Quality assessment of distributions in transport modelling

Dipl.-Ing. Eric Pestel

eric.pestel@isv.uni-stuttgart.de

University of Stuttgart
Institute for Road and Transport Science
Department for Transport Planning and Traffic Engineering
Pfaffenwaldring 7, 70569 Stuttgart, Germany

Key words: Quality assessment, macroscopic travel demand modelling, trip distance distributions, trip time distributions

1 Motivation

In macroscopic travel demand modelling, distributions of the time travelled or the distance travelled are used to evaluate how good a model fits in with observed sample data. Therefore, the comparison of distributions is an essential part in the model validation process. Despite its importance the common modelling guidelines from UK (WebTAG [1]), USA (Travel Model Validation and Reasonableness Checking Manual [2]) or Austria (draft of Qualivermo [3]) furnish little information about the correct structure and handling of such distributions. The following questions inevitably appear when validating a model with distributions of time and distance:

- What indicator should be used for the classification?
- Which area is covered by the distribution?
- What is the reference for a comparison of distributions?
- How many classes should be distinguished? What is the appropriate size of each class? Should the class size increase with the distance travelled? How to proceed with empty classes?
- What quality measure should be used when comparing two distributions? Should absolute or relative frequencies or both be evaluated?

The lack of rules leads to individual solutions, which complicate a model validation and the comparison of models. For example, when comparing two distributions the quality measure strongly depends on the number of classes. Therefore, guidelines for model validation need to suggest an appropriate way to determine the number of classes.

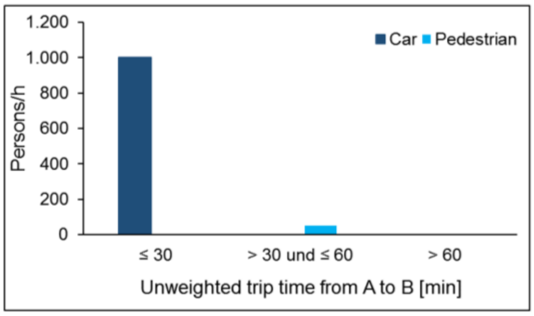
The paper addresses the questions above in order to provide a general guidance on how to handle trip distance distributions and trip time distributions within the model validation process.

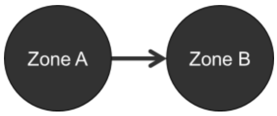
2 Selection of indicators for a classification

In a trip distance and a trip time distribution, the trips are assigned to discrete classes. This requires trip volumes, distance and time values on the level of origin-destination-pairs (OD-pairs). If different modes are considered, indicator values for distance and time may vary for each mode. This can lead to cases where the trips of one OD-pair with mode-specific indicator values fall in different classes. Figure 1 (left) shows a case, where the demand for car and pedestrian trips is assigned to different classes although they belong to the same OD-pair. Such cases occur frequently for trip time, but they can also occur for trip distance, as public transport trips are often longer than car trips.

For comparing the distance-dependent modal share, the trips of one OD-pair should fall into the same class. To achieve this, it is necessary to use a mode-independent indicator for the classification. In case of trip distance distributions, the direct distance between the two centroids can serve as a "natural" mode-independent indicator. For trip time distributions, such a "natural" mode-independent indicator does not exist. An indicator weighted with the number of trips provides a solution for this. Figure 1 (right) extends the example and determines a mean weighted trip time. As a result, both car and pedestrian trips of one OD-pair now fall in the same class.

OD-Pair A to B:	**Mode Car**	**Mode Pedestrian**
Distance	5 km	
Speed	50 km/h	5 km/h
Trip time	0.1 h = 6 min	1 h = 60 min
Demand	1,000 Persons/h	50 Persons/h





OD-Pair A to B:	**Mode Car**	**Mode Pedestrian**
Distance	5 km	
Speed	50 km/h	5 km/h
Trip time	0.1 h = 6 min	1 h = 60 min
Demand	1,000 Persons/h	50 Persons/h
Mean weighted trip time	$\frac{0.1\,[h]\cdot 1{,}000\left[\frac{Persons}{h}\right]+1\,[h]\cdot 50\left[\frac{Persons}{h}\right]}{1{,}000\left[\frac{Persons}{h}\right]+50\left[\frac{Persons}{h}\right]}$ = 8.6 min	

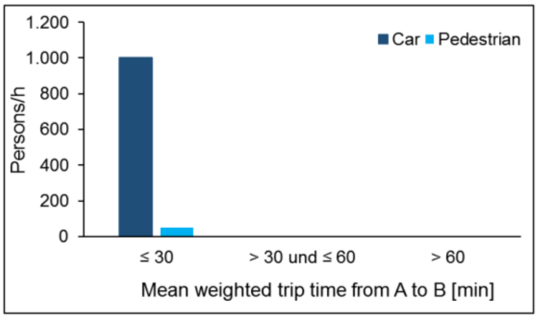


Figure 1: Classification with a mode-specific indicator (left) and with a mode-independent indicator (right).

A standardized classification requires rules for handling intrazonal trips. In a macroscopic travel demand matrix the main diagonal represents the intrazonal trips. Since this part of the demand is not assigned to the network model, indicators like trip time or trip distance cannot be calculated, but must be estimated in such a way that they represent an average movement within a zone. Therefore, it is advisable to exclude the main diagonal when comparing an observed and a modelled distance or time distribution.

Nevertheless, this does not release the modellers from their obligation to check the intrazonal trips in a separate analysis. Such an analysis should investigate the number and the modal split of intrazonal trips. A separate examination for different zone sizes may also be possible.

3 Selection of the study area and a reference distribution

A travel demand model computes demand for a defined study area. The model is calibrated with observed data from a household travel survey. This survey ideally covers a sample of the population from the entire study area. If the study area of the survey and the study area of the travel demand model do not match completely, the comparison of observed and modelled values should only include trips starting and ending in the common study area. This also means that observed trips leaving the common study area have to be excluded from the survey when comparing the observed and the modelled distributions. A comparison of two distributions is only meaningful if observed and modelled trips relate to the same area.

Calibrating the destination choice of a travel demand model requires observed distance or time distributions from a household travel survey as a reference distribution. Since interviewed persons tend to round estimates of distance and time for their reported trips, it is recommended not to use reported distance and time values directly from the survey. Instead, the distance and time values should be computed with the model.

Besides a household travel survey, the base case of a calibrated travel demand model can also serve as a reference. This is the case for comparisons of modelled scenarios, e.g. when analysing the changes from a matrix estimation or checking the model behaviour in the context of realism or sensitivity tests.

4 Selection of a classification method

4.1 Class width versus class size

Usually distribution classes are determined in such a way that the class width and a number of classes are predefined. Based on the classification indicator the demand is then assigned to the resulting classes. If the class width is the same across all distribution classes (and the number of classes is infinite in an extreme case), the term "equidistant" distribution is used. However, in many cases insufficiently populated classes are aggregated, resulting in classes of varying width.

A different approach defines the quantity of demand in each class and uses it to determine the class width. The resulting classes vary in width but they are equally populated. This method of classifying distributions is called "equiquantile" (based on e.g. PALUŠ [4]).

In the case of an equidistant classification, three questions need to be addressed:

- How large are the classes?
- How many classes should there be in total?
- Should classes be grouped together and, if so, from which class onwards?

On the other hand, the method of equiquantile distribution only raises the question about the number of classes. For this reason, the equiquantile classification method is presented first. Based on this, a procedure is explained to answer the question about the class width of equidistant distributions.

4.2 Analysis with equiquantile classes

The class boundaries of this classification method are calculated using weighted quantiles, with the elements of the indicator matrix representing the classification variable and the demand matrix elements representing the weight.

Although all classes should have the same quantity, the actual demand per class may differ slightly from the desired quantity. This may be due to the following reasons:

- A discrete demand does not necessarily have to be distributed completely and evenly across all classes ($(demand) \bmod (number\ of\ classes) \neq 0$). As a result, at least one class has a greater demand than other classes. The smaller the population, the greater the deviation from theoretically equal classes. However, this effect is reduced, as the sample grows larger.
- An unfortunate weighting, i.e. a very strong demand on a relation, can also lead to a distortion.
- A rounding of class boundaries prior to classification can lead to shifts in demand between neighbouring classes.

As mentioned above, the number of classes must be defined for the equiquantile classification. Dividing the demand into 10% steps, i.e. into ten classes, is one possible pragmatic assumption. Figure 2 shows a sample of a trip distance distribution. It is shown that each class has approximately the same number of trips and the resulting graph matches an approximate linear line.

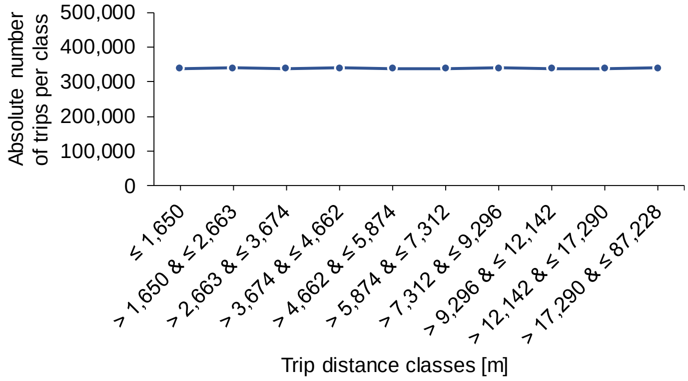


Figure 2: Example of an equiquantile trip distance distribution.

Demand segments are a subset of the total demand. The different modes are often used as a basis for segmenting. Other segmentations are also possible, e.g. based on the trip purpose or based on the trip purpose with a certain mode. For reasons of simplicity, the following sections use the term demand segment synonymous with the term mode.

Depending on the observed quantity of demand on the individual relations, varying class widths will result. For example, the classification in short distance ranges is more detailed for the pedestrian mode than for the mode car. For the comparison of a demand segment in different scenarios, it is advisable to classify the demand segment on a uniform basis for the demand segment in question. Therefore, the classification is determined once for the reference distribution and then preserved for all scenarios.

Figure 3 shows an example of an equiquantile trip distance distribution of the mode car for a realism test, in which the number of inhabitants is changed by ±10% and ±20%. The reference distribution is the base case of a calibrated travel demand model. In the example it can be seen that in the case of less inhabitants (80% or 90%) the demand, i.e. the amount of trips, decreases across all classes, whereas in the lower distance classes there is a tendency to make fewer trips and in the larger distance classes there is a tendency to make more trips. The opposite applies to the case with more inhabitants (110% or 120%).

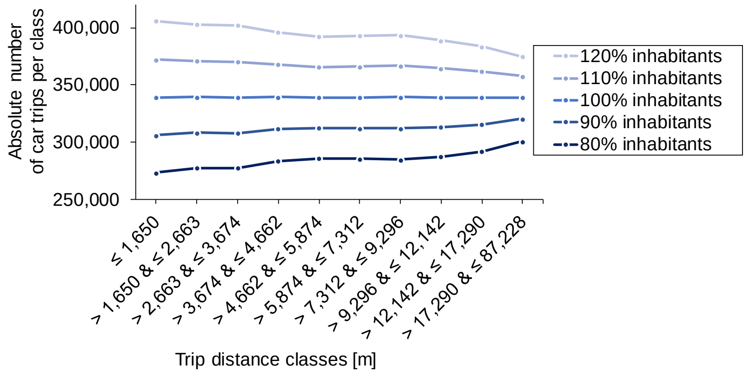
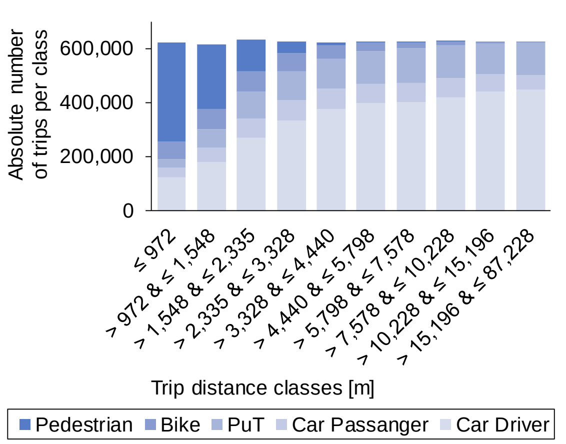


Figure 3: Example of an equiquantile trip distance distribution of mode car with different numbers of inhabitants.

Similar to the evaluation for each demand segment, it is possible to perform an evaluation of the total demand (demand across all modes) to analyse the changes of the total demand in the different distribution classes.

In addition, individual demand segments can also be presented with the classification of the total demand. For example, a distance-dependent modal split (see Figure 4 (left)) or the development of individual modes can be displayed in comparison to the total demand for different scenarios (see Figure 4 (right)).



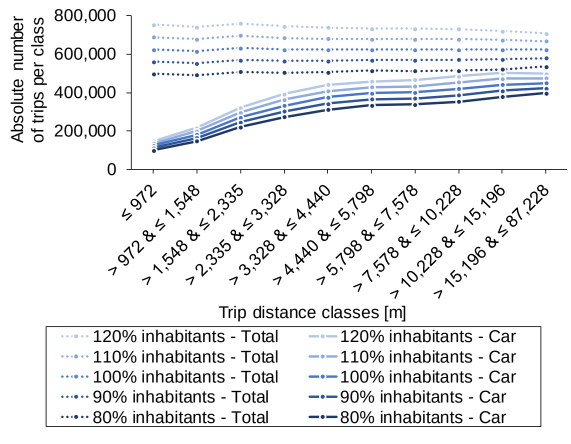


Figure 4: Left: example of an equiquantile trip distance distribution of all modes (distance-dependent modal split),
Right: example of an equiquantile trip distance distribution of mode car and for the total demand with different numbers of inhabitants.

Instead of an approximately equal number of trips per class, it is also possible to use the combination of the number of trips and the trip distance (person distance travelled). This may be useful, as in some cases of the calibration process, it is more important to match the person distance travelled than the actual number of trips. An equal number of trips in all distance classes leads to an increasing person distance travelled with increasing distance class, as is shown in Figure 5 (left). Contrary Figure 5 (right) shows a distribution where each class contains the same amount of person distance travelled, while the number of trips is decreasing monotonously with increasing distance class. In comparison to a "normal" equiquantile distribution (see Figure 5 (left)), it becomes obvious that the class width in the short range distances increases while the long distance classes are displayed in greater detail.

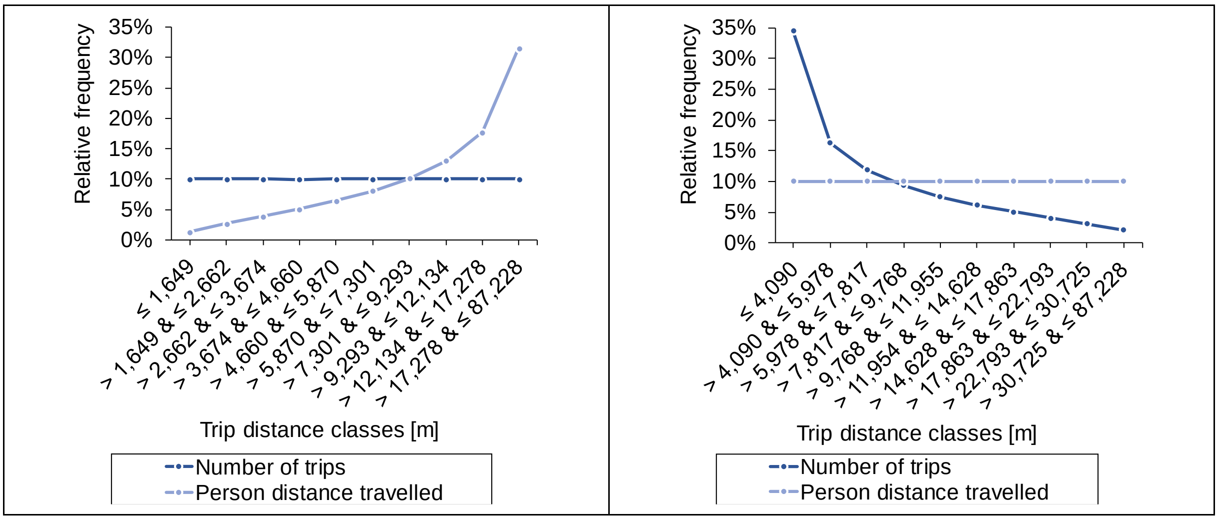


Figure 5: Comparison of an equiquantile distribution according to the number of trips (left) and an equiquantile distribution according to the person distance travelled (right).

4.3 Analysis with equidistant classes

As previously mentioned, the number of classes and the class width must be predefined for the equidistant classification method. A common procedure to determine those parameters is Sturges' rule for constructing histograms, which is discussed e.g. by Hyndman [5].

Another way of determining the class width is to define the smallest class of the equiquantile distributed total demand as the equidistant class width. This class width represents the smallest class with 10% of the total demand. The number of classes is not important for a strictly equidistant distribution, i.e. classes are created until the largest classification indicator is reached. However, in this case empty classes may occur.

The segmentally equidistant classification can be regarded as a variation of the equidistant classification. Here, the statistically unreliable, i.e. low-occupancy, classes are aggregated. However, in order to use this classification method the questions about the number of classes, the size of the grouped classes and the minimum occupancy of a class must be answered once again. QUALIVERMO [3] specifies class widths for these cases:

- for metropolitan traffic in short distance areas: 2 km or 5 min,
- for regional and long-distance traffic: 5 km or 10 min.

5 Selection of a quality measure for distribution comparisons

To check the congruence of two distributions they can be plotted in a common diagram. However, a mere visual examination of congruence is often not adequate. A comparison of the distributions parameters (e.g. mean, standard deviation) which describe their position, appearance and properties may be insufficient, too. Statistical tests (e.g. Kolmogorov-Smirnoff-test) are not practical for their use with macroscopic travel demand models, because due to the large sample size covered by a travel demand model, the test criterion is very strict, which often results in a negative test result (i.e. the distributions do not originate from a common population).

Measures of conformity consider the similarity of two distributions. The following measures are presented and evaluated in the paper:

- Correlation coefficient,
- Euclidean distance,
- Root mean squared error,
- Theil’s Forecasting Accuracy Coefficient,
- Vortisch’s Measure of Distance and
- Coincidence Ratio.

It will be shown that Theil's Forecast Accuracy Coefficient and Vortisch’s Measure of Distance are recommended because of their sophisticated analysis options. The Coincidence Ratio is supported by the relatively user-friendly handling and by specific threshold values from the TRAVEL MODEL VALIDATION AND REASONABLENESS CHECKING MANUAL [2]. Furthermore, its value range from 0 to 1 speaks for a straightforward interpretability.

A comparison of two distributions can either use relative or absolute frequencies. The choice of relative or absolute frequencies influences the value of the measure of conformity. For this reason, a maximum acceptable threshold for considering two distributions as similar requires a clarification, whether to use absolute or relative frequencies. As distance and time distributions derived from surveys are usually provided only as relative values, it is advisable to use relative frequencies. The absolute number of generated trips should be checked directly after trip generation.

6 Conclusion and recommendations

Distributions are an important quality feature of macroscopic travel demand models. In the paper, various classification and quality determination methods have been discussed. The paper concludes with a suggestion of a procedure for creating and evaluating distributions:

- Recommendations for selecting a classification indicator for trip distance distributions and for trip time distributions.
- Recommendations for determining the study area.
- Recommendations for choosing the reference distribution for each application.

- Instructions for quality checks with equiquantile classes and how to proceed with equidistant distributions.

7 Acknowledgements

This paper was developed as part of the following research projects:

- „Influencing factors on the quality of macroscopic travel demand models“ [6] commissioned by the German Research Foundation (DFG).
- „Quality assurance for transport model calculations“ [7] commissioned by the Swiss Association of Transportation Engineers and Experts (SVI) and

Special thanks to PD Dr. Christian Schiller and Robert Simon from the Technical University of Dresden, Dr. Nadine Rieser and Bence Tasnády from EBP Schweiz AG as well as Prof. Dr. Markus Friedrich from the University of Stuttgart. The in-depth discussions within the scope of the above-mentioned research projects have greatly enriched this paper. I also wish to thank Dr. Juliane Pillat (PTV AG) for her thoughts on equiquantile distributions according to the person distance travelled.

8 References

[1] Department for Transport, Ed., *TAG UNIT M3.1: Highway Assignment Modelling*, 2014. Online available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427124/webtag-tag-unit-m3-1-highway-assignment-modelling.pdf. Accessed on: Nov. 03 2015.

[2] Cambridge Systematics Inc., “Travel Model Validation and Reasonableness Checking Manual: Second Edition,” Federal Highway Administration, Cambridge, Massachusetts, 2014. Online available: http://www.camsys.com/pubs/FHWA-HEP-10-042.pdf. Accessed on: Nov. 04 2015.

[3] G. Sammer, G. Röschel, and C. Gruber, *Qualitätssicherung für die Anwendung von Verkehrsnachfragemodellen und Verkehrsprognosen: Quality Management and Model Validation for Application of Transport Demand Modelling and Forecast*. Wien: Österreichische Forschungsgesellschaft Straße, Schiene, Verkehr, 2012.

[4] M. Paluš, “Testing for nonlinearity using redundancies: quantitative and qualitative aspects,” *Physica D*, vol. 1995, no. 80, pp. 186–205, http://www.sciencedirect.com/science/article/pii/0167278995900799, 1995. Online available: http://www.sciencedirect.com/science/article/pii/0167278995900799. Accessed on: Aug. 09 2017.

[5] R. J. Hyndman, “The problem with Sturges’ rule for constructing histograms,” 1995. Online available: https://robjhyndman.com/papers/sturges.pdf. Accessed on: Mar. 02 2018.

[6] M. Friedrich, E. Pestel, M. Schimpf, and C. Schiller, “Einflussgrößen auf die Qualität von makroskopischen Nachfragemodellen im Personenverkehr: DFG-Projekt,” 2015-2019. Online available: http://www.isv.uni-stuttgart.de/vuv/projekte/QTDM/index.html. Accessed on: Dec. 19 2017.

[7] N. Rieser, B. Tasnády, M. Friedrich, and E. Pestel, “Qualitätssicherung von Verkehrsmodellberechnungen,” Forschungsprojekt SVI 2015 / 001 auf Antrag der Schweizerischen Vereinigung der Verkehrsingenieure und Verkehrsexperten, 2018. Online available: http://www.transport-research.info/project/quality-assurance-transport-model-calculations. Accessed on: Aug. 04 2016.