Building and Calibration Transport Demand Model in Anamorava Region

Ramadan DURAKU, Vaska ATANASOVA, Nikola KRSTANOSKI

Abstract: The main objective of this research is to develop a model and to calibrate it in order to apply it to transport forecasting for Anamorava region. The synthetic model has been developed, and it is composed of a transport network model and a demand model, which is enabled using PTV Visum software as well as using the following variables as input: number of residents, number of people employed, working places available as well as the volume of vehicles entering and leaving certain locations surrounding the Anamorava region at "peak hour". Required coefficients are used for converting the traffic volumes from 12 hours to 24 hours expressed, such as AADT. As a criterion, for initial verification R^2, RMSPE, Percentage Deviation and Regression parameters are used. Then, a calibration of the demand model is conducted using the TFlowFuzzy algorithm by employed GEH test, comparing it with the observed data of traffic volumes accomplished simultaneously at some locations inside and in the vicinity of this region for time intervals of 12 hours on two different days in one week in May 2016. In order to fulfil all the criterions, it has been found out that the final model may be used for transport demand forecasting in the future in this region.

Keywords: logistic; methodology; performance analysis; simulation modelling; transport demand

1 INTRODUCTION

Nowadays, transport development is directly connected to the economic and the social development of a given country or region. The rapid advancement of IT in the last years has produced possible orientation of engineering studies as precondition of transport development and perspectives towards the progress of transport demand models [1].

Transport demand models are classified as: macroscopic, mesoscopic and microscopic [2]. They can also be classified as aggregate and disaggregate levels. Based on the modelled period, there are models which represent the period of 24 h and models for every other significant period of a day or so-called "peak hour models" [3].

This research relates to 24-hour, macro type of models because it involves regional as well as aggregate level due to the collection of data according to variables at the zone level. These types of models are used for strategic studies and trips, and they are expressed as Annual Average Daily Traffic (AADT) volumes in the road transport network [4].

This study represents a possibility to develop a four step (synthetic) model in order to estimate transport demand using data on macro variables as an input and its outcome is the generation of initial origin destination (O-D) trip matrix for traffic volumes. The next step is to compare this matrix to *OD* matrix trips based on observed trips through traffic counting. An initial matrix *OD* depends on economic and demographic macro variables, the design of network traffic as well as on traffic volumes in peak hour at incoming/ outgoing of the region is used as input. The observed *OD* matrix of trips directly depends on the accuracy of counting and interviews accomplished with road side traffic participants and their processing. The adjustment of the initial *OD* matrix to an acceptable level is done using various statistical indicators and the further correction procedure is continued using GEH test and TFlow Fuzzy technique which is a special tool in PTV Visum software.

The key purpose of this research is to develop and to calibrate a road transport demand macro model through modelling techniques by PTV Visum software at the Anamorava region. At the same time, the focus is oriented towards the application of the methodology of data input to be used as well as steps to follow when development and calibrating a synthetic model and its relevance to transport demand forecasting. Another purpose concerns sensible transport planners and decision makers undertaking studies of this nature.

2 LITERATURE REVIEW

Pursuant to the literature referred, the assessment of initial *OD* matrix is done based on the following two approaches [5]:
a) Directly estimated are usually known as "origin-destination surveys" which require large surveys in field studies such as: roadside, household and terminal surveys;
b) Indirectly estimated where *OD* trips are correlated to a number of known explanatory variables to the study area. This includes: gravity approach, network equilibrium and entropy maximisation.

The first method is expensive, time consuming and with the tendency to loss validation in the short term, while the second one is less expensive because it requires low-cost data, less time and less man power [6].

Gravity approach assessment of *OD* matrix is used by some researchers such Robillard (1975), Hogberg (1976), Tamin and Willumsen (1989) and Willumsen (2003). These models work relying on linear and non-linear regression [7]. Its disadvantages can be that, since it requires zonal data for calibration of parameters on the transport demand, and also it cannot be accurate to handle external-external trips [8].

Network equilibrium approach is used in cases where the number of alternative roads between two points is lower and the congestion does not play an important role in selecting the road to follow. This is more suitable for interurban roads. Selection of static estimators for O-D matrix by researchers, done according to three ways [9]. Maher (1983), Bell (1983), Casceta and Nguyen (1988) have done it relying on Maximum Likelyhood (ML), while Casceta et al. (1984) according to Generalized Least

Squares (GLS) and Maher (1983) by Bayesian Approach (BA). These models have conceptual differences compared to gravity one because it uses impedance of travel time, and it stratifies social economic variables in order to determine the probability of trip completion. These models according to this approach bring these variables close to utility or impedance function and require more information for assessment and computational time [10].

Entropy maximization approach consists of minimized information and maximal entropy [11]. In order to define OD matrix of trips there is a need for getting the lowest numbers possible of information; when receiving this information it is not possible from counting of traffic. The entropy is an expression from physics and it relies on the supposition that many similar matrices of trips OD have more "microstates" relation. Models set up according to this approach are affected by the total number of trips and in most of the cases they do not take into account variables such as population and number of vehicle ownership so that makes them not so suitable for long term planning applications [12].

This approach has been used for the first time in transport field by Wilson (1970). Wilson and MacGill (1977) have proven that this method is partially useful to the system with bigger number of components with complex and disorganized appearance. Their theory was confirmed also by Hogberg (1975) [13]. This approach has been used and upgraded later by Von Zylen and Willumsen (1977) in order to assess OD trip matrix based on traffic counting [14].

Estimation OD matrix according to the above approaches have a serious weakness because traffic counting used as input have an inaccuracy and it is subject to sampling errors. In this regard, Xu and Chan (1993) were pioneers in assessing OD matrix by applying fuzzy weights [15]. A year later Rosinowski's (1994) developed an algorithm named TFlow Fuzzy (TFF) which traffic counting in the model considers imprecise value based on fuzzy sets theory functioning according to route choice proportion for each link calculating only one time and keeping constant throughout the process of calculation [16]. Reddy and Chakraborty (1998) developed a fuzzy bi-level inference applying maximum entropy in the upper level using assignment algorithm [17]. Shafahi and Faturechi (2009) proposed a new method for estimation of fuzzy OD matrix which takes into account link data as a fuzzy value which varies to the certain bandwidth [18].

Yousefkia et al. (2013) proposed a modified TFlow Fuzzy (TFF) for the estimation of OD matrix taking into account counting like not precious based on fuzzy sets theory. This approach has been used by Mamdohi et al. (2013) in intercity roads of Teheran province as an uncongested network [19]. This method is considered very practical and very realistic because it works according to route choice proportion for each link within the road network in which proportion is updated successively in each iteration.

3 METHODOLOGY OF RESEARCH

For the development and the calibration model for transport demand forecasting, study area is determined as well as research techniques and procedures, development

synthetic model, establishment OD matrix of actual trips, preliminary verification of results, calibration of model and conclusion.

3.1 Study Area

The development of the model has been done for Anamorava region which is situated in south-east part of Kosovo, which is composed of six municipalities (Gjilan, Partesh, Vitia, Kllokot, Kamenica and Ranillug) with 166 settlements and a total surface of 1331 km^2. It is surrounded by other regions of Serbia in east and by Macedonia in the south [20].

3.2 Research Technique and Procedures

Modelling is a process in a number of iterative steps, consisting of: design, verification, calibration and validation [21].

This research begins with the design of an initial macro model at the aggregate level of the region which includes the preparation of graphical substrate, geometric elements of the observed segments on the network, establishment OD matrix of trips through variables identified with impact in generating it etc.

Verification is a visual check of the model's operation and a confirmation of the logic of the modelling results. The initial verification of the model is a very important check whether the model settings are accurate and realistic as well as whether the results are within in range of expectations.

Apart from visual, verification can be done also by statistical indicators such as determination coefficient (R^2), Relative Mean Square Percentage Error ($RMSPE$), Percentage Deviation ($PcrDev$) and regression parameters.
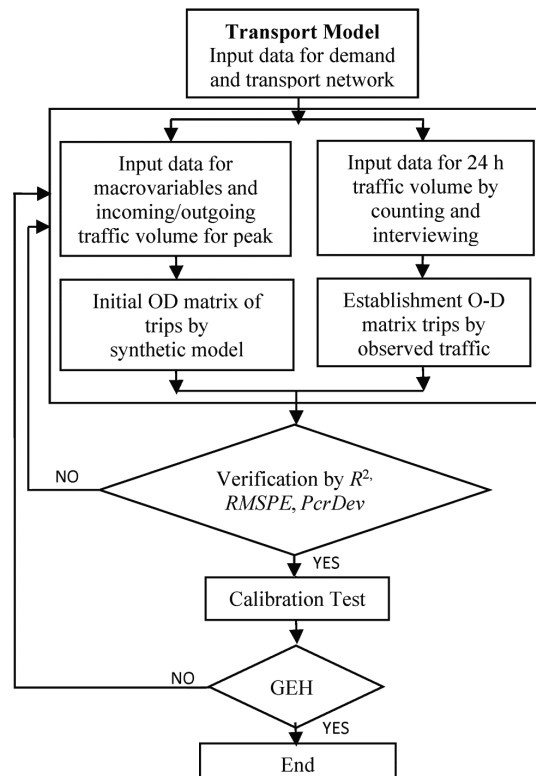

Figure 1 Flow chart of model development

This makes it possible to detect incompliance of model at early phase of establishment and in this way bigger errors, which could potentially appear in the latter phase, may be avoided.

The next step is known as calibration process. According to the Highway Capacity Manual (HCM 2010), calibration is the process of comparing model parameters with actual data obtained by counting and observed in the road network [22].

The aim is to reduce the difference between the output results of the simulation model and the data obtained by counting. It is important to identify the key input parameters that the model uses in the simulation and their impact on the output modelling results.

The procedure methodology of calibration is not finally approved but it is thought that calibration of model is the most significant requirement for the success and application of results in practice in line with new circumstances. The use of models with default values as input or by minimal requirements of calibration will bring unreal results of the model and it will contradict it.

The last step is the validation of the model which has to do with assessment of calibration process comparing model values calibrated with real data of traffic. Being aware of the impossibility to get new data in the lack of Automatic Traffic Counters (ATC), this step is ignored.

The model development according to this approach for transport demand in Anamorava region is given in Fig. 1.

3.3 Synthetic Model Development

In order to develop a synthetic model to estimate, transport demand relies on a demand and offer model [14]. For a demand, data for macro-variables and traffic volumes are required while for offer side data for current facility of road network are needed. Data for variables such as resident population, number of employed persons, number of working places have been collected through competent institutions for year 2016 [23], while traffic data have been obtained through counting. In order to develop a model, a special "meta model" is used in PTV Visum software, which is dedicated for modelling in the transport field [24], using as input data presented in the following Tab. 1.

Table 1 Data for variables in Anamorava region – 2016

No	Variables	Gjilan	Partesh	Viti	Kllokot	Kamenic	Ranillug
1	Resident population	80 525	1758	47 958	2739	33 590	3835
2	No. of employees	16 895	555	6500	665	5411	560
3	No. of workplaces by activity	14 209	318	6301	1195	3435	540
3.1	Administration	1829	129	530	128	587	139
3.2	Manufacturing + Industry	1706	6	605	184	325	5
3.3	Education	2308	95	1125	59	975	183
3.4	Recreation	105	0	98	20	10	0
3.5	Business	2197	13	1375	101	374	35
3.6	Services	4490	32	1500	120	914	96
3.7	Health	1011	9	470	438	187	39
3.8	Agriculture	762	34	653	146	63	51
3.9	Incoming traffic	559					
3.10	Outgoing traffic	591					

Clarification: Units are in thousands at annual level and they have to do with traffic at incoming or outgoing on peak hour.

Furthermore, there is a need to conduct other actions such as zoning and additional attributes related to types of trips, placing road network according to road category and their connection to centroids through connectors etc. Zoning is split into six inner zones and seven outer zones. In the inner zone there are given municipalities of the region which include Gjilan, Vitia, Kamenica, Kllokot, Partesh and Ranillug and the outer part includes 101 – Novoberdo, 102 – Llabjan, 103 – Lipjan, 104 – Ramjan, 105 – Kacanik, 106 – Preshevo, and 107 – Bujanovc, which are presented in Fig. 2.
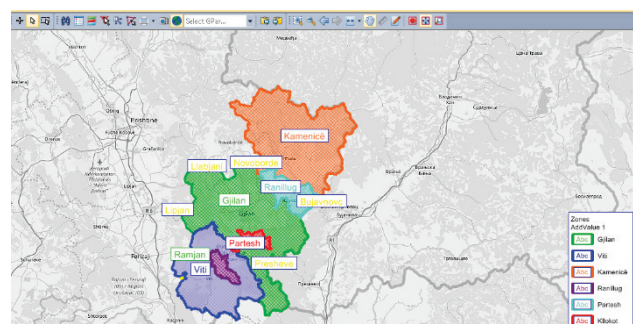

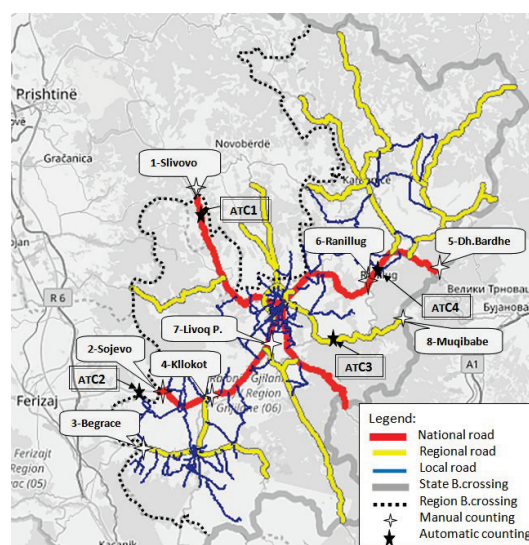**Figure 2** Inner/outer zones of Anamorava region

Figure 3 Current road network of Anamorava region

The main road network in which the traffic volume is dominant is taken into consideration. Categorisation of the road network as well as the capacity on lanes per hour relies on literature [25]. In this regard, based on traffic

volume expressed, such as AADT, this road network is composed by main, regional and local roads, as presented in Fig. 3. Each road network is expressed through nodes and links which (when connected) create a transport network. Each link is associated thereafter with defined parameters, such as: capacity per lane, number of lanes, length, speed restrictions, allowed types of transport, etc. In this regard, the road network is also composed by connectors related to so-called centroids.

One of the most important actions in the zones is inputting data attributes. Attributes are a key component in creating the model [26]. For the normal functioning of traffic attributes of networks, elements must be defined for all activities and trips for each traffic zone in the region, which comprise 10 attributes inserted in PTV Visum software, as presented in Fig. 4.
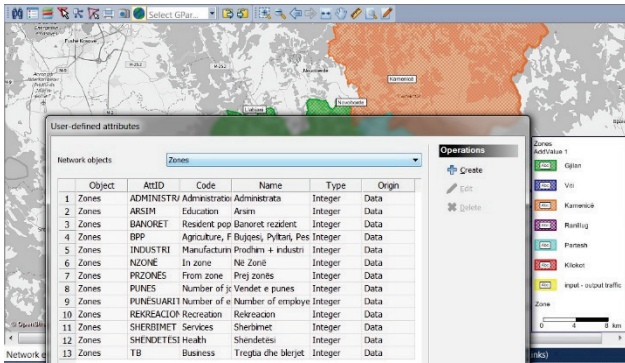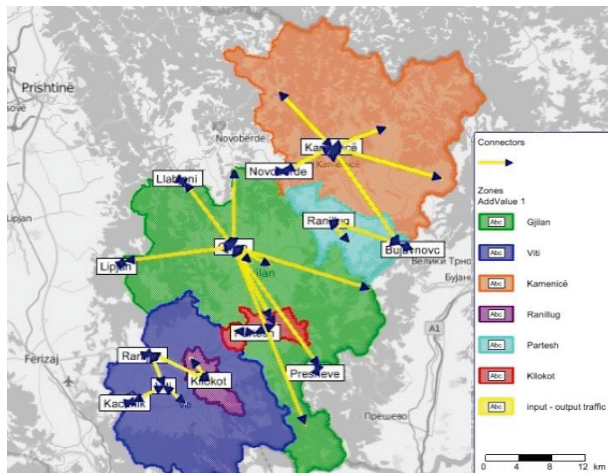

Figure 4 All attributes inserted

Figure 5 Zones, centroids and connectors

In the computer presentation of the models, zones are shown as a single point at which attributes are concentrated and which are called zone centroids. Centroids are joined with networks through connectors of centroids, which include time, distance and cost connecting the transport system with trips from origins to destinations for each zone. The position of centroids is done automatically in the centre of gravity of the zone. Centroids and their connectors play an important role in maintaining the quality of public and private transport in the road network trips from origins to destinations. Connectors connect zones to the link network. They represent the distance to be covered between a zone's centre of gravity and the connector nodes. The number of connectors depends on the load of traffic in the area, and the routes represent the

traffic volumes. Despite entities' connectors for all 13 zones (six municipalities and seven outer zones), which direct traffic volume incoming and outgoing the region and connect it to the remaining Kosovo municipalities and territories of other neighbouring countries should be taken into consideration. The connectors are presented in Fig. 5 and are shown in yellow colour.

With the intention of collecting data on traffic volumes incoming (orange) and outgoing (green) this region, roadside counting traffic volume for peak hour 16:00–17:00 was made on 15th January 2016, assisted by students of the Traffic and Transport Department at the University of Prishtina. Counting was done manually, recording each vehicle in line with the type of vehicle going through a road section as well as registration plates in each direction. The summary data are presented in Fig. 6.
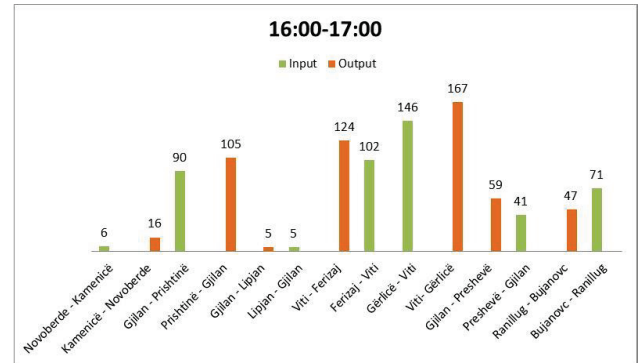

Figure 6 Traffic incoming and outgoing the region during peak hour

In this regard, it has been necessary to define outer zones, which play an important role. Moreover, for the proper functioning of the model it has been required to create a matrix for transiting this territory, particularly for passenger and good vehicles. Values on the intensity of transit traffic between outer zones are shown in Tab. 2 and Tab. 3.

Table 2 Transit of passenger vehicles – C transit (vehicles/h)

Entry/Exit	101	102	103	104	105	106	107
101	0	0	0	0	1	0	0
102	0	0	0	0	1	7	12
103	3	0	0	0	2	0	0
104	0	0	0	0	0	0	0
105	0	0	0	0	0	0	0
106	0	6	0	0	0	0	0
107	0	10	0	0	0	0	0

Table 3 Transit of commercial vehicles – TT transit (vehicles/h)

Enter/Exit	101	102	103	104	105	106	107
101	0	0	0	0	0	0	0
102	0	0	0	0	0	3	4
103	0	0	0	0	0	0	0
104	0	0	0	0	0	0	0
105	0	0	0	0	0	0	0
106	0	2	0	0	0	0	0
107	0	3	0	0	0	0	0

It is known fact that various categories of vehicles have a different influence on traffic volume. Therefore, it is done equivalent of vehicles per unit in order to have a uniform traffic influence. Obtained values are multiplied with an average coefficient $K_e = 1{,}4$ in order to produce the Passenger Car Space Equivalent (*PCSE*). This coefficient is used as recommended by literature [27].

Table 4 Converting vehicles to PCSE

Category of vehicle	Cars	Vans & Pickups	Small Buses	Large Buses	Small Trucks	Medium Trucks	Heavy Trucks
PCSE	1,0	1,2	1,2	1,6	1,4	1,6	1,8
Average PCSE	$K_e = 1,4$						

Since counting was conducted on 15th January 2016 during peak hour, it was necessary to find out AADT as well as nonlinear seasonal and weekly coefficients. Based on the results obtained by ATC administered by the Road Directorate of Kosovo which were fixed in this region, the results indicate that the average value of this coefficient is $K_s = 1,34$. Using the same data, it was found also that the weekly nonlinear coefficient for Friday was $K_w = 1,01$. In this regard, taking into account counted traffic volumes that were converted for 24 hours, multiplying them with the above coefficients, the value for traffic volume incoming and outgoing this region [14] was found. These values can be obtained by using Eq. (1)

$$AADT = VOL \cdot K_e \cdot K_s \cdot K_w \quad \text{(vehicle/24 h)} \tag{1}$$

where $AADT$ – is the annual average daily traffic, VOL – is the traffic volume, K_e – is the passenger car space equivalent, K_s – is the seasonal coefficient and K_w – is the weekly nonlinear coefficient.

When generating trips, six groups of activities that are embedded in PTV Visum software, presented in Tab. 5 are taken into account [28].

Balancing of production and attraction of trips is done by gravity model for a single constrained which is taken into account in calculating distribution of trips. The parameters assessment of gravity model is done by the realization of iterations using Tanner deterrence function of travel time as input in processing distribution as expressed in equation Eq. (2) [21].

$$f(t_{ij}) = a \cdot t_{ij}^b \cdot e^{c \cdot U_{ij}} \tag{2}$$

where t_{ij} – is the value of travel time from zone i to zone j; U_{ij} – is the travelling utility which is $U_{ij} = 1 / t_{ij}^b$; a, b, c – are the parameters.

Factors of generation for production and attraction for each one of the six purposes of trip regarding the parameters a, b and c are determined as presented in Tab.5.

Table 5 Factor distribution for parameters *a*, *b* and *c*

		Parameters		
Purpose		*a*	*b*	*c*
Purpose 1	n1 – home-work	4,3	1	−0,08
Purpose 2	n2 – home-school	3,8	1	−0,08
Purpose 3	n3 – home-business	3	1	−0,08
Purpose 4	n4 – home-shopping	3	1	−0,08
Purpose 5	n5 – home-other	4,2	1	−0,08
Purpose 6	n6 – not from home	3,8	1	−0,08

Applying iterations for every step, the matrix is established which approximate is given by distribution based on trip distance function. Initial assignment of trips

takes place afterwards in the road network in line with equilibrium assignment technique according to Wardrop's first principle. As the result of completion of all previous activities and modelling of trips by PTV Visum software according to four step (generation, distribution and assignment) traffic volume is generated and classified according to road categories expressed like AADT, as presented in Fig 7. Modal split here is ignored because vehicles are converted into unit vehicle.
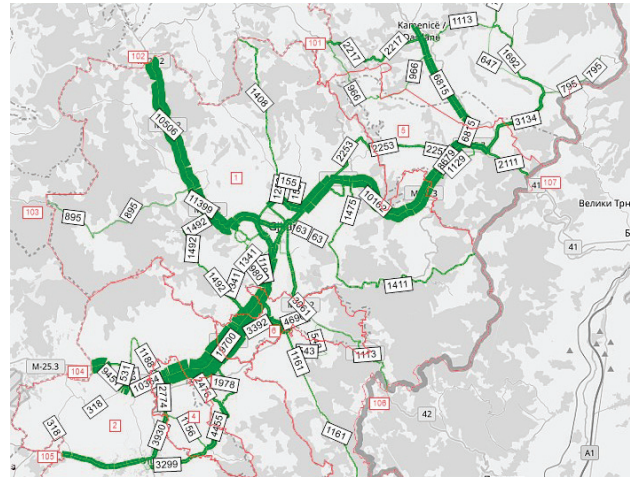

Figure 7 Traffic demand according to current situation

3.4 Establishment O-D Matrix Based on Observed of Trips

It can be easily understood that trips generated according to the synthetic model do not describe actual trips from every zone to any other zone, respectively traffic volumes to the main road segments due to the lack of accuracy of data, the lack of appropriate road designs in software or other attributes. For this purpose, research has been undertaken in order to establish an **OD** matrix of trips based on observed actual of trips using counts and interviews. The establishment of an **OD** matrix is the most important part of this research, because results gained serve for verification and calibration of model developed according to synthetic model at AADT level. The additional reason to undertake this research is also the lack of the study which should have been completed in the past in this field for this region or else called "prior matrix" which could be used as starting point for this analysis.

This is made possible by completing manual counting and interviews in roadside in some selected locations in the main road network with higher load of traffic volume. Forms and special questionnaires have been prepared in advance in order to complete counting and interviews, which contain the date and the time, the location of the interview, the occupation transport mean used for the trip, the direction of trip (OD), the purpose of the trip, the trips frequency, the number of passengers in vehicle, etc.

Depending on the origin and the destination of trips, internal–external, external–internal, internal–internal and transit trips have been taken into consideration. The trips within one zone have not been a subject in this research. Counting and interviews were completed simultaneously on 18.05.2016 and 21.05.2016 and continuously for 12 hours from 07:00 a.m. until 19:00 p.m. As a result, there were 11 523 interviews completed, comprising 19,43% of total traffic volumes, and also showing 59 317 vehicles

counted manually. Interviews were completed according to the "face-to-face" method, stopping at the maximum number of vehicles and reflecting the questionnaire answers received from traffic participants. Even here, counting and interviews were conducted assisted by approximately 100 students of the Transport and Traffic Department-University of Prishtina who were trained in advance and supported by Kosovo Police. The exact positions of locations used for manual counting and interviews are given in Fig. 3. After the research has been finalized, the collection of counting and interviews has been completed, using MS Excel. Then it has been possible to establish the **OD** matrix of trips for type of passenger vehicle as unit. It was not possible to continue further without conducting additional calculation coefficients in relation to K_e, K_{int} and $K_{convert(12 h/24 h)}$ in order to obtain a final matrix of daily trips by unit vehicle. This is the reason why vehicles have been converted into units using the coefficient $K_e = 1,4$ as reflected in Tab. 4.

Also, it is not possible to interview every vehicle driver and passenger. In order to find the number of trips realized by a vehicle within the period of 12 hours the coefficient of interview K_{int} was used. This coefficient is obtained by dividing the traffic volume and the interviews conducted within the given period of time for each location separately, as an average of two days as given by Eq. (3).

$$K_{int} = \frac{VOL_{12}}{I_{12}} \qquad (3)$$

where K_{int} – is the coefficient of interview; VOL_{12} – is the number of vehicles counting in 12 hours interval and I_{12} – is the number of interviews in 12 hours interval.

Table 6 Interview coefficient for each roadside location

No.	Location	VOL_{12}	I_{12}	%	K_{int}
1	Slivovo	5494	1131	18,01	4,86
2	Sojevo	5981	1092	15,42	5,48
3	Begrac	3137	880	24,23	3,56
4	Kllokot	4749	925	17,73	5,13
5	Dheu and Bardhe	1148	278	57,43	4,14
6	Ranillug	5350	894	16,47	5,99
7	Livoc and Poshtem	7173	1354	18,41	5,30
8	Muqibabe	571	115	59,62	4,96

Summary results regarding this proportion and interview coefficient for each location are given in Tab. 6.

By using this coefficient, it has been possible to establish OD matrix of trips for the actual total number of vehicles for a time interval of 12 hours, by applying Eq. (4):

$$OD_{matrix(12 h)} = I_{12} \cdot K_{int} \quad \text{(vehicle/12 h)} \qquad (4)$$

where $OD_{matrix(12 h)}$ – is the origin-destination matrix of trips realized by vehicle in time interval 12 hour.

After using data for traffic volumes by ATC for an interval of 24 hours, which are fixed in four locations in this region, it has been possible to obtain a converting coefficient of traffic K_{con} from 12 h to 24 h by Eq. (5).

$$K_{con} = \frac{VOL_{24}}{VOL_{12}} \qquad (5)$$

where K_{con} – is the converting coefficient of traffic from 12 h to 24 h, VOL_{24} – is the traffic volume by ATC counting in time interval 24 h and VOL_{12} – is the traffic volume by number of vehicles counting in time interval 12 hour.

As using data for four locations with ATC, the average value of this coefficient is $K_{con} = 1,5$.

With regard to the application values of traffic volumes and value of coefficients gained K_e, K_{int} and K_{con}, the final **OD** matrix of trips is obtained for vehicles for the period of time 24 h, by Eq. (6).

$$OD_{matrix(24 h)} = VOL_{12} \cdot K_e \cdot K_{int} \cdot K_{con} \quad \text{(vehicle/24 h)} \quad (6)$$

where $OD_{matrix(24 h)}$ – is the origin-destination matrix of trips realized by vehicle in time interval 24 hour.

The **OD** matrix has been established with dimensions 13×13, like an average of two days when counting and interviews are done with an intention not to require application of a weekly nonlinear coefficient of trips.

This **OD** matrix includes all municipalities of Kosovo as well as the municipalities of Bujanoc and Preshevo due to their vicinity to this region and there has been a high impact on generating trips. The obtained results according to zones which are incorporated in PTV Visum software are summarized in Tab. 7.

Table 7 OD matrix transport demand for 24 h

13x13	Name	Sum	1 Gjilan	2 Vti	3 Kamenicë	4 Kllokot	5 Ranillug	6 Partesh	101 Novobërdë	102 Llabjani	103 Lipjan	104 Ramjan	105 Kacanik	106 Presheve	107 Bujavnovc
			20396.00	5441.00	5150.00	1849.00	647.00	2661.00	403.00	6021.00	53.00	5199.00	2746.00	616.00	1292.00
1	Gjilan	18425.00	0.00	3310.00	2591.00	958.00	401.00	1653.00	198.00	4508.00	18.00	2834.00	1201.00	504.00	249.00
2	Viti	6195.00	4197.00	0.00	175.00	375.00	10.00	44.00	0.00	84.00	0.00	276.00	1014.00	0.00	20.00
3	Kamenicë	6907.00	4623.00	202.00	0.00	47.00	97.00	104.00	194.00	747.00	27.00	390.00	35.00	10.00	431.00
4	Kllokot	1630.00	642.00	312.00	27.00	0.00	0.00	18.00	0.00	57.00	0.00	454.00	66.00	5.00	49.00
5	Ranillug	988.00	344.00	11.00	202.00	0.00	0.00	58.00	0.00	71.00	0.00	150.00	0.00	0.00	152.00
6	Partesh	2554.00	1928.00	64.00	86.00	11.00	26.00	0.00	0.00	197.00	0.00	136.00	64.00	0.00	42.00
101	Novobërdë	620.00	479.00	0.00	125.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.00
102	Llabjani	5027.00	3575.00	143.00	684.00	22.00	69.00	130.00	0.00	0.00	0.00	61.00	4.00	77.00	262.00
103	Lipjan	114.00	72.00	0.00	31.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	11.00
104	Ramjan	6063.00	3456.00	392.00	760.00	332.00	29.00	622.00	0.00	63.00	0.00	0.00	354.00	15.00	40.00
105	Kacanik	2437.00	492.00	969.00	24.00	81.00	0.00	9.00	0.00	29.00	0.00	821.00	0.00	0.00	12.00
106	Presheve	484.00	275.00	0.00	69.00	0.00	0.00	0.00	0.00	101.00	0.00	31.00	0.00	0.00	8.00
107	Bujavnovc	1030.00	313.00	38.00	376.00	23.00	15.00	23.00	11.00	164.00	8.00	46.00	8.00	5.00	0.00

After the above methodology, taking into consideration all values gained this **OD** matrix is imported directly into to PTV Visum software. Modelling procedure here is simple because it does not require a complicated procedure of entering data compared to the synthetic model [29]. Modelling of traffic volumes conducted for unit of

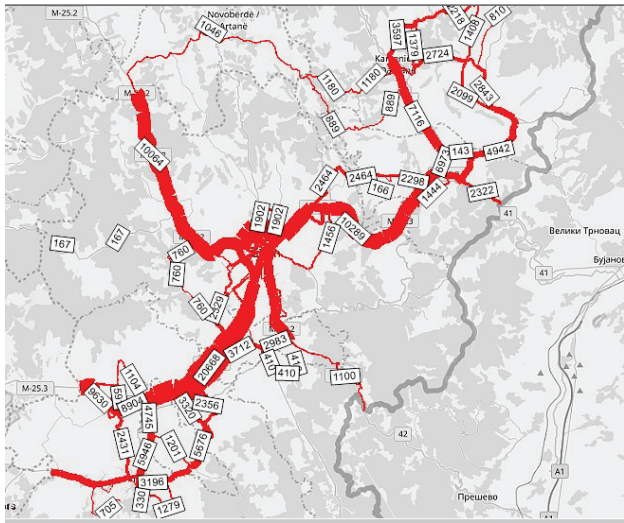vehicle category in the current road network of this region is shown in Fig. 8.

Figure 8 *OD* matrix of actual trips based on observed data

4 COMPARSION RESULTS

Once the synthetic model is developed and the ***OD*** matrix is established, the comparison of traffic volumes modelled and the ones observed is done for every location in the main road network, as presented in Tab. 9. The comparison of traffic volumes in magistral roads is done by the following indicators: $R^2 > 0.85$, *RMSPE < 20%*, *PrcDev* < 20% and regression line parameters [30].

Given indicators are useful to measure the scale performance in order to approach results and to assure an indication of the error of a model. This way is considered as quick correction. Initially there have been significant differences of traffic volumes in certain parts of the road network but after applying the process of balancing based on production and assign by equilibrium method to the traffic network it has been made possible that results are approximate to the certain level of acceptability according to iteration approach. Furthermore, based on the value gained by indicators respectively on the less relative error in percentage the synthetic model is selected with best performance and it describes closely the current situation of trips in the road network as presented in Tab.8.

Table 8 Comparison of synthetic model vs OD trip matrix

R^2	RMSPE	PcrDev	Linear regression
0,8642	17,46%	15,87%	$y = 0.9788x + 217,82$

This can be justified also by the value gained on the coefficient of determination $R^2 = 0,8642 > 0,85$, which means that in the regression equation $y = 0.9788x + 217,82$ there is a good spread of data which are displayed on the Fig. 9 and they indicate that trips by model best fit with ***OD*** matrix which are based on the observed counts.

Similarly, referring results presented in Tab. 8, also by two other indicators $RMSE = 17,46\%$ and $PcrDev = 15,87$ it is clear that traffic volumes generated by synthetic model compared to the ones according to ***OD*** matrix have much less deviation and get an acceptable level less than <20%. Results gained above are a clear indicator that there is no need to go back and adjust input data of the model or also its parameters of functioning as presented by algorithm in

Fig. 1. With the intention to approximate and further correct results of the model, there is a requirement to calibrate it using *GEH* test.
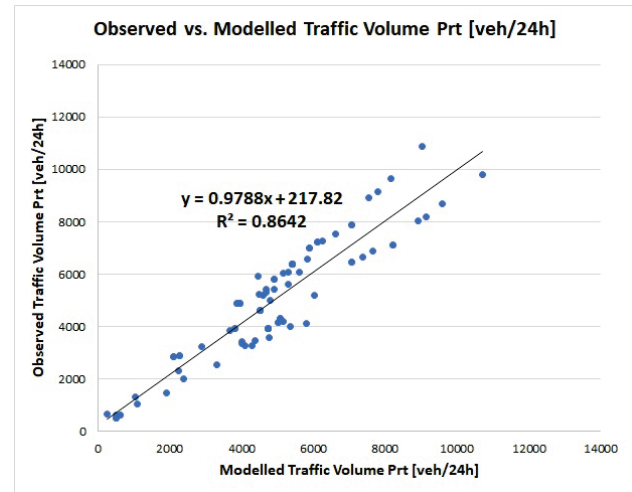

Figure 9 Demand model according to counting and interviews

5 CALIBRATION MODEL

The calculation of values of the *GEH* test is similar to statistical method χ^2 (chi-square), which takes into account the difference value obtained according to the model vs. observed. *GEH* test can be calculated by Eq. (7), as follows:

$$GEH = \sqrt{\frac{2 \cdot (O_i - M_i)^2}{(O_i + M_i)}} \qquad (7)$$

where O_i – is the value obtained by observation, M_i – is the value obtained from the model.

The *GEH* test can be applied for each segment as well as for all segments together. Based on [31], the criterion for model calibration of transport model is fulfilled if *GEH* < 5 on 85% of all location taken into account.

Firstly, the ***OD*** trip matrix has been built, and regarding it assignment to the links in traffic network, the traffic load of links f_{ij} has been obtained. Transport demand is usually expressed through the ***OD*** matrix, but due to the exclusion of all non-zero OD trips within the most convenient areas, the vector form should be used, as shown below in Eq. (8) [32]:

$$\begin{pmatrix} 0 & f_{12} & f_{13} & \cdots & f_{1n} \\ f_{21} & f_{22} & f_{23} & \cdots & f_{2n} \\ f_{31} & f_{32} & f_{33} & \cdots & f_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ f_{n1} & f_{n2} & f_{n3} & \cdots & f_{nn} \end{pmatrix} = \begin{pmatrix} f_{12} \\ f_{13} \\ \cdots \\ f_{1n} \\ f_{21} \\ f_{22} \\ \cdots \\ f_{2n} \\ \cdots \end{pmatrix} \qquad (8)$$

If they cannot provide data for all directions in ***OD*** matrices, there is the opportunity to use the observed values

in the individual location of the road network. These can comprise both **OD** matrix traffic volumes as well as link volumes. Therefore, the observed values can be presented by Eq. (9) with another vector, such as [32]:

$$v^r = (v_1 \quad v_2 \quad v_3 \quad ... \quad v_m) \tag{9}$$

where v^r – is the value of traffic volumes on links expressed in vector form.

The trips of any **OD** matrix pair provide a certain share to each traffic count location. Therefore, the observed volumes correspond to the sum of all **OD** matrix trips which are carried out on this link. Thus, by taking it into account, the general form is given by Eq. (10):

$$A \cdot f = v \tag{10}$$

where A – is the pair of **OD** matrix share on a certain link, f – is the **OD** trip matrix, v – is the traffic volume of links.

For OD traffic observed values, A is especially constant. As a result, very large number of combinations in the **OD** matrix values is obtained according to the directions f_{ij} that should match best with the counting performed for the traffic volumes. In order to select the best matrix, out of all possible matrices according to [33], by using the method of entropy maximisation, the objective function $q(f)$ given by Eq. (11) and Eq. (12) is applied:

$$\text{maximise } q(f) = -\sum_{ij}^{p} f_{ij} \cdot \ln \frac{f_{ij}}{\tilde{f}_{ij}} - f_{ij} \tag{11}$$

So that $A \cdot f = v$ (12)

where $\tilde{f}_{ij}$ – is the traffic demand on one OD pair of the initial **OD** matrix, f_{ij} – is the traffic demand on one OD pair of the corrected **OD** matrix, p – is the number of non-negative elements of the **OD** matrix.
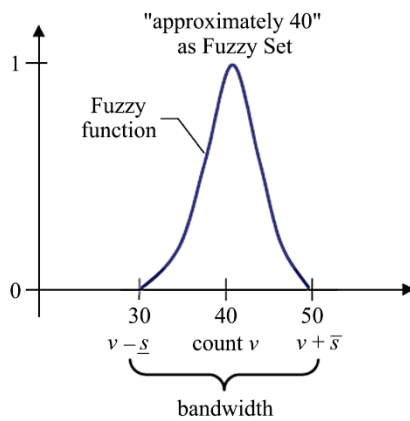

Figure 10 Fuzzy logic set [32]

With regard to the description of the above procedure used to find the function of maximum $q(f)$, so as to find the best fit, there are also shortcomings because the observed traffic figures expressed with symbol v and it can be treated as fuzzy sets, as they can express non-linearity from the area of origin up to 20% from day to day [32]. Considering the case of the correction matrices of OD, observed traffic values should be replaced by Fuzz Sets $\tilde{v}_s$ with varying

bandwidths, as given by Eq. (13) and Eq. (14) and presented in Fig. 10.

$$\text{max } q(f) = v \tag{13}$$
$$\text{such that } A \cdot f = \tilde{v}_s \tag{14}$$

Comparing it to simple intervals representing observed as Fuzzy Sets, allows favouring observed values close to the mean value. Values which are closer to the bandwidth border are also accepted if the decrease in the fuzzy membership function is more than the offset by the increase in the entropy function. Unfortunately, in general, fuzzy membership functions do not have the analytical properties (continuous, differentiable) that are required for an efficient solution algorithm. The counted values of the traffic load are substituted by fuzzy set $\tilde{q}$ with the bottom value of deviation $\underline{s}$ and the upper value of deviation $\overline{s}$ [32]. Now, the problem of maximisation takes the form:

$$\text{maximise } q(f) + q(\underline{s}) + q(\overline{s}) \tag{15}$$
$$\text{such that } A \cdot f + \overline{s} = \overline{v} \tag{16}$$
$$A \cdot f - \underline{s} = \underline{v} \tag{17}$$
$$\overline{s} \geq 0 \tag{18}$$
$$\underline{s} \geq 0 \tag{19}$$

where:

$$q(\underline{s}) = -\sum_{i=1}^{m} \underline{s}_1 \cdot \ln \frac{\underline{s}_1}{\hat{\underline{s}}_1} - \underline{s}_1 \tag{20}$$

$$q(\overline{s}) = -\sum_{i=1}^{m} \overline{s}_1 \cdot \ln \frac{\overline{s}_1}{\hat{\overline{s}}_1} - \overline{s}_1 \tag{21}$$

$\overline{v}, \underline{v}$ – are the maximum and the minimum of Fuzzy set,

$\overline{s}, \underline{s}$ – are the deviations from the observed traffic count,

$\hat{\overline{s}} = \overline{v} - v$ – is the upper bandwidth for observed traffic count,

$\hat{\underline{s}} = v - \overline{v}$ – is the lower bandwidth for observed traffic count.

Incorporating slack variables $\overline{s}, \underline{s}$ into the weighted entropy maximisation gives preference to matrices which achieve $A \cdot f = v$ within the bandwidth "as good as possible". The formulation that produced the problem with the help of the fuzzy set is structurally identical to the original one of the problems which had the exact limitations.
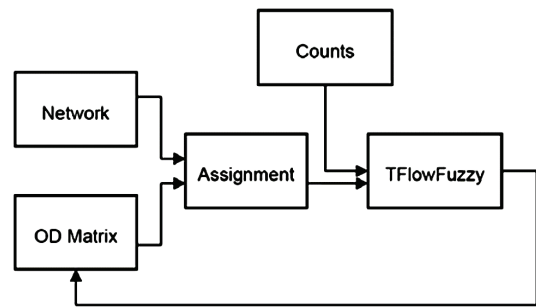

Figure 11 TFlowFuzzy for the correction of the initial synthetic matrix

The table below shows that the GEH index for all locations, except for segment No. 28 is lower than 5 (*GEH*

< 5). The total of seven locations gives 14 directions corresponding to this criterion. Thus, it may be concluded that the traffic model is suitable. Due to incompatibility, TFlowFuzzy is used for calibration in which correction is carried out in these seven links, as presented in a schematic way through Fig. 11 [34].

Tabs. 9 and 10 show general assessment, e.g. *GEH* test values, with absolute difference of traffic flow modelled and that which was counted, expressed as AADT levels.

Table 9 Summary of *GEH* test indicators

Evaluation aggregate	
GEH: Avg.	2
GEH: <5.0	86%
Deviation: Avg.	3%
Deviation: Avg. weighted	4%

Table 10 Presentation of results of *GEH* test

Link		Observed	Model	Evaluation	
No.	From node	Count 2016 (veh/24h)	Volume *PrT* (veh/24)	*PcrDev.* (%)	GEH
28	109	13 123	14 009	6%	5,5
28	110	12 960	11 932	−8%	6,7
88	225	1999	1938	−3%	1
88	339	2166	2073	−4%	1,5
92	8	10 178	10 119	−1%	0,4
92	11	11 411	11 223	−2%	1,3
145	403	5210	5273	1%	0,6
145	404	4326	4264	−1%	0,7
921	309	9916	10 309	4%	2,8
921	739	9397	8997	−4%	3
1048	607	10 004	9836	−2%	1,2
1048	827	9831	10 125	3%	2,1
1071	278	918	960	4%	1
1071	843	1159	1155	0%	0,1

Besides this, in order to calculate the quality of calibration, the other statistical parameter R^2 is used. If the value is closer to 1 it shows better compliance with model values the counted ones, whereas if the value is closer to 0 it shows incompliance. A value from 0,9 to 1 is considered acceptable. This value is $R^2 = 0,9912$, which means that it is exceptionally good.

In the Fig. 12, the relation between this characteristic is expressed also by equation $y = 0,9981x − 13,568$.
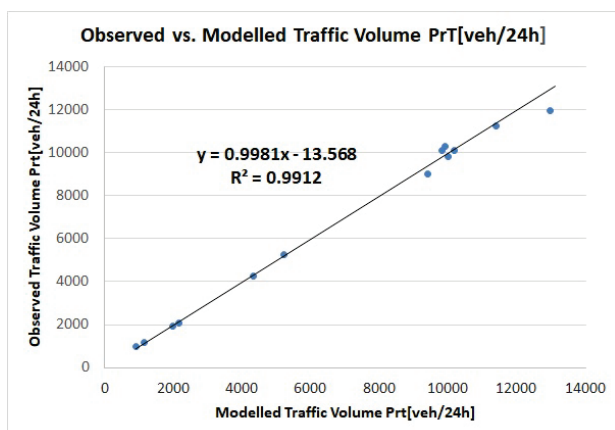

Figure 12 Demand model according to counting and interviews

6 CONCLUSIONS

The aim of this study has been to develop a synthetic model for transport demand in the Anamorava region using PTV Visum software.

The development of this model has been done by employing demographic, social economic and traffic input variables describing transport demand on one side and road and transport network on the other side which describe the offer.

OD actual trips matrix has been established in order to verify the model through field research conducted in observing the traffic through counting and interviews.

Results of this *OD* matrix are used for rough verified of traffic volumes with the ones generated by synthetic model according to statistical indicators R^2, *RMSPE*, *PcrDev* and regression line.

Once these criterions have been fulfilled, the model calibration has been done using *GEH* test using traffic volumes data based on observed traffic at the seven locations of this region.

Pursuant to the results of analysis it has been clear that criteria according to parameters of the *GEH* test as well as coefficient R^2 have been fulfilled and the result is that the final model is reliable and it can be used for transport demand forecasting in this region in the future.

This study can also be used like the feasibility study to prepare the transport strategy for this region which could also include maintenance and reconstruction of actual network, or facilitate decision making for new investments in the transport infrastructure, further upgrading of road network to the level of motorway and finding the best variants for its situating etc.

6 REFERENCES

[1] Mitran,G., Ilie, S., & Nicolae, V. (2014). The Calibration of Macroscopic Transport Sub-models through Traffic Counting. *The 2nd International Conference on Traffic and Transport Engineering (Beograd)*, 132-140.

[2] Kutz, M. (2004). *Handbook of Transportation Engineering, Volume II: Applications and Technologies*. Second edition, McGraw-Hill Professional.

[3] Zdenek, R. & Dytrych, J. (2012). Peak-period Models and 24-hour Models Comparison. *Proceedings of the 2nd International Conference on Environment, Economics, Energy, Devices, Systems, Communications, Computers, Mathematics (EDSCM'12)*, 175-180.

[4] Luburić, G., Šoštarić, M., & Slavulj, M. (2011). Measuring of Transit Traffic in Cities. *Tehnicki vjesnik, 18*(4), 619-625. https://hrcak.scrce.hr/75412

[5] Zochovska, R. (2012). Dynamic Approach to the Origin-Destination Matrix Estimation in Dense Street Networks. *The Archives of Transport*, XXIV(3), 390-413. https://doi.org/10.2478/v10174-012-0025-1

[6] Kresnanto, N. C. & Tamin O. Z. (2009). The Impact of Location and Number of Traffic Counts in the Accuracy of O-D Matrices Estimated from Traffic Counts. *Proceedings of the 14th HKSTS International Conference: Transportation and Geography.*

[7] Bera, S. H. & Rao Krishna, K. V. (2011). Estimation of Origin-Destination Matrix from Traffic Counts: The State of the Art. *European Transport University of Trieste*, XVII(49), 2-23.

[8] Lo, H. P., Zhang, N., & Lam, W. H. K. (1999). Decomposition Algorithm for Statistical Estimation of OD Matrix with Random Link Choice Proportions from Traffic Counts. *Transportation Research Part B: Methodological, 33*(5), 369-385. https://doi.org/10.1016/S0191-2615(98)00042-3

[9] Oyang, L. Q. & Huang, D. (2018). An Activity-based Model for the Estimation of Origin-Destination Matrix on Road Networks. *Tehnicki vjesnik, 25*(5), 1530-1537. https://doi.org/10.17559/TV-20180527042248

[10] Abdel-Aal, M. M. M. (2014). Calibrating a Trip Distribution Gravity Model Stratified by the Trip Purposes for the City of Alexandria. *Alexandria Engineering Journal, 53*(3), 677-689. https://doi.org/10.1016/j.aej.2014.04.006

[11] Deng, Q. & Cheng, L. (2013). Research Review of Origin-Destination Trip Demand Estimation for Subnetwork Analysis. *Proceedia - Social and Behavioural Sciences*, 96, 1485-1493. https://doi.org/10.1016/j.sbspro.2013.08.169

[12] Timms, P. (2001). A Review of Methods for Estimating Origin-Destination Trip Matrices using Link Counts. *Institute of Transport Studies, University of Leeds*, Working Paper 556. https://doi.org/10.1080/01441640118851

13) Hogberg, P. (1976). Estimation of Parameters in Models for Traffic Prediction: A Non-linear Regression Approach. *Transportation Research, 10*, 263-265. https://doi.org/10.1016/0041-1647(76)90059-9

[14] Willumsen, L. G. (1978). Estimation of an O-D Matrix from Traffic Counts-a Review Working Paper. *Institute of Transport Studies, University of Leeds, Leeds, UK.*

[15] Xu, W. & Chan, Y. (1993). Estimation an Origin-Destination Matrix with Fuzzy Weights. Part 1: Methodology. *Transportation Planning and Technology, 17*(2), 127-144. https://doi.org/10.1080/03081069308717506

[16] Gangeraj, E. N., Behzadi, G. A., & Behzad, R. (2017). Estimation of Origin-Destination Matrix from Traffic Counts Based on Fuzzy Logic. *Civil Engineering Journal, 3*(11), 1166-1175. https://doi.org/10.28991/cej-030946

[17] Reddy, K. H. & Chakroborty, P. (1998). A Fuzzy Inference Based Assignment Algorithm to Estimate O-D Matrix from Link Volume Counts. *Journal of Computers, Environment and Urban Systems, 22*(5), 409-423. https://doi.org/10.1016/S0198-9715(98)00037-4

[18] Shafahi, Y. & Faturechi, R. (2009). A New Fuzzy Approach to Estimate the O-D Matrix from Link Volumes. *Transportation Planning and Technology, 32*(6), 499-526. https://doi.org/10.1080/03081060903374700

[19] Yousefkia, M., Mamdoohi, A. R., Maridpour, S., Noruzoliaee, H., & Mahpour, A. (2013). A Study on the Generalized TFlowFuzzy O-D Estimation. *Proceedings of Australasian Transport Research Forum 2013*, 2-4 October.

[20] Kosovo Census Atlas. (2013). *Geographic and Administrative Division of Kosovo*, ASK, Prishtina.

[21] Ortuzar, J. & Willumsen, L. (2009). *Modelling Transport*, 4th Edition. Chichester: Wiley.

[22] TRB, Highway Capacity Manual 2010. *Transportation Research Board-Special Report 209*, Edition 2010.

[23] General Statistics of Kosovo (2014). ASK, Prishtina.

[24] PTV Vision Visum 15 User's Manual.

[25] Grujić, A. & Jokanović, I. (2013). Classification and Categorization of State Roads – Comparative Analysis of Regulations. *Proceedings of Papers of Civil Engineering Faculty of Subotica*, 22, 159-171. https://doi.org/10.14415/zbornikGFS22.015

[26] Atanasova, V., Angelevska, B., Krstanoski, N., & Cvetanoski, I. (2017). Calculation of Transport Demand by Applying Software PTV Vision Visum. *International Scientific Journal. Machines, Technologies and Materials (MTM),* 3, Sofia, Bulgaria.

[27] Mikolaj, J. & Remek, L. (2015). Traffic Flow Modelling based on ISOHDM Study. *Procedia Engineering, 111*, 522-529. https://doi.org/10.1016/j.proeng.2015.07.126

[28] Haxhipetkovska, K. (2015) Creation of Macro Model for Transport Demand for City and Region. *Thesis*. University "St Kliment Ohridski" Faculty of Technical Sciences Bitola, Macedonia.

[29] Trost, D. (2005). Calibration and Validation of Traffic Models. *Thesis*. Univerza v Ljublani, Faulteta za Gradbenistvo in Geodezijo. UDC.519/61/.64:656(043.2)

[30] Savrasovs, M. & Pticina, I. (2016). Methodology of OD Matrix Estimation Based on Video Recordings and Traffic Counts. *Proceedia Engineering, 178*, 289-297. https://doi.org/10.1016/j.proeng.2017.01.116

[31] TII Publications PE-PAG-02015. (2016). *Project Appraisal Guidelines for National Roads Unit 5.1 – Construction of Transport Models*. Transport Infrastructure Ireland (TII).

[32] Friedrich, M., Nökel, K., & Mott, P. (2000). Keeping Passenger Surveys Up to Date: A Fuzzy Approach. *Transportation Research Records*, No. 1735, 35-42. https://doi.org/10.3141/1735-05

[33] Novačko, L., Šimunović, Lj., & Krasić, D. (2014). Estimation of Origin-Destination Trip Matrices for Small Cities. *Promet – Traffic & Transportation*, 26(5), 419-428. https://doi.org/10.7307/ptt.v26i5.1501

[34] PTV AG. (2008). *How to work with TFlow Fuzzy*? Official Manual, PTV Vision Visum, Karlsruhe.

Contact information:

Ramadan DURAKU, PhD
University of Prishtina "Hasan Prishtina"
FIM – Department of Traffic and Transport
10 000 Sunny Hill, n.n., Prishtina
ramadan.duraku@uni-pr.edu

Vaska ATANASOVA, Prof. Dr.
Faculty of Technical Sciences,
Department for Transport and Traffic Engineering,
Ivo Lola Ribar bb, 7000 Bitola, Macedonia
vaska.atanasova@uklo-edu.mk

Nikola KRSTANOSKI, Prof. Dr.
Faculty of Technical Sciences,
Department for Transport and Traffic Engineering,
Ivo Lola Ribar bb, 7000 Bitola, Macedonia
nikola.krstanoski@uklo-edu.mk