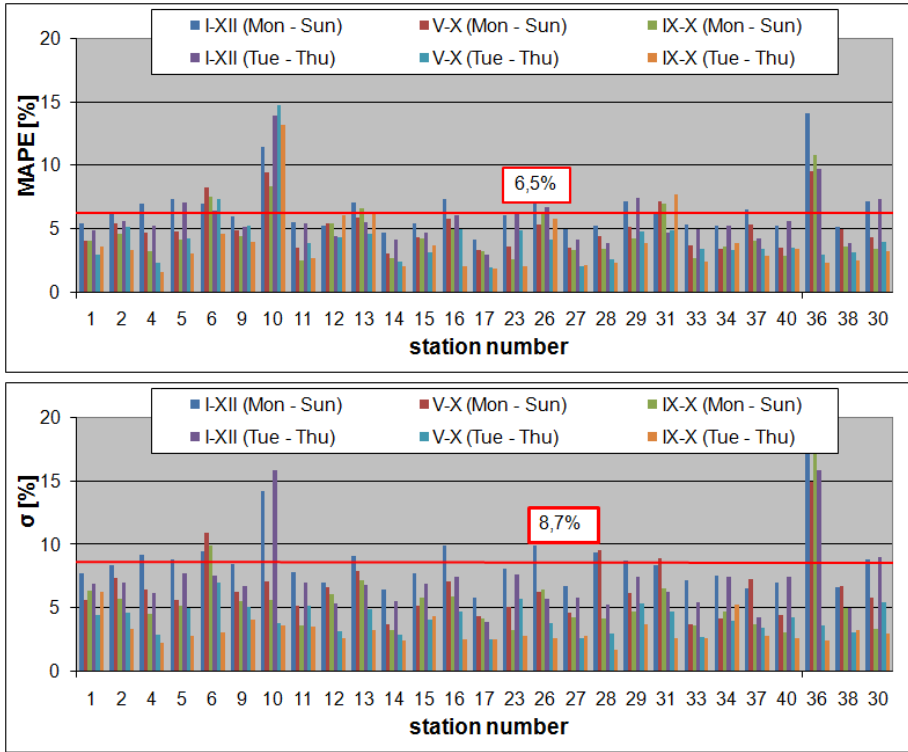


Fig. 2. Comparison of the accuracy of estimates of AADT using factor approach - FA

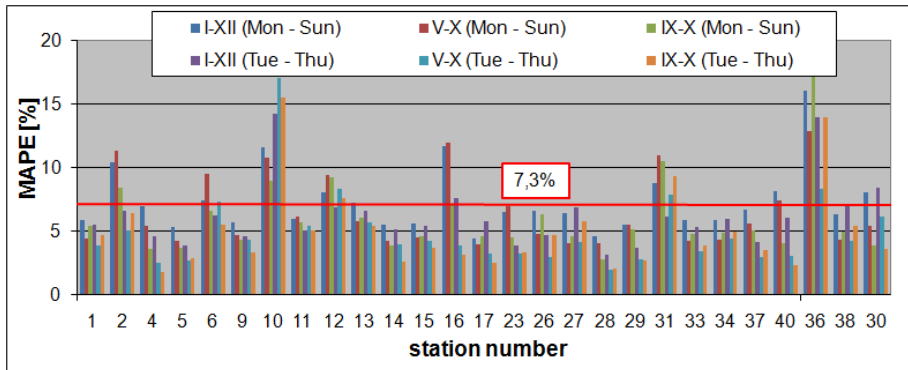


Fig. 3. Comparison of the accuracy of estimates of AADT using factor approach - FB

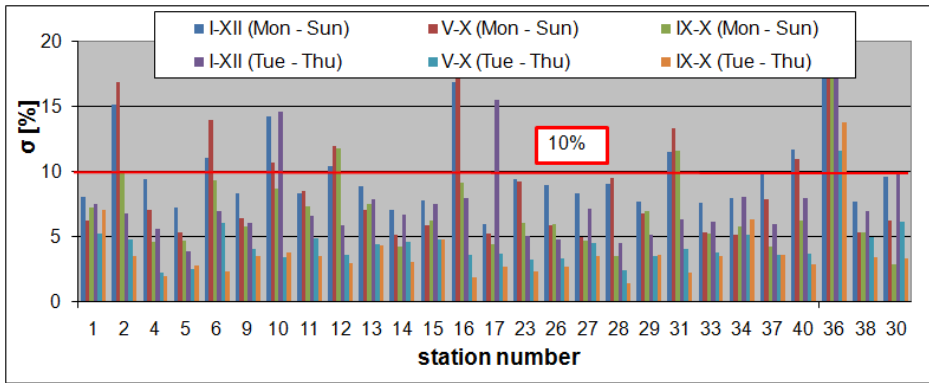


Fig. 3. Comparison of the accuracy of estimates of AADT using factor approach – FB – continued

4. MULTIPLE REGRESSION APPROACH OF AADT ESTIMATES

Explanatory variables used:

- quantitative variables – the value of the daily traffic volume (VOL), the proportion of heavy vehicles (hv – the period of analysis IX – X was not taken into account due to the lack of statistical significance).
- qualitative variables – day of the week, month, Polish region (Pomerania, Warmia – Masuria – Podlasie, Greater Poland, Mazovia, Silesia, Lesser Poland), spatial relationship (direct impact of urban areas at a distance < 20 km, the impact of urban areas at a distance > 20 km, suburban zones; outside the influence of urban areas), the level of AADT (< 10000 P/d or > 10000 P/d – included only if it shows statistical significance, that is, for the whole year, all days of the week and in the months V – X and IX – X for days from Tuesday to Thursday), the cross-section (1x2 or 2x2) and the nature of traffic patterns (long distance or not). Other variables include the technical class of the road (A or GP), tourist area and international traffic were removed from the analysis due to high correlation with other variables (Spearman’s correlation coefficient is: highway/Greater Poland – 0.74; tourist area/Pomerania – 0.69; long-distance/international traffic – 0.51).

In order to use qualitative variables, they were transformed into zero-one variables (1 – the case has a particular feature, 0 – the case does not have this feature).

For example, for the period IX – X, Tuesday – Thursday, the model takes the following form:

$$(4.1) \quad \begin{aligned} AADT_E = & 3240 + 0,85 \cdot VOL - 234,5 \cdot Wed - 656,3 \cdot Thu - 158,1 \cdot IX - 792,9 \cdot P \\ & - 1996 \cdot WMP - 2303 \cdot W - 1125,9 \cdot Maz - 1937 \cdot S + 680,5 \cdot di \quad [\text{Veh}/24\text{h}] \\ & - 348,1 \cdot sz - 245,5 \cdot oi - 397,9 \cdot AADT - 335,7 \cdot c1x2 + 361,7 \cdot ld \end{aligned}$$

where: VOL – the value of daily traffic, Wed – Wednesday, Thu – Thursday, IX – September, P – Pomerania, WMP – Warmia – Masuria – Podlasie, W – Greater Poland, Maz – Mazovia, S – Silesia, di – direct impact of city agglomerations at the distance < 20km, sz – suburban zone, oi – outside impact of city agglomerations, AADT – the level of AADT > 10000 P/d, p1x2 – cross section 1x2, ld – long distance traffic.

The variables representing the daily traffic volumes have strong positive correlation with the explanatory variable. Data on the days and months of the year have a very low coefficient of correlation with the dependent variable. However, due to the conclusions arising from the review of the literature, it has been added to the set of explanatory variables. Table 3 shows the statistics for sample linear multiple regression models. Small estimation errors, high significance of the parameters obtained for $p < 0.05$ (explanatory variables showing statistical significance in Student's t-test were marked in red) and the value of the coefficient $R^2_s = 0.99$ prove a very good fit for the model. In order to validate the model, its verification was made based on the Snedecor's F distribution, analysis of variance – ANOVA and analysis of residues – Fig. 4 (positive verification of the model and a good fit of residues to a normal distribution were achieved). In addition, the required sample size necessary to estimate AADT was checked with a maximum error equal to half the confidence interval (177 Veh/24h for all points and 257 Veh/24h for the selected ones) at a confidence level of 0.95. The required number for all points amounts to 8232 (it is 8272) and for the selected points 3064 (it is 3075). Figure 5 shows the average value of the MAPE and σ at each measurement point. For the annual period, all days of the week, the average value obtained was: MAPE equal to 8.1% and σ equal to 7.9%. In this case, the improvement of the accuracy of AADT estimation based on the daily traffic volumes from the most favorable period of traffic measurements was achieved as well. For months V – X, IX – X (all days of the week), a respective reduction was achieved (reference to the whole year, all days of the week): 22.4% and 26.7%.

For the year, months V – X and IX – X for the days Tuesday – Thursday, a respective reduction was obtained (reference to the whole year, all days of the week): 30.7%; 53.4%; 52.8%.

Table 3. The results of the statistical analysis of the regression model
(month IX – X, day of the week: Tuesday – Thursday)

explanatory variables	R= 0,995 R ² =0,991 corrected R ² =,990 F(15,660)=4597,9 p<0,0000					
	BETA	Standard Error	β	Standard Error	t(660)	p
			3240,27	323,32	10,02	0,000
VOL	0,97	0,01	0,85	0,01	108,39	0,000
Wednesday	-0,01	0,00	-234,51	77,48	-3,03	0,003
Thursday	-0,04	0,00	-656,30	77,82	-8,43	0,000
September	-0,01	0,00	-158,14	62,87	-2,52	0,012
Pomerania	-0,04	0,01	-792,92	124,64	-6,36	0,000
Warmia – Masuria – Podlasie	-0,08	0,01	-1996,04	192,28	-10,38	0,000
Greater Poland	-0,11	0,01	-2303,21	142,41	-16,17	0,000
Mazovia	-0,07	0,01	-1125,91	138,00	-8,16	0,000
Silesia	-0,05	0,01	-1936,98	229,31	-8,45	0,000
impact < 20km	0,04	0,01	680,48	125,39	5,43	0,000
suburban zone	-0,01	0,00	-348,09	118,33	-2,94	0,003
outside impact	-0,01	0,01	-245,51	113,11	-2,17	0,030
AADT > 10000	-0,02	0,01	-397,89	134,79	-2,95	0,003
1x2	-0,02	0,01	-335,71	167,86	-2,00	0,046
long distance traffic	0,02	0,01	361,71	122,54	2,95	0,003

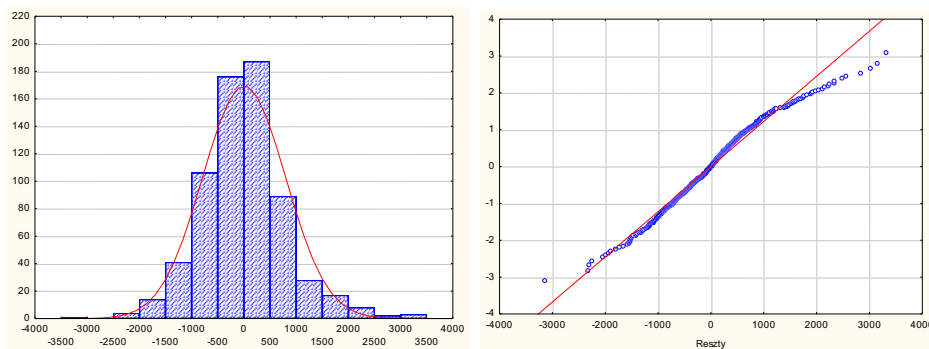


Fig. 4. Analysis of the residuals for the model regression
(month IX – X, day of the week: Tuesday – Thursday)

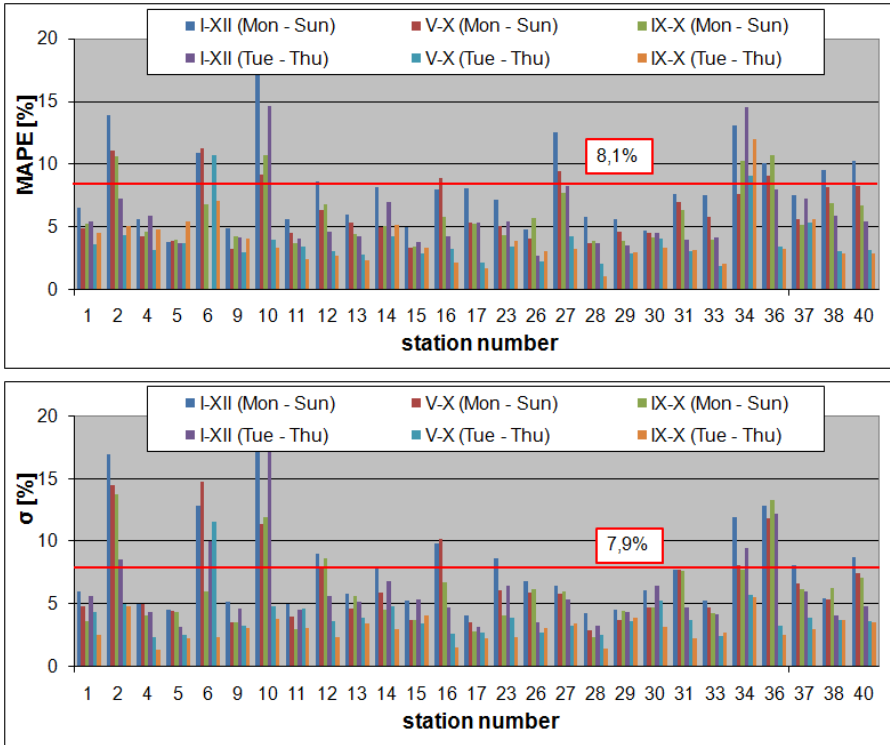


Fig. 5. Comparison of AADT estimation accuracy using multiple regression models

5. ARTIFICIAL NEURAL NETWORK APPROACH OF AADT ESTIMATES (ANN)

For the construction of models based on artificial neural networks, input variables selected for multiple regression models were used (daily traffic volume, the proportion of heavy vehicles, day of week, month, region of Poland, spatial relationship, level of AADT, cross-section and the nature of traffic patterns). The selection of final optimal set was made automatically using a genetic algorithm (Table 4 – blank field for an error ratio means a missing variable). Below, in Table 4, the results of modeling for the top 3 models obtained were shown (one-way multi-layer MLP network with one hidden layer), including: network structure (the number of independent variables

at the input and output of the network and the number of neurons in each layer of the network), the value of deviations quotient, correlation, the value of MAPE and the error ratio obtained from a sensitivity analysis for each variable in the model (impact assessment of independent variables). The high correlation value and low deviations quotient value prove a good fit for all of the developed models. In all cases the highest value of the error ratio was obtained for a variable of spatial relationship or cross-section and, subsequently, VOL and Polish region, which testifies to their greatest impact on the quality of the ANN built. Furthermore, other variables (depending on the measurement period) show significance, which proves their proper selection. Figure 6 shows the average value of MAPE and σ for the best model achieved (highlighted in bold in the table 4). For the annual period, all days of the week, the average value was achieved: MAPE equal to 4.6% (range from 1.2% to 18.9%) and σ equal to 2.4%. As it can be seen, the improvement of the accuracy of AADT estimation based on the daily traffic volumes from the most favorable period of traffic measurements was obtained. For months V – X, IX – X (all days of the week) a respective reduction was achieved (reference to the whole year, all days of the week): 14.8% and 29.4%. For the year, months V – X and IX – X for the days Tuesday – Thursday a respective reduction was obtained (reference to the whole year, all days of the week): 25.1%; 44.2%; 51.9%.

Table 4. Descriptive statistics for the artificial neural network models

The period of analysis	Type and structure of the network	Quotient deviations	Correlation	Mape [%]	Error quotient of the sensitivity analysis									
					VOL	hv	day	month	region of Poland	spatial linking	AADT > 10000	cross-section	nature of traffic	
I-XII	6:14-2-1:1	0,12	0,99	5,20	1,63					2,85	4,50	1,52	4,18	2,15
	8:27-3-1:1	0,12	0,99	5,26	1,83	1,14		1,10	3,46	4,05	1,68	4,94	2,67	
	9:34-4-1:1	0,11	0,99	4,63	2,60	1,33	1,13	1,19	2,66	2,98	1,78	4,07	1,52	
V-X	4:12-2-1:1	0,11	0,99	5,07	3,81					2,54	3,77		4,42	
	4:12-5-1:1	0,10	1,00	4,46	3,83					2,92	4,88		5,42	
	5:13-6-1:1	0,09	1,00	3,94	2,85					3,69	6,48		5,43	4,45
IX-X	4:12-5-1:1	0,10	0,99	4,75	3,52					3,03	5,52		5,75	
	5:13-4-1:1	0,09	1,00	4,13	3,46					3,17	5,21		5,31	1,67
	8:22-4-1:1	0,08	1,00	3,27	4,86			1,24	1,00	3,02	5,58	1,16	5,47	1,33

The period of analysis	Type and structure of the network	Quotient deviations	Correlation	Mape [%]	Error quotient of the sensitivity analysis									
					VOL	lv	day	month	region of Poland	spatial linking	AAADT > 10000	cross-section	nature of traffic	
I-XII - Tue-Thu	5:24-4-1:1	0,09	1,00	3,77	6,43				1,35	3,04	5,96		6,74	
	6:25-5-1:1	0,09	1,00	3,47	3,23				1,27	3,06	5,90	2,49	6,13	
	8:27-4-1:1	0,09	1,00	3,58	3,62	1,14			1,32	2,71	4,66	1,47	6,01	1,60
V-X - Tue-Thu	8:23-3-1:1	0,08	1,00	3,68	7,48	1,19	1,03	1,07	1,72	3,22	1,20	3,96		
	9:24-4-1:1	0,06	1,00	2,91	7,30	1,88	1,04	1,18	3,24	5,12	2,05	6,49	2,11	
	8:23-6-1:1	0,06	1,00	2,58	6,54	2,22	1,02	1,24	5,97	7,03		7,35	4,51	
IX-X - Tue-Thu	9:19-7-1:1	0,06	1,00	2,78	10,34	2,59	1,09	1,04	2,57	2,28	1,25	3,75	1,50	
	7:15-5-1:1	0,05	1,00	2,23	6,49	4,85			5,44	7,15	1,67	8,44	3,93	
	7:15-8-1:1	0,06	1,00	2,50	7,81	2,95			5,54	5,45	1,83	7,47	3,71	

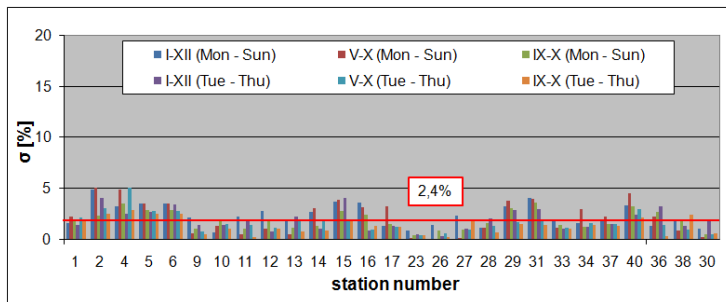
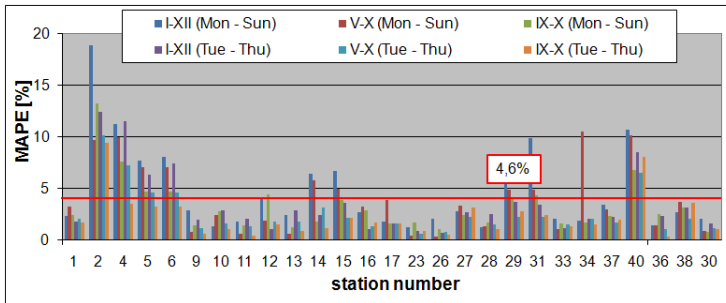


Fig. 6. Comparison of estimation accuracy of AADT using ANN model

CONCLUDING REMARKS

In all models used, variables with most statistical significance were: VOL (daily traffic volume), cross-section of the road (in the case of the indicator method FA – technical class of the road), region of Poland and spatial relationship. These variables also have the greatest impact on the explanatory variable (AADT). In the case of multiple regression models, the following variables are also statistically significant: the nature of traffic patterns, day of the week and month of the year, whereas the significance of the rest of the variables depends on the period of analysis (AADT level – for the whole year, every day of the week and the months V – X and IX – X for days Tuesday to Thursday; hv – for all cases apart from the months IX – X for all days of the week and days from Tuesday to Thursday). Other statistically significant variables are also: technical class of the road, tourist area, international traffic – however, due to high correlation with other variables, they were removed from the models (the variables which lowered the accuracy of the model were removed). In the case of artificial neural networks, for the best obtained model, in 6 cases (total number of vehicles), statistically significant proved to be respective variables: the nature of traffic patterns 5 times, the level of AADT and month of the year 4 times as well as day of the week and has 3 times. In addition, in the ANN models, unlike regression models, a greater impact of qualitative variables on the explanatory variable was noted, even exceeding the impact of the VOL. In the case of the FA indicator method, tourist areas are an important factor.. These results plus the achieved high accuracy of models (Table 5) show the correct assignment of the impact of the highlighted features of roads and the environment on AADT. Although the best accuracy (the entire database) was obtained for the ANN models, in particular for a larger time range data (the whole year, days from Tuesday to Thursday), due to the lack of possibility of mathematical writing of interrelationships between models and the need for a correspondingly large sample, it is difficult to use this method. In the case of multiple regression models, they ensure high accuracy for estimating AADT only for selected data (days Tuesday – Thursday), which greatly limits their use. In practice, identified interrelationships can be used to divide the road network into homogeneous groups of traffic variability and to estimate AADT with sufficient accuracy using the FA indicator method (a 24% reduction of the average value of MAPE compared to the current FB method on days Tuesday – Thursday, months IX – X). This method, due to its simplicity and versatility (it enables the estimation of the value of AADT for the roads of higher technical classes with one and two lanes, with varying nature of traffic patterns, with

distinction on the directions of movement and it does not require a large sample), has the fewest limitations of all the methods used and it is, therefore, the most useful. Furthermore, because of the profile's stability over a long period of time proven by M. Spławińska [21], it can be concluded that certain interrelationships will also occur in other years. Accordingly, recommended methodology of estimating AADT is indicator method. It should be noted, however, that the FA method requires the creation of a reliable database taking into account the characteristics of the road and its surroundings. It was also shown that the used tools are useful in the analysis of traffic data and may be applied in future studies thanks to the successive increase in available traffic data. Can also be used by road administration for databases updates, including verification of the date from traffic measurements.

Table 5. Comparison of estimation accuracy of AADT

METHOD	AVERAGE MAPE [%]					
	MONDAY - SUNDAY			TUESDAY THURSDAY		
	I - XII	V - X	IX - X	I - XII	V - X	IX - X
Indicator - FA	6,5	5,1	4,6	5,8	4,2	3,8
Indicator - FB	7,3	6,5	6,0	6,2	4,9	5,0
Multiple Regression	8,1	6,3	5,9	5,6	3,8	3,8
Artificial neural networks	4,6	3,9	3,3	3,5	2,6	2,2

REFERENCES

1. A. Olma, Określenie współczynników przeliczeniowych do szacowania natężeń ruchu drogowego w obszarach miejskich, Praca doktorska, Politechnika Śląska, Gliwice 2005
2. Ruch Drogowy 2010, Transprojekt - Warszawa Sp. z o.o.
3. AASHTO Guidelines for Traffic Data Programs, American Association of State Highway and Transportation Officials, 1992.
4. Federal Highway Administration (FHWA), Traffic Monitoring Guide, 2013.
5. L. Jin, Ch. Xu, J. D. Fricker, Comparison of Annual Average Daily Traffic Estimates: Traditional Factor, Statistical, Artificial Neural Network, and Fuzzy Basis Neural Network Approach, TRB 87th Annual Meeting DVD, Washington DC 2008.
6. M. Spławińska, Ocena szacowania Średniego Dobowego Ruchu w roku (SDR) przy wykorzystaniu wskaźników przeliczeniowych, Drogownictwo, 7-8/2010, 267 -272.
7. S. Granato, The impact of factoring Traffic Counts for Daily and Monthly Variation in Reducing Sample Counting Error, 1998 Transportation Conference Proceeding, Iowa State University, 08.1998, 122 ÷ 125.

8. R. T. Luttinen, O. Virpi, Quality Assessment of AADT Estimates, TRB 88th Annual Meeting CD – ROM, Washington DC 2007.
9. M. Lewis, D. Albright, Evaluating the 2000 Highway Capacity Manual adjustment factor for AWDT to AADT by applying a consistent traffic data methodology, Transportation Research Record: Journal of the Transportation Research Board 1993, Washington DC 2007, 117 ÷ 123
10. Highway Capacity Manual-HCM-2010, Transportation Research Board, National Research 2010
11. J. K. Eom, M. S. Park, T. Y. Heo, L. F. Huntsinger, Improving the Prediction of Annual Average Daily Traffic for Non-Freeway Facilities by Applying Spatial Statistical Method, Transportation Research Record: Journal of the Transportation Research Board 1968, Washington DC 2006, 20 ÷ 29
12. M. Sabry, H. Abd-El-Latif, S. Yousef, N. Badra, Determination of AADT from Short Period Traffic Volume Survey. Journal of Applied Sciences Research, 3(7): 607 - 612, 2007, s. 607 ÷ 612
13. M. Sabry, H. Abd-El-Latif, S. Yousef, N. Badra, Use of Box and Jenkins Time Series Technique in Traffic Volume Forecasting, Research Journal of Social Sciences, 2007, 83 ÷ 90
14. B. Selby, K. Kockelman, Spatial Prediction of AADT in Unmeasured Locations by Universal Kriging, TRB 90th Annual Meeting DVD, Washington DC, 2011
15. X. Wang, K. Kockelman, Forecasting Network Data: Spatial Interpolation of Traffic Counts Using Texas Data, Transportation Research Record: Journal of the Transportation Research Board 2105, Washington DC 2008, 100 ÷ 108.
16. H. Wu H., Z. Zhang, Framework for Estimating AADT Using Co-clustering Based Collaborative Filtering, TRB 88th Annual Meeting DVD, Washington DC 2009.
17. Y. Zhang, Y. Xie, Forecasting of Short – Term Freeway Volume with v-Support Vector Machines, Transportation Research Record: Journal of the Transportation Research Board 2024, Washington DC 2007, 92 ÷ 99
18. Zbieranie, archiwizacja i analiza danych ze stacji ciągłych pomiarów ruchu w roku 2008, Etap III, Analiza roczna i edycja wyników pomiarów prowadzonych w stacjach GR i PAT w roku 2009, Transprojekt – Warszawa Sp. z o.o., Warszawa 2010.
19. M. Splawińska, Określenie korzystnego czasowego zakresu pomiarów wyrwykowych do wyznaczania średniego dobowego ruchu w roku, Drogownictwo, 1/2015, 23-28.
20. M. Splawińska, Podział sieci dróg na odcinki jednorodnie ruchowo w celu zwiększenia dokładności szacowania SDR, Zeszyty Naukowo - Techniczne Stowarzyszenia Inżynierów i Techników Komunikacji Rzeczpospolitej Polskiej Oddział w Krakowie, Nr 1(103)/2014, 361-370.
21. M. Splawińska, Analiza stabilności charakterystyk zmienności natężeń ruchu w dłuższym okresie, 9902-9911, Logistyka 6/2014.

Received 01. 06. 2015

Revised 16. 06. 2015

LIST OF FIGURES AND TABLES:

Fig. 1. Map of the locations of the continuous measurement stations (source: based on [18])

Rys. 1. Mapa lokalizacji stacji pomiaru ciągłego (źródło: na podstawie [18])

Fig. 2. Comparison of the accuracy of estimates of AADT using factor approach - FA

Rys. 2. Porównanie dokładności szacowania SDR metodą wskaźnikową – FA

Fig. 3. Comparison of the accuracy of estimates of AADT using factor approach – FB

Rys. 3. Porównanie dokładności szacowania SDR metodą wskaźnikową – FB

Fig. 4. Analysis of the residuals for the model regression (month IX – X, day of the week: Tuesday – Thursday)

Rys. 4. Analiza reszt dla modelu regresji (miesiąc IX – X, dzień tygodnia: wtorek – czwartek)

Fig. 5. Comparison of AADT estimation accuracy using multiple regression models

Rys. 5. Porównanie dokładności szacowania SDR przy wykorzystaniu modelu regresji wielorakiej

Fig. 6. Comparison of estimation accuracy of AADT using ANN model

Rys. 6. Porównanie dokładności szacowania SDR przy wykorzystaniu modelu SSN

Table 1. The list of public holidays and related dates in 2010

Tabela 1. Wykaz dni ustawowo wolnych od pracy oraz bezpośrednio z nimi związanych w 2010 roku

Table 2. Characteristics of Golden River measuring stations

Tabela 2. Charakterystyka stanowisk pomiarowych Golden River

Table 3. The results of the statistical analysis of the regression model

(month IX – X, day of the week: Tuesday – Thursday)

Tabela 3. Wyniki analizy statystycznej modelu regresji (miesiąc IX – X, dzień tygodnia: wtorek – czwartek)

Table 4. Descriptive statistics for the artificial neural network models

Tabela 4. Statystyki opisowe dla modeli sztucznych sieci neuronowych

Table 5. Comparison of estimation accuracy of AADT

Tabela 5. Porównanie dokładności szacowania SDR

MODELE DO WYZNACZANIA ŚREDNIEGO DOBOWEGO RUCHU W ROKU NA DROGACH KRAJOWYCH

Słowa kluczowe: drogi, zmienność natężeń ruch, Średni Dobowy Ruch w roku (SDR), regresja wieloraka, sztuczne sieci neuronowe

STRESZCZENIE:

Jednym z podstawowych parametrów opisujących ruch drogowy jest Średni Dobowy Ruch w roku (SDR). Jest on wykorzystywany do różnych celów między innymi do projektowania i planowania rozwiązań drogowych, obliczania hałasu drogowego czy do studiów wypadkowości. Jego nieprawidłowe oszacowanie i prognozowanie może prowadzić do licznych błędów, przykładowo do niewłaściwego doboru typów skrzyżowań i niewłaściwego ich projektowania czy do przecięcia tras projektowanych na natężenie ruchu mniejsze niż to, które rzeczywiście może się pojawić. Uzyskanie dokładnych i wiarygodnych wielkości SDR możliwe jest jedynie na podstawie danych pochodzących z ciągłych automatycznych pomiarów ruchu. Niestety z większości odcinków drogowych nie ma takich danych, więc SDR musi być wyznaczany w oparciu o krótkie okresy wyrwykowych pomiarów. W tym celu najczęściej stosuje się metodę wskaźnikową.

W Polsce, wskaźniki reprezentatywne dla całej sieci zamiejskich dróg krajowych o określonym charakterze ruchu (gospodarczy, turystyczny, rekreacyjny), zamieszczane są raz na pięć lat w opracowaniu Ruch Drogowy. Na ich podstawie można wykonywać uproszczone przeliczenia natężeń dobowych na SDR. Jednakże, jak wykazano we wcześniejszych pracach, reprezentatywne typowe profile zmienności rocznej i tygodniowej dobowych natężeń ruchu są zbyt uproszczonym obrazem zmienności ruchu, co w konsekwencji oznacza zbyt małą dokładność szacunków SDR. Istnieją także inne, nowe metody, umożliwiające dokładniejsze wyznaczanie SDR między innymi: liniowa regresja wielokrotna, Analiza Wariancji (ANOVA), Szeregi Czasowe, Data Mining w tym Sztuczne Sieci Neuronowe i Co-clustering Collaborative Filterig, geostatystyczna metoda Kriginga, modele hybrydowe (FBN, ARIMA – MLANN, GA + TDNN). Trudno jednak znaleźć porównania tych metod i wniosków z nich wynikających.

W artykule przedstawiono trzy metody szacowania SDR na podstawie dobowego natężenia ruchu (Qd) tj: tradycyjna metoda wskaźnikowa (wskaźniki przeliczeniowe powiązane z cechami dróg i ich otoczenia – metoda FA; wskaźniki przeliczeniowe zamieszczone w opracowaniu Ruch Drogowy – metoda FB), opracowany model regresyjny oraz wykorzystujący sztuczne sieci neuronowe (SSN). Analizy prowadzono na danych pochodzących ze stacji pomiaru ciągłego, Golden River z roku 2010 w odniesieniu do ogółu pojazdów w przekroju drogi w trzech okresach tj. dla całego roku oraz w miesiącach dla których uzyskano największą dokładność szacowania SDR: maj – październik, wrzesień – październik dla wszystkich dni tygodnia oraz dla dni o największej dokładności szacowania SDR: wtorek - czwartek. Jako zmienne objaśniające użyto: zmienne ilościowe (Qd, udział pojazdów ciężkich [%] - uc) oraz zmienne jakościowe (dzień tygodnia, miesiąc w roku, poziom SDR, przekrój poprzeczny, klasa techniczna drogi, powiązanie przestrzenne, charakter obszaru, region Polski, charakter przenoszonego ruchu). Badania prowadzono przy różnej kombinacji zidentyfikowanych cech drogi i otoczenia. We wszystkich zastosowanych modelach statystycznie istotne okazały się zmienne: natężenie dobowe, przekrój poprzeczny drogi, region Polski i powiązanie przestrzenne. Zmienne te mają także największy wpływ na zmienną objaśnianą (SDR). W przypadku modeli regresji wielorakiej statystycznie istotne są także zmienne: charakter przenoszonego ruchu, dzień tygodnia i miesiąc w roku, natomiast pozostałe

w zależności od okresu analizy (poziom SDR - dla całego roku wszystkie dni tygodnia oraz dla miesięcy V-X i IX-X dla dni od wtorku do czwartku; uc - dla wszystkich przypadków poza miesiącami IX-X dla wszystkich dni tygodnia oraz dni od wtorku do czwartku). Statystycznie istotne są także zmienne: klasa techniczna drogi, obszar turystyczny, ruch międzynarodowy - jednakże ze względu na wysokie skorelowanie z innymi zmiennymi usunięto je z modeli. W przypadku sztucznych sieci neuronowych dla najlepszego uzyskanego modelu na 6 przypadków statystycznie istotne okazały się odpowiednio zmienne: charakter przenoszonego ruchu 5 razy, poziom SDR i miesiąc w roku 4 razy, dzień tygodnia i uc 3 razy. Ponadto w modelach SSN, w odróżnieniu od modeli regresyjnych zauważono większy wpływ zmiennych jakościowych na zmienną objaśnianą, przekraczający nawet wpływ Qd.

W przypadku metody wskaźnikowej FA istotnym czynnikiem jest dodatkowo obszar turystyczny. Wyniki te plus wysoka uzyskana dokładność modeli wskazują na prawidłowe przypisanie wpływu wyróżnionych cech drogi i otoczenia na SDR. Pomimo, że najlepszą dokładność uzyskano dla modeli SSN w szczególności w odniesieniu do większego zakresu czasowego danych (cały rok, dni od wtorku do czwartku) jednakże ze względu na brak możliwości matematycznego zapisu zależności zachodzących w modelach oraz konieczności zastosowania odpowiednio licznej próby utrudnione jest stosowanie tej metody. W przypadku modeli regresji wielorakiej uzyskano wysoką dokładność szacowania SDR jedynie dla wyselekcjonowanych danych (dni wtorek - czwartek) co znacznie ogranicza jej zastosowanie. W praktyce można wykorzystać zidentyfikowane zależności do podziału sieci dróg na jednorodne grupy zmienności ruchu i z wystarczającą dokładnością szacować SDR przy wykorzystaniu metody wskaźnikowej FA (redukcja średniej wielkości MAPE w stosunku do obecnie stosowanej metody FB w dniach wtorek - czwartek, miesiącach IX -X o 24%). Metoda ta ze względu na swoją prostotę i uniwersalność (umożliwia szacowanie wartości SDR dla dróg wyższych klas technicznych, jedno i dwu jezdniowych, o zróżnicowanym charakterze przenoszonego ruchu, w rozróżnieniu na kierunki ruchu oraz nie wymaga licznej próby) ma najmniejsze ograniczenia ze wszystkich zastosowanych metod i tym samym jest najbardziej użyteczna. Ponadto ze względu na wykazaną we wcześniejszych pracach stabilność profili w dłuższym okresie czasu, można wnioskować że określone zależności będą zachodzić także w innych latach w tym w roku prognozy.

W związku z powyższym, rekomenduje się do szacowania SDR, metodę wskaźnikową FA. Należy jednak nadmienić, że metoda ta wymaga stworzenia wiarygodnej bazy danych uwzględniającej charakterystyki drogi i jej otoczenia. Wykazano także że użyte narzędzia są przydatne w analizach danych ruchowych i mogą znaleźć zastosowanie w przyszłych analizach w miarę sukcesywnego zwiększania dostępnych danych ruchowych. Mogą być także wykorzystane przez administrację drogową do aktualizacji baz danych w tym do weryfikacji wyników pomiarów ruchu.