COMMUNICATIONS

Jan Celko - Marek Drliciak - Igor Ripka*

DATA STRUCTURE OF REGIONAL TRANSPORT FORECAST MODEL

There are several ways to define basic transport relations in an studied area. Generally, the values have to represent relevant and real situation. The quality of data sources is the determining factor of transport forecast quality. The basic purpose of transportation planning and management is to match transportation supply with travel demand, which represents 'need'. A thorough understanding of existing travel pattern is necessary for identifying and analysing existing traffic related problems. Detailed data on current travel pattern and traffic volumes are needed also for developing travel forecasting/prediction models. The prediction of future travel demands is an essential task of the long-range transportation planning process for determining strategies for accommodating future needs.

The transport forecast process estimates the number of vehicles or people that will use a specific transportation facility in the future. The article describes gathering and analysing process of available data for Slovak regions. The mathematical model of forecast transport relations needs a unique form of input data and special computing parameters. The aim of paper is to describe the implementation of the existing data to the transport forecast calculation.

The output data, which are evaluated from the forecast transport model, have to describe a real vision of the region development. The sets of function parameters (measurable or unmeasurable) are edited by using different results of surveys that are often complicated and very time-consuming.

Keywords: Transport forecast, 4-step model, modelling.

1. Introduction

Solutions of traffic problems in the town, which are formatted by the influence of strong inter-zone relations and other development intentions, are necessary to be carried out complexly in a general context in order to ensure sustainable regional development. This can be ensured also by creation of a traffic model. The correctly defined traffic model allows evaluating new development activities impact on the road infrastructure loading and allows evaluating alternative solutions of increased traffic problems, too. We can monitor, in addition, the impact of individual phases of highways and highway feeder construction on the urban network loading. The transport analysis and transport prediction models are generally solved by using the Four-step model that was for the first time implemented in Detroit at the Metropolitan Area Traffic Study in 1950. In spite of the time span, the model has still been unchallenged [1].

2. Four-step model

Trip Generation is the first step in the conventional four-step transportation forecasting process, followed by Trip distribution, Modal Split (Mode Choice), and Route Assignment (Fig. 1). Trip generation predicts the number of trips originating in a particular traffic analysis zone. For the residential side of things, trip generation is thought of as a function of the social and economic attributes of households. Households and housing units are very similar measures but sometimes housing units have no households, and sometimes they contain multiple households. Clearly, housing units are easier to measure, and thus are often used in model. Clear definition of the used parameter is very important for an exact model creation. The second step of four-step model - Trip distribution (or destination choice or zonal interchange analysis) - matches trip makers' origins and destinations to develop a "trip table". The trip table is described by the O-D Matrix (Origin – Destination matrix) that express the number of trips between start and final zones. Historically,

* ¹Jan Celko, ¹Marek Drliciak, ²Igor Ripka
¹Department of Highway Engineering, Faculty of Civil Engineering, University of Zilina, Slovakia
²Institute of Transportation, Technical University of Vienna, Austria
E-mail: jan.celko@fstav.uniza.sk

trip distribution has been the least developed component of the transportation planning model [2].

The trip distribution's zonal interchange analysis yields a set of origin - destination matrices that inform where the trips will be made. The mode choice (Modal split) is a third step of the model, which determines the transport mode using [3].

To determine facility needs, costs, and benefits, the number of trips on each route and link of the network is necessary to determine. The route like a simple chain of links between origin and destination zones is assigned to the network of highways and transit systems. The first required information is the present pattern of travel times and flows, followed by determination of the activity [2].
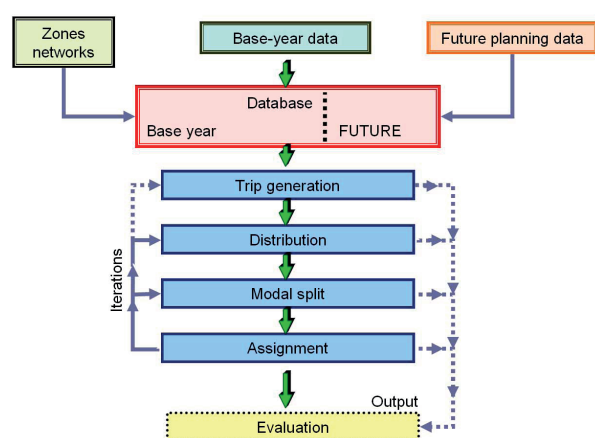

Fig. 1 Four-step model [4]

The solved region is divided into zones dependent on selected parameters which are defined by a detailed analysis of urban, transport, economic and sociological conditions. The basic requisite parameter for traffic forecast is generation of the trips by origin zones [4]. The number of trips depends above all on the zone character. The greatest source of the trips is a living zone. The majority of the trips are connected with a household. The home based trips mean that one end of the trip is connected with home (home – work, sport - home, etc.). On the other hand, the industrial zones are the most attractive for trips induced by a destination zone.

3. Data gathering

The gathering of social and traffic data is the essential step in the process of traffic relation calculations in the studied area. The choice and quantity of questioned samples of population indicates the quality and reliability of results of the transport demand model. The quantity of data collected will be a function of the number of respondents in the final dataset and the amount of information obtained from each respondent. This, in itself,

presents a trade-off situation because any attempt to collect more information from each respondent (beyond a threshold level of information) may result in fewer respondents responding. The total number of respondents will obviously depend on the size of the sample drawn from the population and the response rate obtained from that sample.

3.1. Input data

There are several ways to define basic transport relations in the studied area. Generally, the values must represent relevant and real situation. The quality of data sources of transport behaviours of inhabitants is the determining factor of transport forecast quality. It is possible to use several sources to fill the database in Slovakia [3]. The first useful source is the Statistical Office of the Slovak Republic (SO SR).

The system of demographic statistics in Slovakia based on monthly processing of data on demographic events drawn from the exhaustive population survey, regularly performed censuses of population and housing and supplementary population surveys (microcensuses); provides the decisive bulk of information on population as a whole, its spatial distribution, number, structures, and characteristics in the specific period of time.

Labour Market Statistics provide information on present situation and trends in employment, unemployment, job vacancies, wages, labour costs and strikes. Primary source of this information consists of regular statistical surveys carried out in enterprises and households. Labour Force Survey (LFS) quarterly monitors labour force supply focusing on economic activity of population, employment and unemployment.

Social statistics aims at obtaining relevant and comparable statistical information on social protection, income and living conditions of households, labour and wages, education and learning, health, culture and criminality. This information is needed for making and monitoring policy on each level of public administration management including international organisations and for meeting users´ requirements on national and international level.

The system of quarterly and annual enterprise surveys represents the key instrument of data collection in the labour and wages statistics. It provides information on the number of employees, hours worked, job vacancies, average nominal and real wage (classified by branches, size of enterprises, regions) for operative measures of central bodies as well as for experts and for general public. Information on a level of average monthly nominal wage of employees published by the SO SR is a basis for calculation of other significant social indicators (e.g. basis of pension assessment, earnings of institutional representatives, limit of deposit protection etc.). For the purpose of a regional overview, the employment and wages have been surveyed in annual periodicity referring to work place. Furthermore, a total

3.2. Data gathering process

The first step in data gathering process is a sharing the sophistic data system DATAcube. Database DATAcube contains multidimensional tables for indicators of economic and socio-economic development. The classification system of individual tables is based on maintaining the structure of domains and fields similarly as in the web Portal. Data from various statistical fields are presented in the form of multidimensional tables in monthly, quarterly or yearly time series and allow to create specific own selections. The outputs can be exported to the select format.

Despite of data foundation the demand characteristics for a demand model are not possible to obtain from available sources. The mobility information is to be determined by a classic questionnaire system, where inhabitants are requested for daily transport activities with determination of number of trips, modal choice, traffic chains, and range of the trips.

The next required data contain information about infrastructure for the network model. The source of data is available in the sources of the Road databank (RDB) that is a branch of the Slovak Road Administration. Through its own facilities, in close cooperation with road owners and administrators, it executes:
• collection and processing of technical data of roads,
• pavement diagnostics,
• development and building of the GIS (Geographical Information System),
• information service of databank.

In addition, RDB operates the application system for assessment and determination of routes for the transport of excessive and oversized cargo. The RDB also offers the summary data of the road network SR, selected data of objects in the road network, length of the roads since 1995 and the basic and thematic maps of the road network (road sections with toll, international road E, TEM, TEN-T), maps of regions, districts, map sheets and maps of intersections. Other urban streets are not included in RDB

4. Parameter setting

The transport demand model is based on a large base of evaluated data which are processed into the mathematical function. The most popular functions are: BoxCox, Logit, Kirchhoff function. Mentioned functions are used in trip distribution, mode choice and assignment process (Fig. 2).

Fig. 2 Definition of Function type [5]

Table 1 presents the example of data evaluation from mobility survey and application to the demand model parameters. Restraint functions were used for trips distribution. The parameters adjust the shape of the logit function.

The example of restraint function for utility travel time [6]   Table 1

Group	town	work	school	shop	private
Student	<5000		0.056	-0.0756	-0.0619
	5000-50000			-0.086	-0.075
	>50000			-0.106	-0.075
Employed with car	<5000	-0.059		-0.070	-0.044
	5000-50000	-0.062		-0.087	-0.049
	>50000	-0.072		-0.082	-0.046

4.1. Volume delay function

The key element of the road assignment process using equilibrium method is the volume-delay function (VDF). There are several different formulas for description of VDF (from a simple linear with the breakpoint to much more complicated). In most traffic assignment methods, the effect of road capacity on travel times is specified by means of volume-delay functions t(v) which expressed the travel time (or cost) on a road link as a function of the traffic volume (V). Usually these functions are expressed as the product of the free flow time multiplied by a normalised congestion function.

$$t_v = t_o * f\left(\frac{v}{q}\right), \tag{1}$$

where
t_o - free flow time,
V - traffic volume,
q - capacity (it may be full capacity, capacity on the level or saturation flow).

The function should have several conditions. For the formulation of VDF the conditions can be divided into two groups – mathematical and behavioural [7].

Mathematical
From the mathematical point of view, having in mind the system optimal principle, the function should be [8]:
• continuous,

- strictly increasing,
- non – negative.

Behavioural

Following facts should be kept in mind [7]:

- in urban traffic most of the trips are made every day – users have a lot of experience with traffic conditions in different times of the day on different routes,
- the user can make also "on-time" decisions – if he sees there is a congestion on the street in front of him or when he hears about congestion on the radio, he can turn and choose another route,
- time spent in the congestion weighs much more for the user than travel time at an acceptable speed – the user is strongly forced to choose another route,
- free flow speed - defined as a speed of the traffic flow without any disturbances on the road; free flow speed decreases after implementing the traffic lights, but many users choose a route without acceptance this decrease.
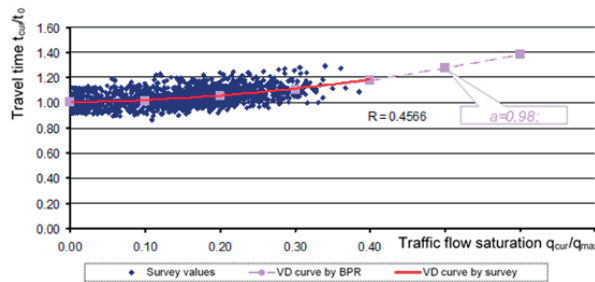

Fig. 3 Definition Volume delay function [9]

Figure 3 presents the example of the parameters of VDF function for for urban main street. Parameters were derived from large measurements in Zilina. The parameter „a" of exponential function was derived from mentioned measurements.

4.2. The impedance

One of the most difficult steps - mode choice - calculates a discrete distribution model. The trip of each route is calculated for a person group first, and one from transport modes is chosen. The modes are divided into two groups: exchangeable modes (generally walk, passenger and public transport), and non-exchangeable modes (car, bike). Changing the mode during the same trip creates a trip chain. If the first mode is a non-exchangeable mode, the entire trip chain is maintained independent of its attributes. If an exchangeable mode was selected for the first trip, mode choice is carried out for the remaining chain trips, however, only within the exchangeable modes.

$$P_{gij}(m) = \frac{e^{U_{gij}(m)}}{\sum_{k=1}^{M} e^{U_{gij}(k)}} \qquad (2)$$

where

i,j indices of origin and destination zones,
$P_{gij}(m)$ choice probability for mode m by person group g,
$U_{gij}(m)$ objective utility value of mode m for person group g,
m number of alternative modes.

The structure of the utility function is of special importance in this respect. The following utility function from [8] illustrates possible components of the function for the mode choice:

$$U_{gij}(m) = p_{1gm} * T_{ijm} + p_{2gm} * Z_{ijm} + p_{3gm} * \ln\frac{D_{ij}}{p_{4gm}} + \\ + p_{5gm} * C_{ijm} + p_{6gm} \qquad (3)$$

with the parameters per group g and mode m:

p_{1gm} marginal utility of one-minute ride time,
p_{2gm} marginal utility of one-minute access/egress time,
p_{3gm} marginal utility of logarithmic relative distance increases (impact of distance advantage),
p_{4gm} advantage distance of mode m,
p_{5gm} marginal utility of one monetary unit of a ticket price,
p_{6gm} constant utility of mode m,

and the mode dependent attributes:

T_{ijm} ride time from i to j with mode m,
Z_{ijm} sum of access time at i and egress time at j for mode m,
C_{ijm} costs of trip from i to j with mode m,
D_{ij} distance from i to j.

The parameters p1, p2 and p5 have a negative utility, so they have a negative sign. The parameter p4 (advantage distance) indicates from what distance a particular mode of transport is considered to be useful by travellers of a specific group. Only when the distance d_{ij} exceeds the advantage distance p4 (d_{ij} > p4) the quotient d_{ij} / p4 is greater than 1 and the logarithmical term ln (d_{ij} / p4) turns positive. As a consequence, distances below the advantage distance result in a negative utility. The distance relevant parameters p3 and p4 are important to give preference to the modes walking and cycling on short distances, as their p4 is defined smaller than for the modes of transport car und PrT and their p3 is negative compared to the others.
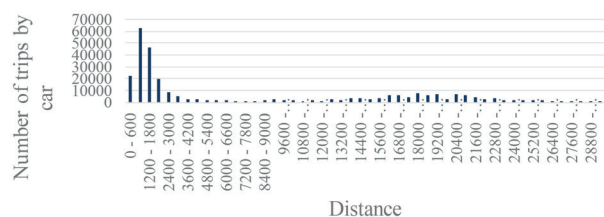

Fig. 4 The final graphical expression of trips by car according to the distance. (Source: Authors)

The parameters of utility function were calibrated by results from mobility survey. The final modal split was evaluated according to many parameters. Figure 4 presents the evaluation according to the length of trips, Fig. 5 presents the observed modal split by the groups of inhabitants defined for Zilina.
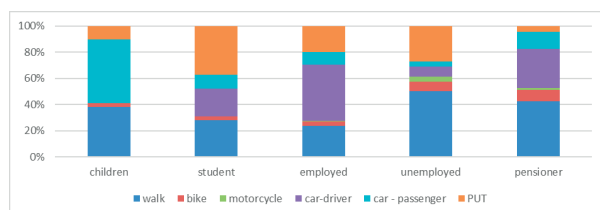

Fig. 5 Modal split in Zilina (Source: Authors)

5. Conclusion

The structure of database of the transport forecast model, especially multimodal, is set by the experience of a model maker. Every step of transport forecast model offers several ways for calculation with different structure of parameters. Therefore, the very large database of the mobility data is necessary for the model solution. The article does not provide detailed information from realised mobility surveys. Authors present only several results of data evaluation from the mobility survey in Zilina.

References

[1] WEINER, E.: *Urban Transportation Planning in the United States: An Historical Overview*, p. 26, ISBN 0-275-96329-2

[2] LEVINSON, D., LIU, H., GARRISON, W., DANCZYK, A., CORBETT, M.: *Fundamentals of Transportation/Mode Choice/Solution,* https://upload.wikimedia.org/ wikipedia/commons/7/79/Fundamentals_of_Transportation.pdf.

[3] CELKO, J., DRLICIAK, M., GAVULOVA, A.: Transportation Planning Model. *Communications - Scientific Letters of the University of Zilina*, vol. 9, No. 3, 2007, pp. 28-32. ISSN 1335-4205.

[4] ORTUZAR, J., D., WILLUMSEN, L. G.: *Modelling Transport*, Wiley, p. 23, 2001, ISBN 978-0-471-86110-2 (H/B).

[5] PTV Visum 15, User manual.

[6] RIPKA, I.: *Determination of the Resistance Functions at Prediction of the Inter-Zonal Relations* (in Slovak). Doctoral thesis, University of Zilina, 2011.

[7] JASTRZEBSKU, W. P.: *Volume Delay Functions*, Proc. of 15th Intern. EMME/2 User's Group Conference, Vancouver, B. C., October 2000.

[8] DRLICIAK, M., CELKO, J.: *Social-traffic Data for Mezoscopic Traffic Model*, Proc. of 12th Intern. Conference Reliability and Statistics in Transportation and Communication (RelStat'12), October 2012, Riga, pp. 204-210. ISBN 978-9984-818-49-8.

[9] BARTOVIC, M.: *Models for Division of Transport Work* (in Slovak), Doctoral thesis, University of Zilina, 2009.