

Validate: A New Method To Generate Nationwide Traffic Data

Thomas Haupt
PTV AG
Stumpfstr.1
76131 Karlsruhe,
Email thomas.haupt@ptv.de

Udo Heidl
PTV AG
Stumpfstr. 1
76131 Karlsruhe
Email udo.heidl@ptv.de

Thomas Friderich
PTV AG
Stumpfstr.1
76131 Karlsruhe,
Email thomas.friderich@ptv.de

Abstract. A nationwide detailed traffic volume calculation has been carried out. Using digital high-resolution navigation networks a trip-end model was used to generate traffic volumes for about 1 million links, which represent the major street network of Germany.

Keywords: road network – traffic volumes – navigation – trip end model – dynamic assignment – travel times

1 Introduction

Validate is a project carried out by PTV AG to build up a nationwide traffic model. *Validate* calculates travel demand (origins and destination of car and HGV trips) from both structural data and activity patterns and distributes the calculated demand to the road network using a capacity restraint assignment. The resulting data indicate the traffic volume per day for the entire major road network of Germany.

In the field of transport planning, procedures of this kind are common practice world-wide and state of the art for strategical planning by governments, counties and municipalities. The new aspect within this project, however, is the large size and high detailing level regarding the road network and demand modeling. Transport models normally cover a specific regional area, e.g. a federal state, or just focus on parts of the road network, e.g. the motorway network. *Validate*, however, is concerned with a nationwide calculation providing a high level of granularity in order to display the entire major road network of all cities and regions. But only the availability of powerful computers and software on the one

hand as well as digital high-resolution road networks and specific – partly commercially collected – data on the other hand make applications like these possible today. This universal approach can be transferred to other European countries. It is also imaginable to carry out similar future projects for several countries simultaneously or for the whole of Europe.

2 The road network

Network modeling is based on the road network. All major roads with traffic-relevant attributes have to be included. The following chapter describes the network structure and the corresponding steps for building the model.

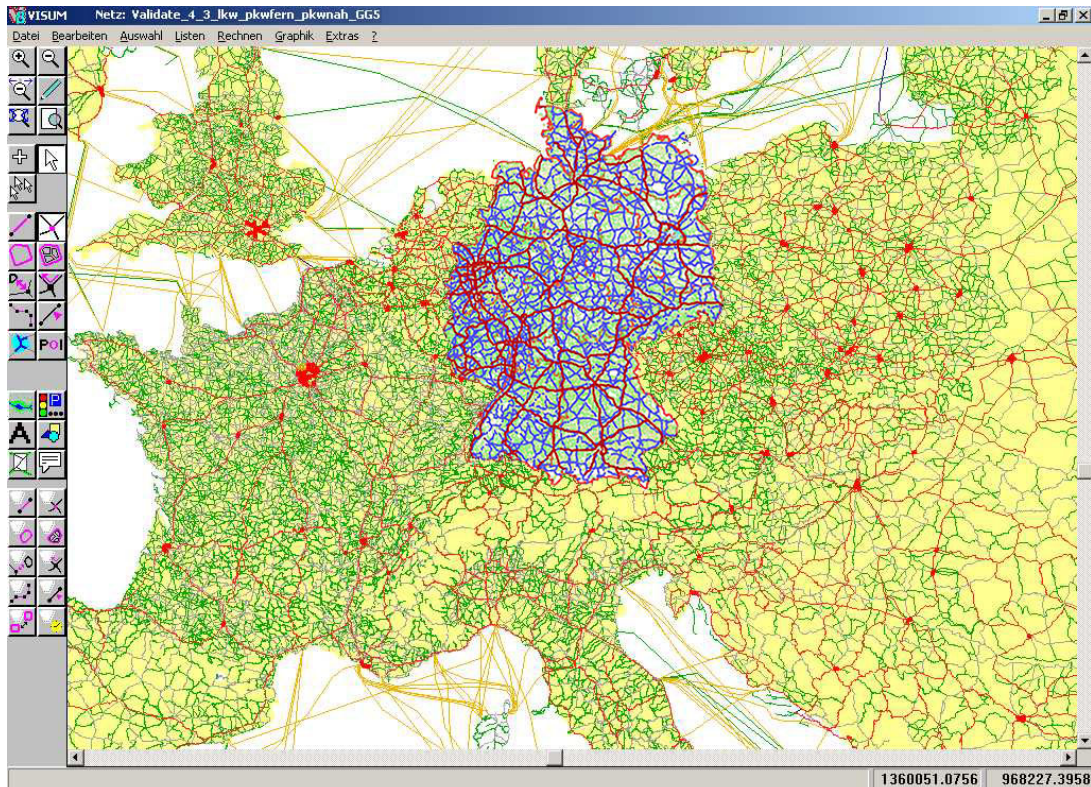


Figure 1: The entire road network of Germany (with Europe)

2.1 Basic data

For a project of this size it is not efficient to carry out one's own data collection. Existing data must be used. The following data sources can be used for the road network:

- Regional road network models
- Digital navigation data

Regional network models are highly detailed and partly up-to-date, however there are no universal quality standards. Moreover, it would be a time-consuming process to acquire the user rights for all data and to manually

integrate the local networks.

The most reliable data among commercial digital maps are those used for in-car navigation systems. Digital navigation data offer several advantages:

- Professional data collection: Guaranteeing a continuous quality standard by collecting road data on a regular basis
- Highly digitized: All streets, even access streets and pedestrian zones, are included. Lanes divided into directions are generally displayed as two links
- Routing: Reliable data collection considering one-way streets and turning regulations
- Current updates: Updates are provided on a quarterly basis. Permanent IDs ensure reference to the last version.

The *Validate* road network has been created on the basis of NAVTEQ's map for Germany covering about 6 million non-directed links.

2.2 Processing

Relevant major roads and trunk roads have been selected from the highly digitized navigation network. Minor side roads and residential streets have not been taken into account. This urban and rural road network able to cover the complete traffic volume except pure residential traffic. The resulting network consists of approx. 3 million links.

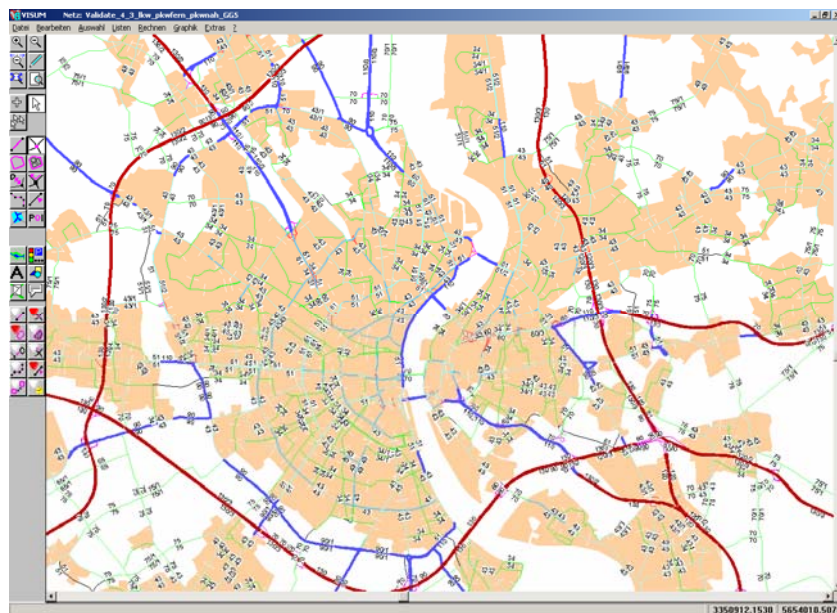


Figure 2: Network Section with speed information (km/h) – urban area of Cologne

Traffic related attributes, especially capacity and free flow speed must be added to perform a capacity dependent route choice procedure. In a first phase, these

data were derived from NAVTEQ's existing attributes describing the road types. In a second phase, this basic supply was refined and verified by matching with other models and data sources.

The so-called generalization is a major step aiming at reducing the number of links within the network without reducing the modeling quality. Links which are connected by two-legged nodes and have identical attributes are aggregated in this procedure. The resolving nationwide network is linked to a reduced European road network to properly display international traffic.

The entire network finally covers about 1 million directed links.

2.3 Updating

The network is based on data from Autumn 2002. Nevertheless, the traffic model has no date of expiry due to the fact that the network database can be updated within a reasonable time frame.

The major advantage of navigation data is their commercial maintenance. NAVTEQ releases data updates on a quarterly basis. They include new streets as well as modifications, i.e. expansion or reduction of existing links. These streets are clearly allocated to the links from the entire network by using permanent IDs. The generalization is recorded in detail and is thus reversible.

As a result, all changes can be adopted within the network topology, even if the data has not been updated with each NAVTEQ release (usually quarterly).

3 Demand model

The demand model is based on the VISEVA-approach developed by LOHSE (1997). VISEVA is a program for solving trip generation and simultaneous origin-destination and mode choice (for model description see <http://www.ersa.org/ersaconfs/ersa02/cd-rom/papers/413.pdf>). The software itself is available within the ptv vision transportation planning suite (see <http://www.ptv-vision.com>).

In the **trip generation step**, planners address the problem where trips begin and end in a network. The modelling area is divided into homogeneous traffic zones. Demand detailed trip purposes are differentiated by origin-activity – destination-activity pairs. In the following figure the origin-destination-groups for the passenger transport in *Validate* are given:

to from	W	K	S	B	S	L	O	R
W	-							WR
K		-						OR
S			-			OO		OR
B				-				BR
S		OO			-			SR
L						-		OR
O							-	OR
R	RW	RO	RO	RB	RS	RO	RO	-

Nr. QZG	Discription
1	RW
2	RB
3	RS
4	RO
5	WR
6	BR
7	SR
8	OR
9	OO

R = residence
W = working place
K = Kindergarden
S = School/University
B = business
S = Shopping
L = Leisure
O = Others

Table 1: Origin-Destination-groups passenger traffic

Trip generation rates per activity pair per day are determined on the basis of the German nationwide travel behavior surveys SrV (1998) and MiD (2002), the former KONTIV.

Based on these categories and in regard of specific characteristics (e.g. average travel times, transport mode) a quite dis-aggregated reflection of the passenger traffic demand can be done.

The whole network is based on about 7,000 traffic zones where as there are about 14,000 communities in Germany. Traffic zones contain several communities in rural areas whereas in urban areas one community is divided into different traffic zones. The traffic zones itself contain several so called market cells which are derived from house inventories (in Germany more than 22 million houses are grouped by road sections in 85000 market cells) and commercially used for marketing purposes (for more information see www.dds.ptv.de). Among other the following structural data is used:

- Inhabitants (with 8 different age groups)
- Working population
- Working Places
- Motorization (cars/1000 inhabitants)

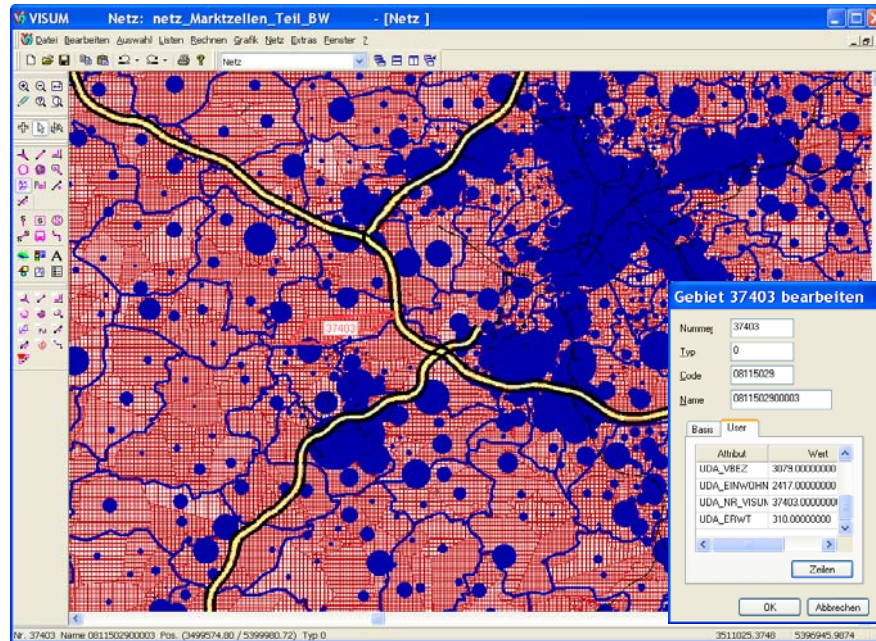


Figure 3: Traffic zones – enriched with working place figures from the market cells

Finally the trip generation step is performed by multiplying the trip rates mentioned above with structural zone data.

For **trip distribution** a special gravity model is used (VISEVA-function). Average trip length and trip length distributions from German National Surveys such as SrV (1998) and MiD (2002) helped to determine the parameter of the VISEVA-function. (for more information see www.tu-dresden.de/srv/ and www.kontiv2002.de).

Commuter trip tables (ca. 14000x14000 cells) from the Bundesagentur für Arbeit (German Federal Employment Agency) have been used to compensate the insufficiency of the gravity model for commuter trip distribution by introducing k-factors.

Finally all resulting OD-sums have been balanced by solving bi- or tri-linear equation systems to get consistent trip tables per purpose segment.

4 Calculating the traffic volume

To get traffic volumes in a network the O-D matrices are assigned with various methods to the road network.

- (1) Heavy Good Vehicles (3 weight classes) are assigned by a multiclass 4 step incremental procedure using 50% of the road capacity
- (2) Long distance (> 100km) car traffic is assigned by 4 step incremental or

alternatively by a stochastic assignment routine

- (3) Short distance (up to 100km) car traffic is assigned using a user equilibrium (UE) on 1% level accuracy (Wardrop's 1st principle)

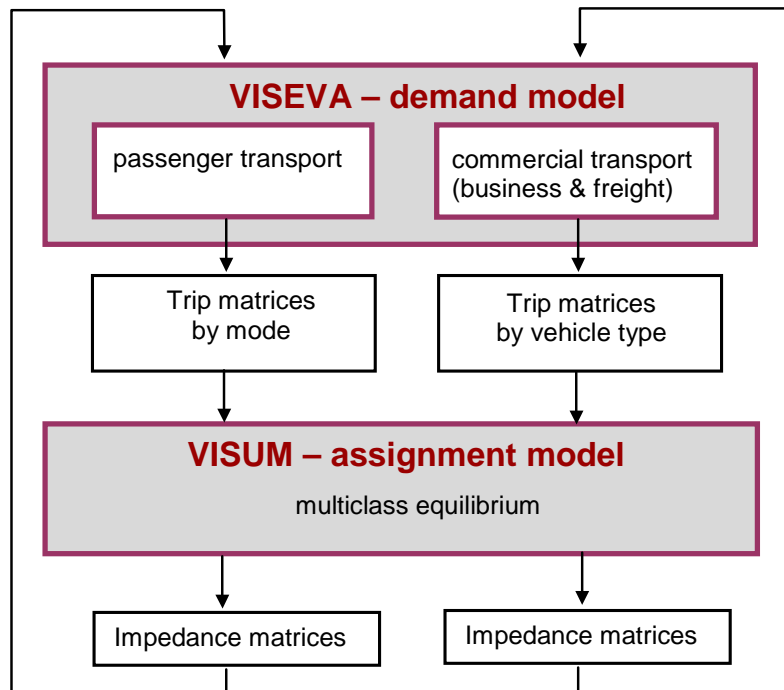


Figure 4: Iteration demand and assignment

Service indicators describing travel time between zones are calculated within the assignment procedure and fed back into the demand model. Here they have an impact on the destination choice and result in a trip length distribution which can be matched with the empirical data. Based on iteration feedback between demand and traffic volume calculation, the model reaches a state of equilibrium between supply and demand.

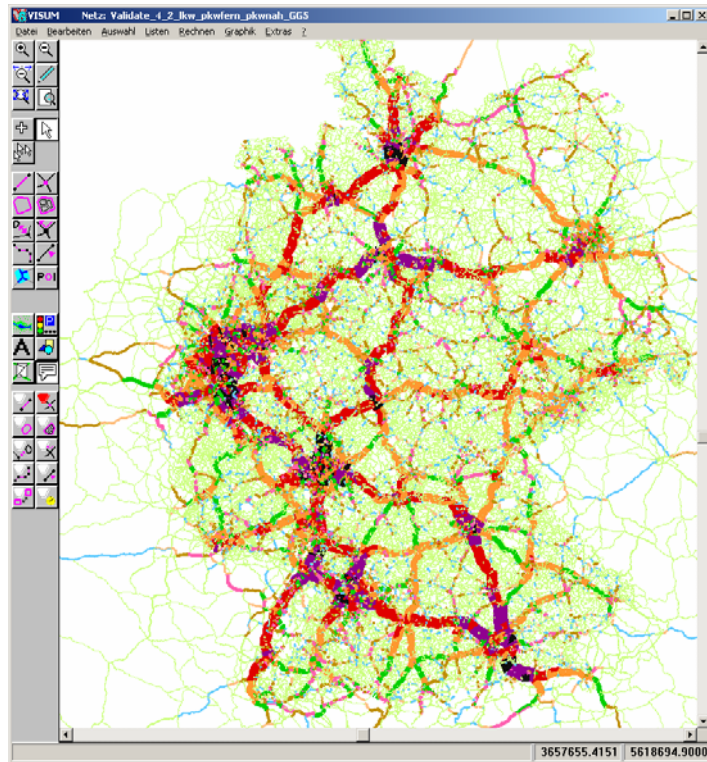


Figure 5: Validate road network with traffic volumes

4.1 Count data on roads

The national traffic counts from the year 2000 delivers around 70,000 survey points. These data and additional automatic recorded count values from 1,600 permanent survey points from BAST (Federal institute for roads) are referred to the Validate network.

At the moment (end of July 2004) a fuzzy based matrix estimation procedure is used to improve the matrices by adapting the OD flows to counted link values.

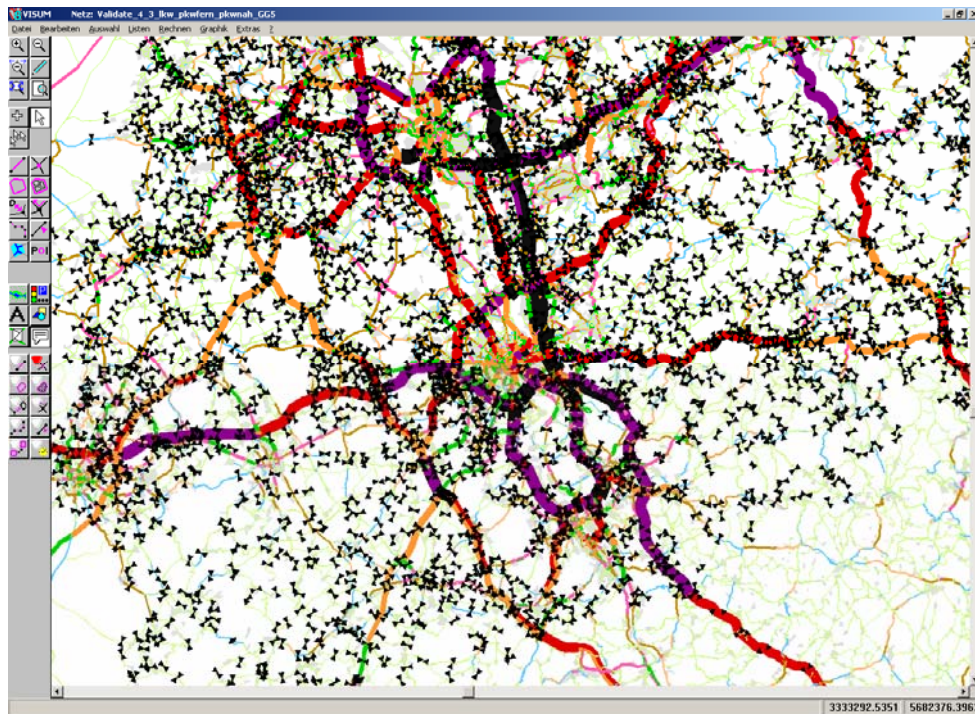


Figure 6: Network section with survey points

4.2 Evaluation of the quality

During the first stage of the project the fundamental feasibility of a model calculation of this size and granularity was of priority and could be proved, while the second fundamental step is the calibration of this model. The general consistency of the model is very good due to the clear linkage between demand and assignment calculations (usually 6 iterations) and the following matrix estimation. But finally the model output users compare assigned and observed link volumes and travel times in their individual local perspective. Therefore the criterion for measuring the quality of the model calculation is the difference between these values.

We are aiming at a correlation coefficient of 95% between survey values and calculated values. The difference between the total number of counted vehicles and assigned vehicles is lower than 1%. This should also be the case for partial networks taken from the global *Validate* VISUM-network as long as they contain at least 200 of the initial observation values.

One of the main difficulties when calibrating lies in the uncertainty of some empirical count data. Various of the survey points only represent a projection of an observed 3-4 h time intervals. Around 1,600 survey cross-sections originate from permanent survey points and can be seen as much more secure statistically.

5 Conclusion and outlook

The model calculations performed for *Validate*, with up to 7,000 traffic cells and over one million links, are, as far as we know, the world's largest model applications which have ever been performed for practical applications. A model iterations, i.e. demand and assignment calculation takes approx. 14 hours on a Pentium 3GigaHz with 3 GigaByte RAM.

The transportation network and traffic volumes gained from *Validate* will be used in many ways: For example as regional network sections with incoming and outgoing traffic, as data for the location evaluation of advertising space or commercial areas, for transportation planning investigations etc. A refinement is currently in process to generate hourly link volumes by a dynamic-stochastic assignment procedure. It is planned to use such detailed link volumes and travel times for dynamic route recommendations, for network disturbance simulations, for road work planning or for logistic tasks.

The results are very encouraging and have already found an entry in transportation planning and geo-marketing. For a transfer of this approach to other countries the general framework should be kept:

- Base the network model on a navigation database.
- Use the National traffic behavior data for calculating the traffic generation rates and calibration of modal split and traveling distances
- Apply structure data such as population and work places on an average zone size < 20,000 inhabitants or work places in urban and < 5,000 in rural areas . Use even more detailed data (e.g. cell information on a base of 1000 inhabitants or work places) to generate shares for multiple zone connectors
- Make use of count data and screen lines to validate the calculations.

References

Lohse, D., et al. (1997), "Ermittlung von Verkehrsströmen mit n-linearen Gleichungssystemen - Verkehrsnachfragemodellierung" Schriftenreihe des Instituts für Verkehrsplanung und Straßenverkehr, TU Dresden, Dresden.

Wardrop, J.G. (1952), "Some Theoretical Aspects of Road Traffic Research", Proceedings of the Institution of Civil Engineers, Part II, 1, 325-378.

Friedrich, M., Mott, P., Nökel, K.: Keeping Passenger Surveys up-to-date – A Fuzzy Approach, presented at TRB 2000.