THE IMPACT OF PUBLIC TRANSPORT NETWORK ACCESSIBILITY ON TRIP GENERATION MODEL

SUMMARY

The most commonly used model in transport planning is the four-step model of transport demand. Although a number of improvements have been made to this model over the past six decades of use, its main weakness remains that the characteristics of the transport network are not included in the sub-model of trip generation. In the research presented in this paper the authors investigated the possibility of improving this key model. Based on the results of correlation and regression analysis it has been proven that the public transport network accessibility significantly affects the total number of generated trips. This opens up new possibilities for improving this model as well as the process of transport planning.

KEY WORDS

transport planning; transport modelling; trip generation model; public transport network accessibility;

1. INTRODUCTION

Since 1950s until today the most commonly used model in transport planning has been the four-step traffic demand model. The basic model consists of four successive sub-models that describe traffic generation, spatial traffic distribution, modal split and assignment of traffic to routes, i.e. transport network.

One of the most often mentioned weaknesses of the four-step model is the lack of the characteristic of the traffic network in the sub-model of trip generation. Since it determines the total traffic volume in a certain area, it consequently does not depend on the characteristics and quality of the respective transport network. This weakness of the four-step model has been emphasised by many authors since the 1980s to date. Manheim [1] emphasises that traffic demand needs to be a component of every phase of the traffic model including the trip generation model. Webster and Bly [2] object that the majority of models assume that the number of trips does not depend on the changes in the traffic system. Oppenheim [3] claims that in the traditional approach (four-step model) the factors that affect the trip generation are limited only to the characteristics of traffic zones or passengers, and that they do not include the characteristics of the transport mode or routes. The Handbook of Transport Modelling [4] in the Chapter on four-step model emphasises that very few models of traffic generation have included the accessibility measure in spite of intuitive objections that such variables affect the frequency of travelling. Ortuzar and Willumsen [5] critically note that the main drawback of the classical four-step model is the assumption that the changes in the transport network have no influence on the generation and attraction of trips.

Along with the said critiques there have been attempts to eliminate the mentioned drawbacks of the four-step model.

One of the first such attempts is the Hansen accessibility measure of the following form [6]:

\[ D_i = \sum_{j=1}^{n} A_j F_{ij} \]

where \( D_i \) is the measure of accessibility of zone “i” according to activities located in other zones, \( A_j \) is the intensity of activities in zone “j” (e.g. number of workplaces, number of citizens, etc.), \( F_{ij} \) is the function of resistance between zones “i” and “j”, and \( n \) represents the number of zones.

The resistance function was usually determined empirically based on the resistance function obtained from the traffic distribution model. After that, the attempts of introducing the variable that represents the
traffic network in the trip generation model were done by Shindler and Ferreri [7] and in the traffic studies of Baltimore and London. It was already from these early studies that the notion of “transport network accessibility” has been imposed as potentially the best representative of its characteristics in the trip generation model. Although there are differences in the definition of this notion, the common idea of these definitions lies in the fact that the transport network accessibility reflects its spatial availability towards the traffic system users in order to be able to realize their trip. Consequently, the accessibility is the measure of the possibility of realizing the trip in a certain area, regardless of whether the need for travelling at a certain moment does or does not exist. Due to the differences in spatial characteristics of the network of routes and network of public transport, the accessibility can significantly differ from area to area. Trip generation models should respect this fact, which is possible only if the accessibility variable is in some way included in the model itself.

One of the few comprehensive studies of transport network accessibility was performed by Leake and Huzzayin [8]. In their studies they defined ten different measures of accessibility and analysed their impact in the trip generation model of regression type at the level of the traffic zone. The analysis used the data from the Middlesbrough traffic study. Comparing the statistical indicators of model validity, they concluded that in a certain number of trip generation models a significant improvement has been achieved by introducing the accessibility variable. This primarily refers to the public transport model, although in the model that encompassed all trips a significant increase of the determination coefficient $R^2$ was achieved. The model of individual traffic showed a significantly weaker influence of the accessibility measure.

More recent academic papers have proven that there was still no satisfactory solution to the problem (Erwing, DeAnna and Li [9]). The researchers involved in transport planning and modelling are aware of the necessity of defining a more adequate variable of transport network accessibility. In spite of all the critiques in the past, the four-step model continues to be the most common one in software packages for transport planning and modelling (Meyer and Miller [10]).

Daly [11] explains that although the trip generation model is crucial in transport demand forecast, too little attention has been paid to improving it in comparison to the models that describe the selection of the transport mode or assignment of traffic to a transport network. By analysing different attempts of including the variables that describe the accessibility in the trip generation model it may be concluded that the correlation between the transport network accessibility and the total number of trips does exist, but it is weak.

Thill and Kim [12] focus on explaining the interdependence of the transport network accessibility and the induced traffic, i.e. traffic that would not exist without the new transport infrastructure. For this they used the data collected for the needs of the Minneapolis and St. Paul traffic studies. They studied the trips performed by passenger cars that originated or ended in the household and conducted a statistical analysis of the obtained results. The authors have concluded that the volume of the generated and attracted trips is under significant influence of accessibility. They noted that positive results are the consequence of different measuring of accessibility and that there is no sense in looking for a generally “best measure” of accessibility for all cases.

Kitamura [13] explains that the traffic supply and land use jointly determine the accessibility, as well as that induced traffic shows the extent of the accessibility effect on trip generation. He notes that in theory the positive correlation between the accessibility and trip generation is to be expected. The current transport planning practice continues to be based on the premise that the number of trips generated by a household is not affected by the density and quality of transport network. According to his analysis of previous studies, there are still no proofs that one can conclude with certainty to which extent the new transport infrastructure provokes induced traffic.

Litman [14] explains comprehensively the notion of accessibility and its importance in transport planning. Defining accessibility is not a simple task, as it is affected by a large number of factors classified by the author into as many as twelve groups.

This paper primarily refers to the determination of the impact of public transport accessibility on trip generation. Such decision is made as the result of the consideration of previously published scientific studies as well as the expectation that the public transport network parameters affect the volume of public transport demand to a greater extent than in the case of road network and private transport. The paper is based on a disaggregated approach, which means the research at the household level.

2. METHOD OF RESEARCH

The data was taken from the household survey, one of the basic methods in analysing the characteristics of transport demand in the Zagreb Urban Transport Study (1998 – 2000). Household survey was conducted in the urban area of the City of Zagreb. By random selection the sample of 200 households was isolated that were evenly distributed in the area. The network of public transport consisted of 15 tram lines and 113 bus lines. The tariff system in the city of Zagreb enabled the equal cost of public transport...
in the urban area regardless of the travelled distance (flat rate).

The questionnaire for the household survey consisted of three parts collecting:
- data about the household,
- data about household members,
- data about everyday trips of the household members.

The data from the first and third part of the questionnaire were used in order to assign the characteristics of the most accessible public transport lines to every household.

For the quantification of the relations between the variables in the model the correlation and regression analyses have been used. By using the correlation and linear regression analyses it is possible to determine interdependencies of all the variables. There is no evidence in the most comprehensive literature [3, 4, 5, 10, 15, 16] that any other type of regression was used in trip generation modelling. Apart from the socio-economic features of a household a certain number of variables that characterize the spatial aspect of the transport network (the public transport network accessibility) were included.

Five different accessibility measures are defined here, led by the principle of following the logic of the transport system users in making decisions on travelling:

1. Accessibility measure
   \[ d_1 = \frac{1}{\text{distance}} \]

2. Accessibility measure
   \[ d_2 = \frac{1}{\text{distance}^2} \]

3. Accessibility measure
   \[ d_3 = \frac{\text{lines frequency}}{\text{distance}} \]

4. Accessibility measure
   \[ d_4 = \frac{\text{lines frequency}}{\text{distance}^2} \]

5. Accessibility measure
   \[ d_5 = \frac{\text{lines frequency}}{\text{number of lines} \times \text{distance}} \]

In the majority of the previous studies the authors had used more complex mathematical expressions to define the transport network accessibility (usually derived from the gravitation model) with data not available to an average transport system user within the short time of decision making. Such approach is in contradiction with the intention of the traffic model to simulate the user’s behaviour.

In the correlation and regression analysis the following statistical indicators were used (\( n = 200 \)):
- correlation coefficient (\( r \)),
- determination coefficient (\( R^2 \)),
- corrected determination coefficient (\( R_k^2 \)),
- multiple correlation coefficient (\( R \)),
- partial correlation coefficient (\( R_{12 \ldots} \)),
- standard forecast error (\( Se \)).

Statistical testing of the model was performed for the level of significance of 5%:
- \( t \) - test for testing correlation coefficient significance (\( r \)),
- \( F \) - test for testing multiple correlation coefficient significance (\( R \)),
- \( t \) - test for testing the regression coefficient significance i.e. significance of independent variables,
- Durbin-Watson test for testing the possible errors of auto-correlation (when errors are neither independent nor non-correlated), occurs in the models where significant variables have been left out.

3. RESULTS

Household data

The basic values of the variables that describe the surveyed households have been presented in Table 1.

The data show large variability of almost all variables, which is expressed through high values of standard deviations. This phenomenon is specific for the disaggregated approach, since the households differ significantly one from the other, and their locations also differ in relation to the public transport network.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average value</th>
<th>Standard deviation</th>
<th>Minimal value</th>
<th>Maximal value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Number of persons in the household</td>
<td>3.09</td>
<td>1.29</td>
<td>1.00</td>
<td>8.00</td>
</tr>
<tr>
<td>2. Number of vehicles in the household</td>
<td>0.79</td>
<td>0.68</td>
<td>0.00</td>
<td>3.00</td>
</tr>
<tr>
<td>3. Distance to PT (minutes)</td>
<td>4.77</td>
<td>3.24</td>
<td>2.00</td>
<td>12.50</td>
</tr>
<tr>
<td>4. Total net monthly income (kuna)</td>
<td>5,160</td>
<td>2,770</td>
<td>1,200</td>
<td>20,000</td>
</tr>
<tr>
<td>5. Total trips per day</td>
<td>5.20</td>
<td>3.19</td>
<td>0.00</td>
<td>21.00</td>
</tr>
<tr>
<td>6. Total trips by PT daily</td>
<td>3.19</td>
<td>2.20</td>
<td>0.00</td>
<td>11.00</td>
</tr>
<tr>
<td>7. Number of PT lines</td>
<td>2.63</td>
<td>1.73</td>
<td>1.00</td>
<td>7.00</td>
</tr>
<tr>
<td>8. Frequency of PT (veh./hour)</td>
<td>19.33</td>
<td>17.14</td>
<td>2.00</td>
<td>63.00</td>
</tr>
</tbody>
</table>

Note: Abbreviation PT stands for Public Transport
Source: Zagreb Urban Transport Study - household survey (data synthesized by authors)
Out of the total number of households in the area of research 48.5% of them were at 1-3 walking minutes from the nearest public transport stop, 27% were at 4 to 6 minutes walking distance, 16% at 7 to 10 minutes and the remaining 8.5% were at more than 10 minutes walking distance.

Regarding passenger car ownership, the household distribution looks as follows:
- 34.5% of households do not own a vehicle,
- 53.5% of households own one vehicle,
- 10.5% of households own two vehicles,
- 1.5% of households own three vehicles.

**Correlation analysis**

The correlation analysis is used to determine the statistical interdependence of different variables, thus discovering the influences of one variable on the other.

Table 2 shows the correlation coefficients that show the strength of the dependence between the group of socio-economic variables (household size, vehicle ownership and total income) and the total travel volume (individual and public transport together), i.e. travel volumes by public transport. The table also includes data on the correlation between the variables describing public transport network accessibility and the total travelling volume, i.e. travelling volume by public transport.

**Figure 1** presents absolute values for "r", since the sign marks only the direction of the strength of the statistical dependence.

In order to study the interrelations of the variables for households of different categories of vehicle ownership, three groups have been derived by sorting (households with no vehicles, households with one vehicle and households with two and more vehicles). For them the correlation matrices have been developed. The results have been presented in Table 3. The accessibility measure d4 has been derived from the accessibility measure d3, but in the previous analysis it did not give any major improvements of the coefficient "r". Therefore, it has been eliminated from further analyses.

**Figure 1** - Correlation coefficient between the variables describing the household, accessibility of the public transport network and the trip volume generated by the household

Since the public transport network accessibility has been described by three basic parameters (distance from the household to the nearest public transport stop, number of available lines of public transport and the frequency of departures on a line), it should be studied how the mentioned parameters affect the number of generated trips by public transport. For this purpose, the households have been divided into two groups. The first group included the households with

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total trips</th>
<th>Trips by public transport</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of persons in household</td>
<td>0.511</td>
<td>0.358</td>
</tr>
<tr>
<td>Number of vehicles</td>
<td>0.298</td>
<td>0.156</td>
</tr>
<tr>
<td>Total income</td>
<td>0.475</td>
<td>0.237</td>
</tr>
<tr>
<td>Distance to PT</td>
<td>-0.234</td>
<td>-0.470</td>
</tr>
<tr>
<td>Accessibility measure d1</td>
<td>0.266</td>
<td>0.458</td>
</tr>
<tr>
<td>Accessibility measure d2</td>
<td>0.255</td>
<td>0.438</td>
</tr>
<tr>
<td>Accessibility measure d3</td>
<td>0.023</td>
<td>0.193</td>
</tr>
<tr>
<td>Accessibility measure d4</td>
<td>0.056</td>
<td>0.242</td>
</tr>
<tr>
<td>Accessibility measure d5</td>
<td>0.142</td>
<td>0.358</td>
</tr>
</tbody>
</table>

**Table 2** - Correlation coefficient "r" between different variables and travelling volume generated by a household

<table>
<thead>
<tr>
<th>Variables</th>
<th>0 veh.</th>
<th>1 veh.</th>
<th>2+veh.</th>
<th>0 veh.</th>
<th>1 veh.</th>
<th>2+veh.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of persons</td>
<td>0.580</td>
<td>0.367</td>
<td>0.595</td>
<td>0.578</td>
<td>0.394</td>
<td>0.397</td>
</tr>
<tr>
<td>Total income</td>
<td>0.568</td>
<td>0.388</td>
<td>0.383</td>
<td>0.542</td>
<td>0.276</td>
<td>0.042</td>
</tr>
<tr>
<td>Distance to PT</td>
<td>-0.274</td>
<td>-0.345</td>
<td>-0.221</td>
<td>-0.288</td>
<td>-0.593</td>
<td>-0.245</td>
</tr>
<tr>
<td>Accessibility measure d1</td>
<td>0.299</td>
<td>0.335</td>
<td>0.288</td>
<td>0.299</td>
<td>0.556</td>
<td>0.336</td>
</tr>
<tr>
<td>Accessibility measure d2</td>
<td>0.300</td>
<td>0.304</td>
<td>0.295</td>
<td>0.297</td>
<td>0.526</td>
<td>0.343</td>
</tr>
<tr>
<td>Accessibility measure d3</td>
<td>0.181</td>
<td>0.080</td>
<td>-0.117</td>
<td>0.158</td>
<td>0.202</td>
<td>-0.054</td>
</tr>
<tr>
<td>Accessibility measure d5</td>
<td>0.280</td>
<td>0.236</td>
<td>-0.055</td>
<td>0.274</td>
<td>0.408</td>
<td>0.111</td>
</tr>
</tbody>
</table>
Table 4 – Correlation coefficients “r” between different accessibility variables and public transport trip volume

<table>
<thead>
<tr>
<th>Variables</th>
<th>Trips by public transport</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 line</td>
</tr>
<tr>
<td>Distance to PT</td>
<td>-0.376</td>
</tr>
<tr>
<td>Accessibility measure d1</td>
<td>0.373</td>
</tr>
<tr>
<td>Accessibility measure d2</td>
<td>0.358</td>
</tr>
<tr>
<td>Accessibility measure d3</td>
<td>0.457</td>
</tr>
<tr>
<td>Accessibility measure d5</td>
<td>0.457</td>
</tr>
</tbody>
</table>

The models have been classified into three groups (Table 5):
1. models of generating all trips (all households) - models 1.1. and 1.2.;
2. models of generating public transport trips (all households) - models 2.1. and 2.2.;
3. models of generating trips by households that own one vehicle - models 3.1., 3.2. and 3.3.

Regression analysis was carried out by applying “stepwise” procedure, which insures that the model includes only the statistically significant variables, while eliminating those that are statistically not significant.

The models consider four independent variables where \( X_1 \) is a total number of persons in a household (older than 6 years of age), \( X_2 \) is a number of vehicles owned by a household, \( X_3 \) is total income of a household and \( X_4 \) is accessibility measure that represents the square root of distances of the household to the nearest public transport stop/station

\[
( d = \sqrt{\text{distance}} ).
\]

Two dependent variables have been considered, where \( Y_1 \) is the total number of trips generated by a household and \( Y_2 \) is the total number of trips by public transport generated by a household.

Independent variables \( X_1, X_2, X_3 \) are representatives of the socio-economic characteristics of the households, whereas variable \( X_4 \) represents the public transport network accessibility. It has been selected based on the previously performed correlation analysis that showed that the variable “distance to the nearest public transport stop” yielded higher correlation coefficients related to the number of generated trips than given by accessibility measures d1-d5. In order to mitigate the differences that are consequences of the relation of small numbers, and not of actual conditions, accessibility measure “d” has been adopted, which is the square root of the distance. Distance to the nearest public transport stop is measured in minutes and if one household is e.g. at a four-minute distance, and the other at a two-minute distance, then in the linear model the influence of the first value would be double. Since this is not a realistic assumption, it is more suitable to extract the root of the values, which has been confirmed by the statistical tests.

Table 5 presents seven regression models of trip generation with statistical indicators of the model quality.

The statistical indicators of model validity make it possible to objectively assess whether and in what extent variables that represent the transport network can improve the trip generation models that do not contain such a variable. The data are found in Table 6. It shows for all three groups of models the index of change of each of them after having introduced the variable that represents the accessibility of the public transport network (Index 100 would mean that there is no change).

Table 6 – Change in statistical indicators of the models that contain the transport network accessibility measure with relation to the models that do not contain it (expressed in indices)

<table>
<thead>
<tr>
<th>Models</th>
<th>( R )</th>
<th>( R^2 )</th>
<th>( R^2 )</th>
<th>F</th>
<th>Se</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st group of models</td>
<td>109</td>
<td>119</td>
<td>66</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>2nd group of models</td>
<td>131</td>
<td>175</td>
<td>170</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>3rd group of models</td>
<td>159</td>
<td>270</td>
<td>392</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>

Table 5 – Regression models of trip generation

<table>
<thead>
<tr>
<th>Model</th>
<th>Form of the model</th>
<th>( R )</th>
<th>( R^2 )</th>
<th>( R^2 )</th>
<th>Se</th>
<th>F test</th>
<th>DW test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>( Y_1 = 0.407 + 0.937 X_1 + 0.037 X_3 )</td>
<td>0.588</td>
<td>0.346</td>
<td>0.339</td>
<td>2.596</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1.2</td>
<td>( Y_2 = 2.896 + 0.923 X_1 + 0.531 X_2 + 0.029 X_3 - 1.212 X_4 )</td>
<td>0.645</td>
<td>0.416</td>
<td>0.404</td>
<td>2.466</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2.1</td>
<td>( Y_2 = 1.165 + 0.699 X_1 - 1.208 X_2 + 0.016 X_3 )</td>
<td>0.504</td>
<td>0.254</td>
<td>0.242</td>
<td>1.917</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2.2</td>
<td>( Y_2 = 4.023 + 0.726 X_1 - 0.916 X_2 + 0.100 X_3 - 1.384 X_4 )</td>
<td>0.661</td>
<td>0.437</td>
<td>0.425</td>
<td>1.670</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>3.1</td>
<td>( Y_1 = 4.115 + 0.717 X_1 + 0.036 X_2 - 1.243 X_4 )</td>
<td>0.544</td>
<td>0.296</td>
<td>0.275</td>
<td>2.727</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>3.2</td>
<td>( Y_2 = 0.040 + 0.627 X_1 + 0.019 X_3 )</td>
<td>0.433</td>
<td>0.187</td>
<td>0.172</td>
<td>2.073</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>3.3</td>
<td>( Y_2 = 4.809 + 0.634 X_1 - 1.842 X_4 )</td>
<td>0.689</td>
<td>0.475</td>
<td>0.465</td>
<td>1.667</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
4. DISCUSSION

Before discussing the results of this research it is necessary to explain the differences in the interpretation of statistical indicators of regression models between the aggregated (at the level of the traffic zone) and the disaggregated approach (at the level of household), which has been implemented in this case. Usually the aggregated regression models have much higher values of the coefficient of multiple correlation and determination than the disaggregated models. Sometimes the amounts are doubled, so that it is typical that good aggregated models have the determination coefficient in the range of \( R^2 = 0.75 – 0.95 \), and good disaggregated models in the range \( R^2 = 0.35 – 0.45 \) [17]. However, this difference does not reflect the quality of the model but is rather a consequence of the difference at the level of aggregation of the collected data. By aggregating data to the level of zone (that can have several hundred households) a number of information on the variations that exist among households are lost. Therefore, the zone regression models explain only the variations that exist among zones, whereas disaggregated regression models explain the variations that exist among households. Since it is obvious that the variations among zones are smaller than the variations among households, it is obvious why the determination coefficients for aggregated models are higher than those for disaggregated models.

Correlation analysis

When considering the total number of trips generated by one household greater dependence with this value is shown by socio-economic variables. Their correlation coefficients range around the value of 0.50 (Table 2). It is interesting to emphasise that the variable “number of vehicles in the household” shows a relatively low dependence, just slightly higher than the dependence expressed by variables that describe the accessibility of public transport network (measures d1 and d2 and “distance to the nearest public transport stop/station”). In case of the latter the sign of the correlation coefficient is negative, which shows inverse proportionality. When the trips realized by public transport are separated from the total number of trips, the picture changes significantly. The highest relation to the volume of public transport generated trips lies in the variables that represent the transport network, i.e. accessibility of the public transport network. Their highest correlation coefficients range from 0.40 to 0.50, whereas correlation coefficients that refer to socio-economic variables are much lower, in the range from 0.15 to 0.36.

In case of households that do not own passenger cars, there is a much higher level of interdependence of the socio-economic variables and the amount of generated trips, than in the case of the variables that describe the transport network (Table 3). There is a logical explanation for this. Since these households depend on the public transport the principal influential factors that affect the volume of generated trips are the number of persons in the household and the total income available to the household.

It should also be noted that in this group of households there are almost no differences in the obtained correlation coefficient for all trips and trips by public transport. This confirms the “non-elasticity of demand” in case when there is no possibility of selecting the transport modes.

Different results than the previously presented were obtained for the category of households that own one vehicle. In this group one can best determine the dependence between the qualitative characteristics of the transport network and the volume of generated trips. These are households whose members can decide whether to perform the trip by private car or by public transport, and it may be assumed that the accessibility of the public transport network will have significant impact on the travel decision. This theoretical assumption has been confirmed by this research. If one considers the impacts of different variables on the total number of trips, then it may be noted that they are almost equal (excluding the accessibility measure d3) for this category of households. That means that trip generation is similarly sensitive to socio-economic features and characteristics of the transport network. However, public transport trip generation features a much greater dependence on the quality of the transport network and its accessibility, which has been expressed in correlation coefficients that are approaching the value of 0.60, and are about 50% higher than the highest correlation coefficient assigned to a socio-economic variable. These data best show the extent of the error made when the trip generation models fail to respect the characteristics of the transport network, taking into consideration only the socio-economic characteristics of a household.

Regression analysis

Since the regression model is better the higher the values of the indicators \( R \), \( R^2 \) and \( F \), and the lower the standard forecast error \( Se \), the indices clearly show that the regression models of trip generation at the level of households have been improved after including the measure of transport network accessibility (Table 6). The majority of improvements is substantial, and only in one case in the first group of models there was a reduction of the value of \( F \), which in the concrete example is of no special significance, since even this reduced \( F \) is still fourteen times bigger than the marginal (acceptable) \( F \).
Comparing the results per groups of models it was found that least improvement was achieved in case of models that include all trips together (increase in the correlation coefficient $R$ is 9%, and determination coefficient $R^2$ 19%). It should be noted that even this improvement is not negligible. If one takes into consideration the earlier mentioned criterion, according to which the good disaggregated regression models of trip generation are found in the range of $R^2 = 0.35 - 0.45$, then it is evident that only by including the accessibility measures in the model it would fail within this range.

The second group of models showed big improvements after introducing the accessibility measure as a new variable. The increase in the coefficients of correlation and determination amounts to 31% and 75% respectively, whereas the standard forecast error decreased by 13%. However, what is even more important to notice is that the trip generation model by public transport that would contain only socio-economic independent variables could not have been used for forecasting purposes at all. Its $R^2$ amounts to only 0.242 (model 2.1). After having introduced the accessibility measure as the additional independent variable, $R^2$ increased to a quite acceptable value of 0.425.

The third group of models recorded by far the biggest improvements of relevant statistical indicators. These are the trip generation models by public transport for the households that own one vehicle, which means the largest category of households. It should be specially emphasised that the model from this group that does not contain the accessibility measure (model 3.2.) is very weak according to all indicators (e.g. $R^2 = 0.172$), whereas the model that has been assigned the variable describing the transport network accessibility has shown very good characteristics ($R^2 = 0.465$), and can be considered the best of all the studied models (model 3.3.).

### 5. CONCLUSION

Based on the quantified results of the correlation and regression analysis it has been proven that the characteristics of the public transport network significantly affect the total number of generated trips at the household level – basic units of trip generation. It should be emphasised that the definition of accessibility variables should reflect a simple transport system users’ logic in making decisions about the travelling. Taking into account this criterion, a number of accessibility measures of the public transport network and their impact on the improvement of the model have been studied in the trip generation models, depending on the type of transport demand the model has simulated. It was found that different measures can satisfy the set criterion and that at the same time they statistically significantly improve the model.

However, the trip generation model in the currently most used commercial software packages for transport modelling do not contain the variable that would describe the transport network. They are rather based on the standard series of socio-economic variables such as the household size, employment structure, total income, vehicle ownership, etc.

The results obtained by this research show that such a standard modelling procedure may be defective, especially when the transport planning studies consider different alternatives in public transport improvements. Since the transport demand model serves primarily for forecasting purposes, it is evident that it would significantly fail in estimating the volume of trips by public transport if it did not include variables of the future public transport network. In some cases such a forecasting error could result in rejecting the decision on investing into the improvement of public transport since no justification would be found in the expected volume of trips. Therefore, apart from the theoretical considerations, the results of this study can be implemented in the practice of transport planning and modelling.

Dr. sc. **DAVOR KRAŠIĆ**  
E-mail: davor.krasic@iztzg.hr  
Institut za turizam  
Vrhevec 5, 10000 Zagreb, Hrvatska

Dr. sc. **LUKA NOVAČKO**  
E-mail: luka.novacko@fpz.hr  
Fakultet prometnih znanosti Sveučilišta u Zagrebu  
Vukelićeva 4, 10000 Zagreb, Hrvatska

**SAŽETAK**

**UTJECAJ DOSTUPNOSTI MREŽE JAVNOG PROMETA NA MODEL GENERIRANJA PUTOVANJA**

Najčešće korišteni model u prostorno-prometnom planiranju je četverostupanjski model prometne potražnje. Iako je u proteklim šest desetljeća uporabe doživio niz poboljšanja, njegova glavna slabost ostaje to što obilježja prometne mreže nisu zastupljena u podmodelu generiranja putovanja. U istraživanjima koja su predstavljena u ovom radu ispitana je mogućnost poboljšanja tog ključnog modela. Na temelju rezultata korelacijske i regresijske analize dokazano je da dostupnost mreže javnog prometa statistički značajno utječe na ukupan broj generiranih putovanja. Time se otvaraju nove mogućnosti poboljšanja ovog modela kao i postupka prostorno-prometnog planiranja.

**KLJUČNE RIJEČI**

prometno planiranje; prometno modeliranje; model generiranja putovanja; dostupnost mreže javnog prometa;
REFERENCES