PODEJŚCIE PRZEDZIAŁOWE DO MODELOWANIA POPYTU Z WYKORZYSTANIEM WEWNĄTRZMIEJSKICH ROZKŁADÓW DŁUGOŚCI PODRÓŻY[1]

Petro Horbachev
prof., ScD,Kharkiv National Automobile and Highway University, ul. Petrovskogostr., 25, 61002 Kharkiv, Ukraine, tel.: +38 057 707 3783, e-mail: gorbachev_pf@mail.ru

Stanislav Svichynskyi
MSc, Postgraduate,Kharkiv National Automobile and Highway University, Petrovskogo str., 25, 61002 Kharkiv, tel.: +38 057 707 3783, e-mail: stas_svichinsky@ukr.net

Streszczenie. *Analiza istniejących metod modelowania popytu pozwala na zdefiniowanie przedziałowego podejścia do obliczania macierzy podróży jako preferowanego do określania popytu w publicznym transporcie. Podejście to uwzględnia probabilistyczną naturę formułowania relacji „dom – praca", które powodują, że pewne podróże mają jednakowe długości. Taka regularność długości tych podróży może być uwzględniona w rozkładzie częstotliwości, która staje się bazą do rozważań nowego podejścia do obliczeń macierzy podróży. Takie podejście pozwala na uzyskanie przedziałów możliwych wskaźników jakości obsługi w miastach na Ukrainie.*

Słowa kluczowe:*popyt na transport, podróże do pracy, rozkład długości podróży*

1. Introduction

Public transport (PT) is one of the major providers of transportation services for citizens at the present stage of society development. It provides a considerable percentage of employers with travellingfrom residences to workplaces. This fact proves the significant role of public transport in the city spatial development that is reflected in trip length distribution (TLD), which defines a relation between trip length and its probability.

There isa number of scientific recommendations to make correct management decisions on developingthe urban transport system, but the key role of their justification is played by theavailability of transport system model. One of the main components of the model is atrip matrix.A number of papers on trip matrix determining has been published, but the proposed approaches either require improvement or put the matter into question.

Both the public transport and its network play animportant role in trip matrix and TLDforming, suites expedient to develop a different approach to determine

1 Wkład autorów w publikację:Horbachov P. 50%, Svichynskyi S. 50%

transport demand.This approach should bebased on **the direct use** of trip length frequency distribution (TLFD).

2. The brief review of existing methods of transport demand calculation

The overload of road network still remains one of the transport problems that modern cities arefacedwith. Almost always it occurs during the rush-hours, especially in the morning. The biggest number of this timetrips constitutes the workones [6], the considerablepart of which is covered by public transport.A great number of papers are devoted to the problem of defining the demand for home--based work tripsthat is caused by stability of this type movements and their time determinacy. There are three basic approaches to the transport demand calculation: statistical, synthetic and probabilistic ones.

The statistical approach involves the use of statistical methods of trip matrix calculation that allow the available trip matrix obtained from the previous study to beapplied tothe future analysis period usinggrowth rates (growth factors) [1, 6]. The advantage of the methods covered by this approach is their simplicity. Among the disadvantages thereare: theneed for factual trip matrix,usefulness only for short-term planning,disregardof the changes in transport supply.

In the synthetic approach in order to obtain the trip matrix synthetic models based on the assumption about similarity between the processes in transport systemand the natural laws are used [1, 6]. The most widespread models used within the given approach are gravity and entropy ones.

The gravity model is based on the analogy of general trip making behaviour in transportation system with the law of gravitation. The number of trips between transport zones is determined depending on disutility of a travel, which is usually represented by distance, time or cost oftravel, i.e. transportation factors[6].

The entropy model is based on the analogy with thermodynamic processes. Trip matrix determination under this model consists in system entropy maximization that leads to receiving the matrix, which is characterized by the greatest number of its states. To make trip distribution according to this model the constraint on total transportation costs in the transport system is usually used [6, 7].

The disadvantage of the entropy model consists in its usefulness only forclosed transport system. The common lacks of both gravity and entropy models are: 1) consideration of the transportation factor as the only one during trip matrix calculation, while this calculation is multifactorial; 2) uncertain similarity of natural processes to transportation ones; 3) point estimate of trip matrix. Also, it should be noted that any results on testing the usefulnessof both models to predict the trips between zones which was conducted bycomparing the calculated trip matrices with factual ones have not been foundyet.

These disadvantages of synthetic models led to the origin of probabilistic approach to transport demand calculation. Within this approach the special attention should bepaid to discrete choice models that are based on the assumption that each individual makes a routechoice from a predetermined set of available alternatives and this choice is based on socio-economic characteristics of the alternatives and their relative utility for an individual [2,6]. At the same time, the accuracy estimation of these models is very complicated by the need for real information about population travelling that can be received by means of proper surveys.

All the reviewed models have univalent algorithm of calculatingthenumber of trips between zones and give the point estimate of trip matrix. Cardinally different approach to determine the demand for public transport services is framed by the interval conception of transport demand modelling which can also be referred to the probabilistic approach [3].

According to this conception the hypothesis that the demand for work trips is random and described by not only one trip matrix but by the list of trip matrices is put forward. Each matrix of the list describes one possible state of transport demand. The interval of possible demand states is defined by certain boundaries thatcorrespond to matrices with minimum and maximum passenger-kilometres performed during the passengers'movementsinthe city [3].

In contrast with all other models of trip matrix calculation, the conception involves the calculation of estimators that define the differences between certain matrices. These estimators are – the difference betweenminimum passenger-kilometres, which isthe result of trip making, and– the deviation of tripnumbers in matrix cells:

$$\Delta W = \sum_{i=1}^{N_A}\sum_{j=1}^{N_A} h_{ij} l_{ij} - \sum_{i=1}^{N_A}\sum_{j=1}^{N_A} h'_{ij} l_{ij} \Delta H = \sqrt{\sum_{i=1}^{N_A}\sum_{j=1}^{N_A} (h_{ij} - h'_{ij})^2} \quad (1)$$

where:
N_A – number of transport zones,
h_{ij} – number of trips between zones*i*and *j*ininitialmatrix,
l_{ij} – shortest distance between zones *i*and *j*, km,
h'_{ij} – number of trips between zones *i* and *j* in experimental (new) trip matrix [3].

The disadvantage of this conception is too wide interval of possible trip matrix states that assesses demand for transport rather 'fuzzy'. The elimination of this disadvantage is possible due to defining the ways of reduction the mentioned interval.

As a result of the review of existing models of trip matrix calculation it should be mentioned that there is no generally accepted approach to the trip distribution process. **Most models give a point estimate of the matrix and do not consider the probabilistic nature of the home-based work trips**('home – work' pair forming). Moreover, **the vast majority of scientific works which deal with the transport demand definition applies only two types of initial data for the trip distribution**among which there arethe total number of trips originated in and attracted to the transport zones as well asa transportation factor (transportation cost).

3. Theoretical background of TLFD-based trip matrix calculation method

Multifactorial formation of transport demand, on one hand, and the prevalence of synthetic models of trip matrix calculation based only on transportation factors, on the other hand, causethe need to assess the practical application of thesemodels. To this purpose it is reasonableto examinethe example of using the gravity and entropy models for calculatingthetransport demand in such Ukrainian cities as Sverdlovsk, Kupiansk and Balakliia (Kharkiv region), which is given in [4]. In this work the accuracy estimation of mentioned models was carried out. The estimation became possible due to therelatively small size of citiesthat allowed obtaining the factual data about trips between all transport zones. These data were used to get the total number of trips, generated and attractedby each zone,being the basis for calculations according to aforementioned models. During calculations, different 'deterrence functions',including those received by direct calibration of parameters in order to achieve the minimum difference between factual and calculated numbers of trips between zones, were used. The trip matrix estimation was carried out by indices given in expression (1). The results concerning the matrices to be the closest to the factual trip matrix are given in Tab. 1.

Tab. 1. Results of accuracy estimation of synthetic models

City	***Model of trip matrix calculation***	***Difference betweenpassenger-kilometres, ΔW***		***Deviation of number of trips, ΔH***	
		pass-km	***%***	***passengers***	***%***
Sverdlovsk	entropy	-3,1	0	452	8,7
	gravity	-1546	-7,1	451	8,7
Kupiansk	entropy	121,9	4,0	759	100,0
	gravity	112,5	3,7	766	100,9
Balakliia	entropy	87,4	2,2	1030	99,3
	gravity	-283,2	-7,3	1012	97,6

The data presented in Tab. 1 show that even though ΔW the difference is small, the values of ΔH are quite high. For the other trip matrices (except those listed in Tab. 1) ΔW takes the values up to 58.1%, and ΔH – up to 166.3% [4]. It means that the most widespread gravity and entropy models are not always adequate to predicttrips between zones.

This fact casts doubt on the appropriateness of using the synthetic models in practice and approves the interval conception that takes into account the probabilistic nature of forming the 'home – work' pairsand covers all possible trip matrix states, which can exist under certain conditions. The reduction of the interval of possible trip matrix states (i.e.elimination of interval conceptiondisadvantage) is possible by **the direct use of TLFD** in trip distribution becauseTLFD is the result of transport demand realization and it connects the trip distance and the probability of trip making. It should be noted that nowadaysthe TLFD is used only for calibration of trip matrix calculated with various models, including gravity and entropy ones. The calibration can be performedeither by means of 'deterrence function' or by changing the total number of trips generated and attracted by zones. The last mentioned method cannot be justified because it supposes the conscious mutilation of initial data.

To understand the possibility of direct use of TLFD for trip distribution it is necessary to consider trip length distribution as an outcome of the trip matrix and the matrix of distances between transport zones. From this point of view the process of formingthe TLFD consists in h_{ij} duplicating each l_{ij} distance in the linear set of trip distances, if h_{ij} is the number of trips between zones i and j, Fig. 1.

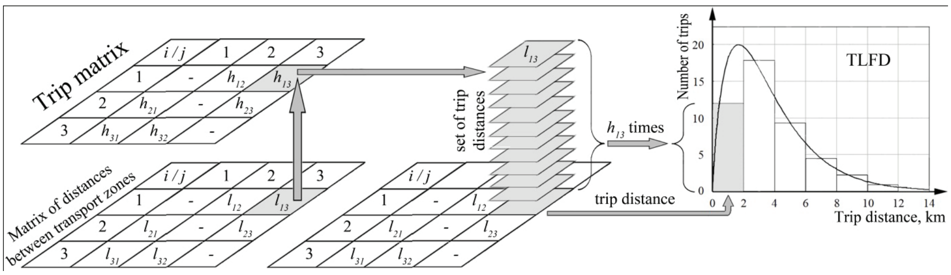

Fig. 1. Mechanism of forming the trip length frequency distribution

The mechanism presented inFig. 1 allows realizing the factthat the same TLFD can be generated by different trip matrices because trip distances grouped in each interval of distributionlead to a total interval frequency at different number of duplications of each separate distancewithin this interval. This explains ΔW small values given in Tab. 1 and indicates the reasonability of the useof interval conception of transport demand modelling and random filling of tripmatrix cells as the basis for developing of a new methodoftransport demand definition.

To develop the method based on the direct use of TLFD the similarity of general views of the distribution pattern and differential form of Gamma distribution should be taken into account[6]. Also, it is necessary to know the reasons causing this form of distribution pattern. At present,these reasons can be revealed by analysing the transport networks of PT due to the presence of such network elements as stops because they are the places of concentration of urban population trips, i.e. the places of interaction of passengers that need to travel (transport demand) and public transport infrastructure (transport supply). All other characteristics of transport networks of PT, including link lengths (distances between adjacent stops) and distances between pairs of stops (or between transport zones), are derived bystop spatial location [5].

The results of theoretical studies presented in [5] showed that the spatial distribution of the stops in the city is close to bivariate Normal. This regularitycauses

Exponential distribution of the PT link lengths. Considering that any distance between pair of stops can be represented as a set (sequence) of the links, it was proved that the distribution of could be described by Gamma distribution:

$$f(l_{ij}) = \frac{\beta^{\alpha} l_{ij}^{\alpha-1} e^{-\beta l_{ij}}}{\Gamma(\alpha)}, (l_{ij} > 0, \qquad \beta > 0, \qquad \alpha > 0) \tag{2}$$

where:

β – a scale parameter;

α – a shape parameter.

Also it was found that the Gamma distribution of distances between pairs of stops is not transformed into another distribution during formationof a trip distance set basedon a trip matrix according to the mechanism shown in Fig. 1.

It allows assuming that the Gamma distribution density function (2) isthe TLFD model in the case if TLFD is generated by the trip matrix where diagonal cells are equal to 0 and all other cells are equal to 1 (intra-zone trips arenot considered in this study):

$$H = \{h_{ij}\} = \begin{pmatrix} 0\,1\,1 \cdots 1 \\ 1\,0\,1 \cdots 1 \\ 1\,1\,0 \cdots 1 \\ \dots\dots\dots \cdots \dots \\ 1\,1\,1 \cdots 0 \end{pmatrix} \tag{3}$$

The given trip matrixcan only be theoretical and it is not applied in practice, so the slope of thecurve of a TLFD model represented by formula (2) may vary depending on thestrategy of trip matrix filling.

The transformation of one random variable (distances between pairs of stops) into a different one (trip distances) shown in Fig. 1 can be explained as follows: empirical probability of making the trip over a certain distance according to the trip matrix (3) can be defined with thefollowing formula:

$$P_{emp}(l_{ij} \in \Delta) = \frac{1}{M} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_A} I_{\{l_{ij} \in \Delta\}}, \qquad M = N_A(N_A - 1), \tag{4}$$

where:

$I_{\{l_{ij} \in \Delta\}}$ – indicator of event that l_{ij} trip distancebelongs to a certain Δ interval.

The theoretical probability of theaforementioned event can be determined by using the formula (2). In general, the trip quantityinmatrix cells represents the frequency of certain distancetrips. In terms of statistical probability and exhaustive events the probability of a certain number of trips between zones can be written as follows:

$$p(h) = \frac{h_{ij}}{H}, \qquad H = \sum_{\substack{i,j=1 \\ i \neq j}}^{N_A} h_{ij} \tag{5}$$

where:
H – a total number of trips in the matrix (during the analysis period).

From expressions (4) and (5) the empirical probability of making the h_{ij} tripsovera distance l_{ij} can be defined as:

$$P_{emp}(l_{ij} \in \Delta) = \sum_{\substack{i,j=1 \\ i \neq j}}^{N_A} \frac{h_{ij}}{H} I_{\{l_{ij} \in \Delta\}} \tag{6}$$

Formula (6) is the probabilistic mixture of indicators. It should be mentioned that if there are no trips over a distance , then probability (6) is equal to 0. Since the distance can be described by Gamma distribution, which is stable to be summarized up, the next hypothesis can be given: the theoretical probability of making the trips overa distance , which is defined by the TLFD, can also be determined by the Gamma distribution density function.

It provides a good basis for determining the possible states of trip matrix based on TLFD model. In this case the most appropriate is to use the interval conception of transport demand modelling. To calculate the desired matrices it is necessary to use the factual TLFD patterns that can be described by Gamma distribution curve according to the research results.

The availability of parameters of distribution of distances between zones and parameters of TLFD model (whichis useful for describing the factual TLFD) allows determining trip matrices reproducingits pattern. The determination of the trip matrix states should be based on finding the numbers of trips made over a certain distancefrom the matrix of distances between transport zones.These numbers of trips between zones, which are, in fact, the frequencies of trip making over the certain distances, shouldreproducethe factual TLFD in total. Thus, availability of the parameters of distribution of distances between zones and parameters of TLFD model allows definingthe ratio of certain trip length frequencies and, thereafter, the number of trips between zones that transform the distribution of distances between zones into TLFD, Fig. 2.

Hence, it is possible to establish the order of trip matrix filling. The TLFD model allows determining the percentage of population making work trips over the distancesthat belong to the fixed intervalsof trip length distribution.

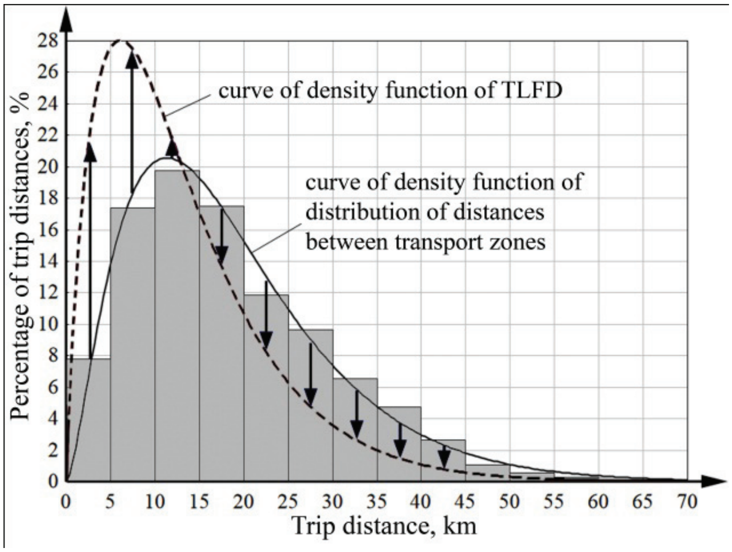

Fig. 2. Transformation of distribution of distances between zones into TLFD

The percentage of residents whose trip distances should belong to a certain interval η_I can be calculated as the subtraction of the values of TLFD model represented by Gamma distribution function at the points that correspond to the interval boundaries:

$$P\left\{l_{ij}^{(\eta_I)}\right\} = P\{l_{ij} \in \eta_I\} = P\{l_{ij} \in (\eta_I^{low}; \eta_I^{up}]\} = W(\eta_I^{up}) - W(\eta_I^{low}) \quad (7)$$

where:

$P\left\{l_{ij}^{(\eta_I)}\right\} = P\{l_{ij} \in \eta_I\}$ – probability of belonging the work trip distance l_{ij} to theinterval η_I;

$\eta_I = (\eta_I^{low}; \eta_I^{up}]$ – an interval of grouping of passenger trip distances,whichis defined by lower η_I^{low} and upper η_I^{up} boundaries,

$I=1,2,\dots,x_{Int}$, x_{Int} – number of group intervals in the TLFD;

$W(\eta_I^{low})$,$W(\eta_I^{up})$ – the values of TLFD model at points η_I^{low} and η_I^{up}, respectively.

This allows calculating the total number of trips that should be made overthedistances from a certain interval η_I. Then this number (frequencies) of trip distances in each interval η_I makes up the factual TLFD (described by TLFD model):

$$H^{(\eta_I)} = P\left\{l_{ij}^{(\eta_I)}\right\} \cdot H \quad (8)$$

where:

$H^{(\eta_I)}$– total number of trips over the distances that belong to the interval η_I.

The grasp ofTLFD and theapply of the matrix of distances between transport zones make it possible to find trip matrix cells that contain the trips $h_{ij}^{(\eta_I)}$ made over the appropriate distances $l_{ij}^{(\eta_I)}$ that belong to the interval $\eta_I = (\eta_I^{low}; \eta_I^{up}]$. It allows toset up the equations defining the total number of trips, which are made over the distances that belong to the certaingroup intervals. Thus, in order to make the trip distribution it is necessary to solve the following system:

$$\begin{cases} h_{ij} \geq 0; \\ \sum_{j=1}^{N_A} h_{ij} = D_i; \\ \sum_{i=1}^{N_A} h_{ij} = A_j; \\ \sum_{\substack{i,j \\ i \neq j}} h_{ij}^{(\eta_1)} = H^{(\eta_1)}; \\ \sum_{\substack{i,j \\ i \neq j}} h_{ij}^{(\eta_2)} = H^{(\eta_2)}; \\ \dots\dots\dots\dots\dots\dots\dots \\ \sum_{\substack{i,j \\ i \neq j}} h_{ij}^{(\eta_I)} = H^{(\eta_I)}; \\ \sum_{i=1}^{N_A} \sum_{j=1}^{N_A} h_{ij} = \sum_{I=1}^{x_{Int}} H^{(\eta_I)} = H, \end{cases} \tag{9}$$

where:

$h_{ij}^{(\eta_I)}$ – number of trips among the h_{ij} which are made over the distances $l_{ij}^{(\eta_I)}$ thatbelong to the interval η_I.

To receive trip matrices meeting the constraints (9), the VBA (Visual Basic for Applications) program in the MS Excel software has been developed. The program run during the filling of the trip matrix cells is based on the random numbers and, therefore, leads to receiving the random matrices meeting the factual TLFD.

The proposed method of trip matrix calculation allowstaking into account the probabilistic nature of the home-based work tripsand receivingsuch a distribution of trips between matrix cells thatprovides formation of the factual TLFD. Also, the developed method corresponds to interval conceptionof transport demand modellingand promotes it. The trip matrices, obtained by using the developed method, make possible the interval estimate of any parameters associated with transport demand realization and, properly, the quality of transport service.

4. Practical implementation of developed trip matrix calculation method

To implement the developed method of transport demand modelling the cities Sverdlovsk, Kupiansk and Balakliia (Kharkiv region, Ukraine) were selected. They belong to the category of small cities (this category includes cities with population up to 100 thousand people) where public transport service is available.

To perform calculations withthe developed method the following initial data are required:1)the total number of trips generated and attracted by transport zones; 2) the matrix of distances between transport zones; 3) the TLFD model.

The total number of trips generated and attracted by transport zones was obtained due to the work [4]. The matrix of distances between transport zones was calculated with PTV Vision VISUM software. Moreover,the trip matrices were found in paper [4].These matrices can be considered as the factual ones with rather high probability,though refer to the survey period. The matrices allowed formingtheTLFDsby the scheme shown in Fig. 1.

To describe the received TLFDs the Gamma distribution was appropriate, Fig. 3. The parameters of TLFD model (Gamma distribution) are presented in Tab. 2.

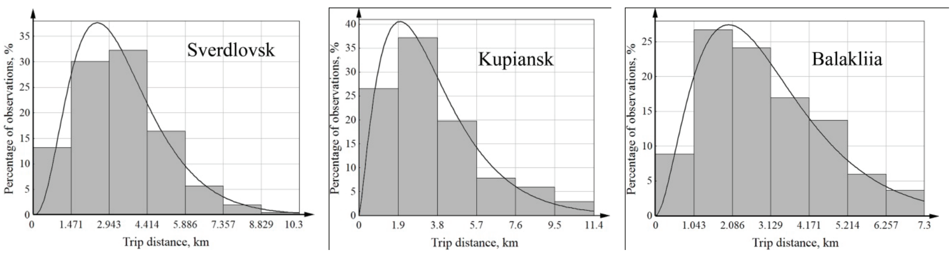

Fig. 3. Trip length frequency distributions of Ukrainian cities

Tab. 2. Parameters of the TLFD models (Gamma distribution) for the Ukrainian cities

City	Parameters		Mathematical expectancy	Kolmogorov-Smirnov test	Chi-square test	Probability of chi-square test
	Scale	Shape				
Sverdlovsk	0.92	3.70	3.403	0.025	6.449	0.257
Kupiansk	1.55	2.28	3.517	0.016	2.711	0.380
Balakliia	1.06	2.87	2.985	0.024	11.263	0.243

Using the fixed boundaries of TLFD group intervals and itsparameters both the probability and numberof trip distances that belong to each interval were calculated by formulas (7) and (8) respectively, Tab. 3.

Tab. 3. Percentage ratio between number of work trips overthe certain distances in Ukrainian cities

Interval No	Sverdlovsk		Kupiansk		Balakliia	
	Probability of trip making	Number of trips	Probability of trip making	Number of trips	Probability of trip making	Number of trips
1	0.109	567	0.269	204	0.092	95
2	0.352	1837	0.362	275	0.253	263
3	0.294	1535	0.211	160	0.252	261
4	0.152	790	0.096	73	0.179	185
5	0.061	320	0.039	30	0.108	112
6	0.022	112	0.023	17	0.059	61
7	0.010	51	-	-	0.057	60

This data hasbecome the initial information for the formation of system (9). The use of program developed in MS Excel allowed generating 10 trip matrices, which meet constraints (9),for every city. The complete compliance with factual TLFDs can be considered as the confirmation of the possibility of using trip matrices in practice.

To estimate the efficiency of public transport functioning in Sverdlovsk, Kupianskand Balakliia the average in-vehicle distance and time of a travel were selected. These indices were calculated withVisum using each state of trip matrixthat allowed determiningthe interval of their possible values, i.e. to give the interval estimate of passengers' travel parameters,Tab. 4.

Tab. 4. The average indices of passengers' service quality in Ukrainian cities

City	Interval of possible values of average	
	in-vehicle trip distance, km	in-vehicle trip time, min
Sverdlovsk	[3.138;3.453]	[17.0;18.7]
Kupiansk	[2.508;2.719]	[8.2;8.9]
Balakliia	[2.192;2.411]	[8.1;8.9]

The obtained results suggest that the developed method of determining the demand for public transport services is practicable. It allows determining the possiblestates of transport demand in the cities and significant reduction of the interval and the number of possible variants of trip matrixaccordingly.That is a major contribution inapplying the interval conception of modellingthe demand for intra--city public transport trips.

5. Conclusions

When modelling the transport demand in cities the intra-city trip length distribution curve should be taken into account, because the regularity described by this curveis directly related to the demand for transport services and itallows taking the probabilistic nature of populationtravelling into account.

The developed method of transport demand definition based on TLFD model allows improving the interval conception of trip matrix calculation by determining the most probable demand states according to the factual TLFD pattern. That makes the interval estimates of passengers' trips parameters possible.

At the same time, the application of the factual TLFD while determining the trip matrix is a serious base for usingthe transportation factors in modellingthe demand for PT servicesas different to common analogies with natural processes.

References

[1] Bell M.G.H. *et al*. Transport Planning and Traffic Engineering. C.A. O'Flaherty, Ed. Oxford: Butterworth-Heinemann, 1997.

[2] Ben-Akiva M.,Bierlaire M., Discrete Choice Models with Applications to Departure Time and Route Choice. R. Hall, Ed., 2nd ed. Kluwer: Handbook of Transportation Science, 2003.

[3] Horbachev P., Kontseptsiiaformuvannia system marshrutnohopasazhyrskohotransportu v mistakh [Conception of the Forming of Public Transport Systems in Cities], Sc.D. dissertation, Dept. Transpotation Technologies, Kharkiv National Automobile and Highway Univ., Kharkiv, 2009(in Ukrainian).

[4] Horbachov P.,LiubyiYev., Otsinkatochnostiisnuiuchykhmetodivmodeliuvanniapasazhyrskykhkorespondentsiinaprykladimalohomista ["The Estimation of Accuracy of Existing Methods of Trip Matrix Modelling by Example of Small City"], Informatsiino-keruiuchysystemynazaliznychnomutransporti – Information-Management Systems on Railway Transport, no. 5-6, pp. 48-52, 2010(in Ukrainian).

[5] Horbachov P., Makarychev O., Svichynskyi S., Obgruntuvannia pokaznykov ohozakonurozpodiludovzhynperehonivmiskoho pasazhyrskoho transportu ["Theoretical Justification of Exponential Distribution Law of Distances Between Stops of City Public Transport"].Avtomobilnyj transport – Automobile Transport, no. 31, pp. 88-92, 2012(in Ukrainian).

[6] Ortuzar J.D.,Willumsen L.G., Modelling Transport. 4th ed. Chichester: Wiley, 2011.

[7] Wilson A.G., Entropy in Urban and Regional Modelling. London: Pion Limited, 1970.