UPDATE OF THE NEAC MODAL-SPLIT MODEL

Leest, E.E.G.A. van der
Duijnisveld, M.A.G.
Hilferink, P.B.D.
NEA Transport research and training

1 INTRODUCTION

The NEAC model and information system consists of models and methods for constructing databases describing the current freight transport flows and for developing forecasts of future freight transport flows. NEAC describes all intra- and inter-regional freight transport in Western and Eastern Europe by transport mode and by commodity type. Results are expressed in total weight of the goods. In NEAC, a classical four step modelling approach is used. This approach consists of the following steps:

1. Trade and transport generation;
2. Regional distribution;
3. Mode choice;
4. Route choice/assignment.

The third step, the model for mode choice, is the subject of this paper. In 2005, NEA carried out a research project to update the NEAC modal-split model. This paper presents the assumptions, highlights the results and also mentions the differences between the previous and the new version. The goal of the research project was to develop a model that could replace the outdated modal-split model. As well as the calibration of updated parameters, the aim was to also incorporate the transport mode short sea in the model. To achieve this NEA has developed and calibrated a completely new model. An analysis of the effects of the new NEAC modal-split model on forecasts of main European transport flows is presented.

The modal-split model for freight transport within the TRANS-TOOLS research project[1] is based on the new NEAC model. The goal of TRANS-TOOLS, which is co-funded by the European Commission (DG-TREN) under the 6th Framework Programme for Research and Development, is to produce European transport network models to overcome the shortcomings of current models.

2 THE NEAC MODEL AND INFORMATION SYSTEM

The NEAC model and information system is the collection of the databases and models which are described below. These modules are linked through their input and output. The basis of the system is the NEAC base year database. Several national and international trade and transport data sources are the input for the construction of the database. Several techniques are used to combine different data sources into one database. The idea behind the construction of the database is that trade flows determine transport flows. A trade flow is the economic activity between a production region and a consumption region. A transport flow is the shipment of commodities from a loading region to an unloading region on some transport mode. A transport chain is the sequence of transport flows from the first origin, the production region, to the final destination, the consumption region, without changing the commodity. Thus, when goods are transported from the production region to the consumption region without transhipment the transport chain consists of just one transport flow. Finally, the set of transport chains between two regions matches the trade flow between those regions. The database consists of the transport chains. These transport chains are divided into commodity groups. The transport mode at the origin and destination of the transport chain are defined together with one transhipment region, if known. To incorporate all this information into the database a top-down approach is used. The construction starts with the country to country trade information and then the trade flows are refined as much as possible using more detailed information where available. This way, when regional information is not available for all countries, it is still possible to include regional detail for the countries for which regional trade information is at hand. In the process of the top-down approach, the following four phases can be differentiated:

1. The building of a country-to-country matrix.
2. Including transhipment regions on the basis of transhipment statistics.
3. Regional division of country-to-country totals.
4. Incorporating domestic transport.

In the database the core countries are divided into regions, based on the NUTS level 2 regional division. Outside Europe the regions are countries or groups of countries. The transport modes specified in the database are transport over road, rail, inland waterways and sea. If the transport is carried out on any other mode or the mode is unknown, it is placed in the rest category. The freight flows are registered separately for each commodity group. In the NEAC database, the commodity group classification is the NST/R one digit classification with crude oil as a separate class.

All the models in the NEAC system need the input of one of the NEAC databases and a model-specific scenario. The output of all models is the NEAC database including data added by the model, like changes in the trade flows, modal-splits, containerization rates, or emissions. The output of the assignment model is the infrastructure network loaded with transport flows. Figure 1 presents a clear view on all the relations and connections between the models, databases, scenarios and external data sources used in the NEAC information system. A short description of the models is presented in the text below . The more elaborate description of the old and new modal-split models is presented in the next two chapters.

The base year database is constructed from the trade flows for all Origin-Destination (OD) relations. Structural economic relations are determined for the base year and it is assumed these will remain valid in the future. Therefore, it is possible to construct a gravity model that forecasts trade flows in future years. This model uses economic variables such as demand and supply attraction and resistance of the spatial divergence of regions caused by transport time, transport costs, and other trade constricting aspects. In the *NEAC trade model* the demand and supply attraction are represented by the added value of the sector that produces the commodity and the added value of the sector that consumes the commodity in the origin region and destination region, respectively. The distance between the trading regions is used to represent the resistance of the spatial divergence on trade. Furthermore, a dummy variable in the model captures circumstances favoring trade between certain regions, such as regions that are both in the same country or free trade zone. A scenario for the future developments of main economic variables corresponding to the required forecast year and the NEAC base year database are the input of the trade model. This model produces the forecast year databases with trade flows in the NEAC system.

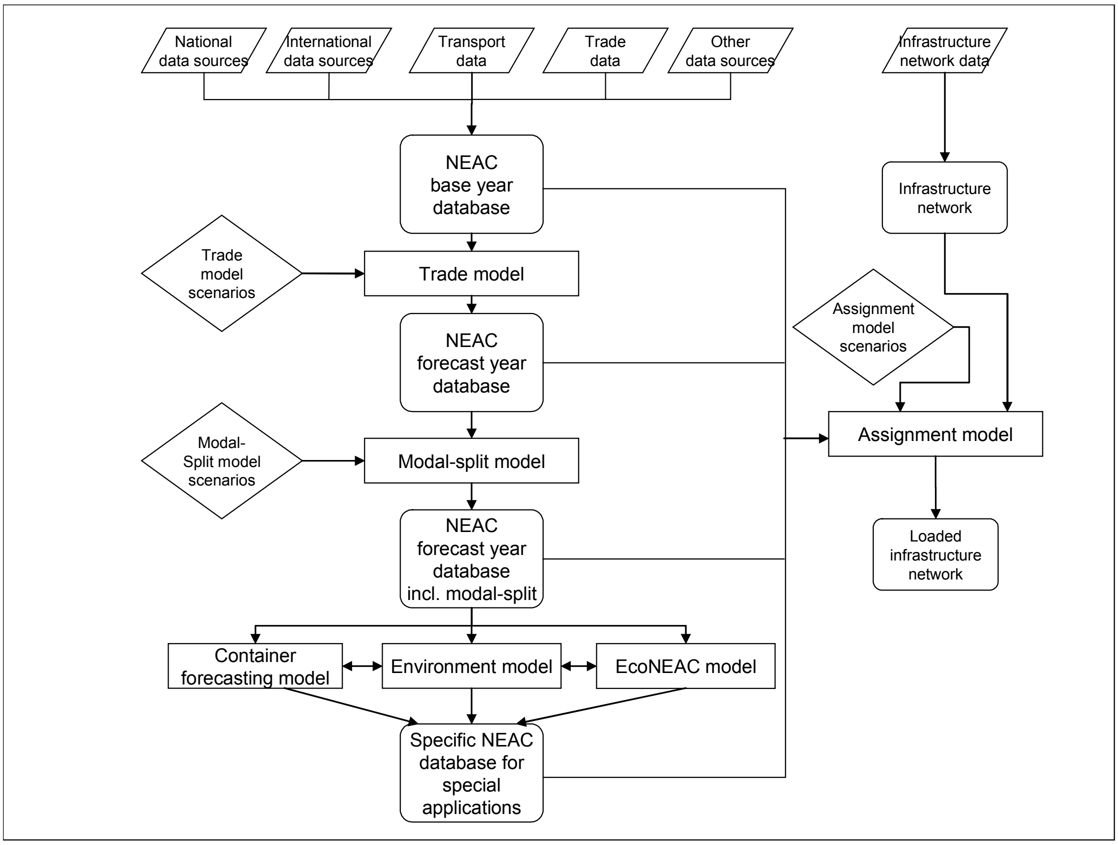

Figure 1: Schematic view of the databases and models in the NEAC system.

The old and the new *modal-split model* are discussed in the next two chapters. Besides the old modal-split model, the so-called NEAC conventional modal-split model, the NEAC modal-split module also consists of the NEAC analogy modal-split model. In the (former) candidate member states of the EU in Eastern Europe the transport markets are undergoing a process of restructuring. This leads to the convergence of those markets to the Western European situation. Therefore the changes in the modal-split in forecast years cannot be explained by just the changes in transport costs and times for these markets. In the analogy model the OD relations within Eastern Europe and between Eastern and Western Europe are matched to segments with comparable Western European OD relations. The rate of convergence is based on the economic growth and the completion and implementation of infrastructure projects in the countries of origin and destination.

The *container forecasting model* determines the development of the containerization rate based on the growth in commodity trade. First, the growth of the containerization rates between the base year and the forecast year is determined. Next, these growth statistics are used to convert the base year containerization rates into the forecast year rates. The last step applies the determined forecast year containerization rates to the forecast year database including modal-split information from the NEAC system.

The *environment model* estimates the energy consumption for the various transport flows considering the volume, transport modes, and OD relation of a flow. Loaded infrastructure networks are also input for this model. Energy consumption figures are used to calculate the emissions of those transport flows.

The *ecoNEAC model* deals with the dynamic relation between economy and transport. In this model transport generates trade, in contrast to the trade model. The ecoNEAC model provides a feedback to the trade flows. Trade flows that increase lead to more transport and to accommodate this transport new infrastructure is realized. This new infrastructure attracts new trade, thus the trade flows increase even more.

The *assignment model* projects the trade flow data on the infrastructure network. It assigns the total freight volume of a trade relation onto the network as a single flow and handles different assignment techniques. Furthermore, it can also assign the number of transport units used on the several possible routes between the origin and destination of a trade flow onto the network for distinct transport modes.

Examples of applications of NEAC are transport flow analyses, corridor analyses, infrastructure analyses, market potential analyses and policy impact analyses. For example, in the TEN-STAC project, the Trans European Network was revised with the help of NEAC. More information on the NEAC databases, models and the structure of the system as well as information on the used data sources and some applications of the NEAC system are available on the NEAC website (www.nea.nl/neac).

3 THE OLD MODAL-SLPIT MODEL

The modal-split model determines the division of the future trade over the different transport modes as a function of the changed transport costs and times in the forecast period for the available modes given the OD relation and commodity group. The new transport costs and times are based on a scenario of changes in infrastructure, policies and regulations, and cost structures. The output of this model is a forecast year database including the modal-split of the trade flows. The idea behind the modal-split model is that every OD relation is a separate market for freight transport. The transported volume on every mode is a function of the relative characteristics of the modes and the total volume for the OD relation. Thus the model is considered a variant of demand models.

The first step in the construction of the old NEAC modal-split model was the segmentation of the data. The data set was split up on commodity group and transport mode. Then for every subset a segmentation analysis was performed based on the distance and total annual volume in tonnes for the OD relation, commodity group, and transport mode. The method used for this segmentation is AID (automatic interaction detection), which uses a hierarchical binary splitting algorithm on a set of categorical variables. One explanatory variable is chosen as predictor for the division of the data set at every stage. The choice of the predictor is based on the minimization of the residual sum of squares of the dependent variable (criterion) of the divided data set. The splitting of the data set in two mutually exclusive and exhaustive subsets results from the division of the categories of the predictor variable in two groups. So observations in the data set with a value of the predictor in the first category go in the first subset and observations with a predictor value in the second category go in the second subset. This process is repeated creating subsets of subsets until no further improvement, that is no lower residual sum of squares of the criterion, can be attained by dividing the data set in more subsets.

The next step is the estimation of the modal-split model for the base year. The modes under consideration in the model are road, rail and inland waterways. The share of each mode is estimated with this equation:

$$P^{b}_{m,i-j} = \alpha \cdot RT_k^{\beta} \cdot RT_l^{\chi} \cdot RC_k^{\delta} \cdot RC_l^{\varphi} \cdot V_{i-j}^{\gamma}$$

$$k,l,m \in \{1,2,3\}; k \neq l \neq m.$$

Where:

$P^{b}_{m,i-j}$: Percentage of total transport on the OD relation with origin *i* and destination *j* transported on mode *m* in base year *b*.

RC_k: Ratio of transport costs per tonne of modes *m* and *k*.

RC_l: Ratio of transport costs per tonne of modes *m* and *l*.

RT_k: Ratio of transport times of modes *m* and *k*.

RT_l: Ratio of transport times of modes *m* and *l*.

V_{i-j}: Total annual volume (tonnes) on the O-D relation with origin *i* and destination *j*.

$\alpha, \beta, \gamma, \delta, \varphi, \chi$: Model coefficients.

A linear regression is performed on the logarithm of the equation for every segment. Forecasts for the shares of the transport modes in future years are computed with the following equation:

$$P^{f}_{m,i-j} = P^{b}_{m,i-j} \cdot \left(\frac{RT_k^f}{RT_k^b}\right)^{\beta} \cdot \left(\frac{RT_l^f}{RT_l^b}\right)^{\chi} \cdot \left(\frac{RC_k^f}{RC_k^b}\right)^{\delta} \cdot \left(\frac{RC_l^f}{RC_l^b}\right)^{\varphi} \cdot \left(\frac{V_{i-j}^f}{V_{i-j}^b}\right)^{\gamma}$$

$$k,l,m \in \{1,2,3\}; k \neq l \neq m.$$

Where:

b: Base year shares and relative characteristics.

f: Forecast year shares and relative characteristics.

For the forecast year the relative costs and times and the total annual volume are taken from the forecast scenario. The changes in these characteristics compared to the base year relative costs and times and total annual volume now are used to determine the forecast year shares through the elasticities of these characteristics. Since the percentages of the modes do not automatically sum to 100 for every OD relation, they must be adjusted to sum to 100 after estimation.

A remark on this model is that for segments with a high share for mode road model results were better when the complement of the dependent variable (100 - $P^{b}_{m=1,i\text{-}j}$) was used in the estimation. In this way the share of mode road is determined through the estimation of the combined share of all other modes. It is allowed to do this if it is assumed that the mode for which the complement is taken is always a competitor of all other modes. Since this is only considered to hold for mode road this is only done for the segments with a high share for mode road.

4 THE NEW MODAL-SPLIT MODEL

The development of the new modal-split model started with the choice of the type of model to use. The possible model types mentioned in literature and used in practical applications were studied and analyzed. Research for the old version of the model led to the conclusion that an aggregate analysis of the modal split on the basis of multinomial logit models renders no satisfactory results. The old NEAC modal-split model therefore contained a segmentation of the transport markets. Within each segment of the market the development of the modal split was dependent on the transport time and the cost of transport of each mode in relation to other modes (cross elasticities). Segments were identified based on commodity groups, distance and total weight transported between regions. In general, the previous model worked in such a way that an increase in relative tariffs of one mode led to an increase in the proportion of the other modes (the magnitude depends on the elasticities). The previous model was based on the theory that the main determinants of mode choice are cost and time of transport, which in their turn are dependent on the state of the infrastructure and policy measures which have an effect on the transport market.

The old NEAC modal-split model is of the demand model type. A disadvantage of this model is that the calculation of the forecast year modal-split with this equation implies that a mode with a zero percentage in the base year never has a positive percentage in the forecast year. It is possible to manually change the percentage of the base year from zero to a very small percentage to see if the share should grow in the forecast year. This change is also necessary in the estimation of the model since the logarithm of shares is used as the dependent variable in the regression equation and ln(0) is defined to equal minus infinity. Another disadvantage of this model is that its specification of the model equation implies the necessity of the ratios of transport costs and times of all modes for every OD relation, even when a mode is physically unavailable for an OD relation. This problem is solved by the definition of artificial cost and time values by the modeler. Since available modes with high costs and times can have a share of zero percent, the idea is that if costs and times of physically unavailable modes are high enough the shares of theses modes will be zero.

In the research project to update the modal-split model it was concluded that the previous version of the model was unsuitable for the more detailed regionalization used in the NEAC database and the added transport mode short sea. The number of cases to which the disadvantages of the old model apply is higher in the NEAC dataset used for the calibration of the new version of the model. The old model was estimated for three modes: road, rail, and inland waterways. Furthermore, the base year database used in the estimation of the old model also had less regional detail in Eastern Europe. The database for the new model includes transport chains between 268 regions in 37 European countries specified for eleven commodity groups: NSTR one digit with crude oil separate; and four modes of transport: road, rail, inland waterways and short sea. Sea transport is available for a substantial smaller number of O-D relations than road and rail transport. Most of the countries in Eastern Europe are connected to the inland waterways network, but usually only a few of the regions in every country are connected. Therefore, in the old database these countries consisted of one region for which the mode was available, whereas in the new database these countries consist of numerous regions of which only a few are connected to the inland waterways network. The number of regions not connected to the sea network is relatively high in Eastern Europe as well. So the improved regional detail and the higher number of modes cause the necessity of using artificial ratios of transport cost and time for more OD relations. For the same reasons there also is an increase in the number of cases of OD relations with a zero share for one or more of the available modes.

These mentioned disadvantages of demand models, together with the availability of software with the capability to estimate discrete choice models conveniently and the broad variety of these models used in the context of mode choice in both literature and practical applications favor this type of models. These are the main reasons for the choice to use discrete choice models for the development of the modal-split model[2]. From the several discrete choice models the logit model is chosen for the development of the NEAC modal-split model since logit models were the most used discrete choice models in literature on mode choice and modal-split models. The simplicity of the logit model structure and the fact that the probit model is computationally more demanding than the logit model were additional reason to use the logit model. Based on this analysis the choice was made to use logit models for the development of the modal-split model. This approach is also based on the theory that the main determinants of mode choice are cost and time of transport, which are dependent on the state of the infrastructure and policy measures which have an effect on the transport market.

It was concluded from the review of literature and applications of relevant models that the best approach for the new modal-split model would be the use of logit models combined with an adjusted segmentation of transport markets.

Without segmentation analysis of the data the construction of a model with ample explanatory power seems impossible. The data is too heterogeneous even if the data is split up for available modes and commodity groups. The cost and time variables have a low explanatory power and the mode choice is not very sensitive to changes in these variables. Such a model is not useful for estimating the effects of policy and price changes on the modal-shift, thus a segmentation analysis of the data set is performed. Each OD - commodity group relation is considered a separate market for transport services. Not all these markets have the same characteristics. The goal of the segmentation analysis is to divide the data set into groups of markets with homogeneous characteristics. For these less heterogeneous partial data sets better models can be estimated.

For the segmentation analysis the CHAID (Chi-Squared Automatic Interaction Detection) technique is used to split the data in homogeneous groups. CHAID is a criterion-based segmentation tool limited to nominal and ordinal categorical variables. It is an evolution of AID (automatic interaction detection), which uses a hierarchical binary splitting algorithm on a set of categorical variables. One explanatory variable is chosen as predictor for the division of the data set at every stage. The choice of the predictor is based on the minimization of the residual sum of squares of the dependent variable (criterion) of the divided data set. The splitting of the data set in two mutually exclusive and exhaustive subsets results from the division of the categories of the predictor variable in two groups. So observations in the data set with a value of the predictor in the first category go in the first subset and observations with a predictor value in the second category go in the second subset. This process is repeated creating subsets of subsets until no further improvement, that is no lower residual sum of squares of the criterion, can be attained by dividing the data set in more subsets. Some drawbacks of AID are a bias for the selection of variables with more than two categories as the best predictor and the possibility that the choice of an optimal split in an early stage can lead to a sub-optimal final solution: it may be possible to reach a lower minimum residual sum of squares for the final division of the data set by choosing a predictor that does not minimize the residual sum of squares at an early stage. CHAID eliminates these drawbacks. Furthermore, CHAID considers the whole distribution of the dependent variable, it is not restricted to binary splits, and it makes no assumptions of normality. It utilizes the chi-square test for independence to assess statistical significance. CHAID is able to automatically identify significant interaction effects between categories of dependent/criterion variables. This gives model developers the opportunity to avoid flaws in model specification, in particular biases resulting from omitting relevant interactions. CHAID operates according a two step algorithm: merging and splitting.

The first step is merging. For each pair of categories eligible to be merged of each predictor, chi-squared tests are computed to test for independence in the subset of data formed by the two categories of the predictor variable and the dependent variable. Among those pairs found to be non-significant, the most similar are merged. These pairs have the smallest chi-square value. In the algorithm, for any subset containing three or more categories, a test of the significance for any predictor associated with a category against the other categories in that subset is performed to see if any predictor should be unmerged. If more than one category is eligible to be unmerged, the one having the highest chi-square is selected to be unmerged. These procedures are repeated until only significant pairs remain and thus no categories are eligible to be unmerged. Then the probability p that the observed sample relationship between the predictor and the dependent variable would occur if the two variables were in fact statistically independent is calculated. The second step is splitting. The predictor with the lowest significant probability p is selected and the group is split on this predictor. If no predictor has a significant probability p, the group is not split. These steps are repeated until all subgroups have been analyzed or contain too few observations. The result is a tree where the data set is the root and the branches are the more homogeneous subsets.

OD relations with a different availability of modes of transport have different markets for mode choice. The modal-split model is estimated separately for these markets. There are five possible combinations of available modes:

1. Modes road and rail available;
2. Modes road and short sea available;
3. Modes road, rail and inland waterways available;
4. Modes road, rail and sea available;
5. Modes road, rail, inland waterways and short sea available.

For each of these five markets a segmentation analysis is carried out. Based on preliminary test results and expert opinion, the decision was made to use the commodity group, the location of the origin and destination region, the distance and the tonnage as explanatory variables in the segmentation analysis. The respective categories used are the commodity groups; Scandinavia, UK & Ireland, Western Europe and Eastern Europe for the location; less than 400 km, 400 to 700 km, 700 to 1000 km, 1000 to 1500 km and more than 1500 km for the distance and less than 5000 tonnes, 5000 – 50,000 tonnes, 50,000 – 500,000 tonnes and more than 500,000 tonnes for the tonnage. The modal-split in the base year is the dependent variable.

A logit model is a discrete choice model and is based on random utility theory. In the context of mode choice an individual with a demand for transport must choose a specific mode from the set of available transport modes to carry out the demanded transport. The individual will pick the most attractive alternative based on the characteristics of the available modes. The most attractive alternative is described as the one that yields the most utility to the individual. The utility of an alternative is a function of the characteristics of that alternative and of the characteristics of the individual making the choice. Because many of the characteristics are not observable, these enter the model as a random component of utility. Now the systematic or observable utility is defined as a function of all observable characteristics of the individual and the alternative. Total utility is defined as the sum of the systematic utility and a random component:

$$U_{m,n} = V_{m,n} + \varepsilon_{m,n}$$

Where:

$U_{m,n}$: Total utility of mode *m* for individual *n*

$V_{m,n}$: Systematic utility of mode *m* for individual *n*
$\varepsilon_{m,n}$: Random component of utility of mode *m* for individual *n*

The individual maximizes utility and thus chooses the alternative with the highest total utility $U_{m,n}$, but in the model only the $V_{m,n}$ can be observed. So now it is possible to define the probability $P_{m,n}$ that individual *n* chooses mode *m* with this equation:

$$
\begin{aligned}
P_{m,n} &= Prob\left\{U_{m,n} \geq U_{l,n}, \forall l \in M, l \neq m\right\} \\
&= Prob\left\{V_{m,n} + \varepsilon_{m,n} \geq V_{l,n} + \varepsilon_{l,n}, \forall l \in M, l \neq m\right\} \\
&= Prob\left\{V_{m,n} - V_{l,n} \geq \varepsilon_{l,n} - \varepsilon_{m,n}, \forall l \in M, l \neq m\right\} \\
&= Prob\left\{\varepsilon_{l,n} - \varepsilon_{m,n} \leq V_{m,n} - V_{l,n}, \forall l \in M, l \neq m\right\}
\end{aligned}
$$

Where:
M: Set of available modes
$P_{m,n}$: Probability that individual *n* chooses mode *m*.

The probability can be calculated if the distribution of all $\varepsilon_{m,n}$ is specified. For logit models it is assumed the $\varepsilon_{m,n}$ are independent and identically distributed Gumbel. Under this assumption ($\varepsilon_{l,n} - \varepsilon_{m,n}$) is logistically distributed. For the probit model it is assumed that the vector $\varepsilon_n = [\varepsilon_{1,n}, \varepsilon_{2,n}, \ldots, \varepsilon_{M,n}]$ is multivariate normal distributed with mean 0 and variance-covariance matrix Σ_ε. Based on this equation and the mentioned assumptions on the distribution of the $\varepsilon_{m,n}$ the multinomial logit model (MNL) can be derived. This derivation is available in a great number of ways in literature[3]. This equation expresses the model:

$$
P_{m,n} = \frac{e^{V_{m,n}}}{\sum_{l \in M} e^{V_{l,n}}} = \frac{e^{\beta' x_{m,n}}}{\sum_{l \in M} e^{\beta' x_{l,n}}}
$$

$$
\left[V_{m,n} = \beta' x_{m,n}\right]
$$

Where:
M: Set of available modes.
$P_{m,n}$: Probability that individual n chooses mode m.
$V_{m,n}$: Systematic utility of mode m for individual n.
$x_{m,n}$: Observed characteristics of mode m and individual n.
β: Logit parameter.

The function for the systematic utility $V_{m,n}$ is defined as the multiplication of the vectors β and $x_{m,n}$. The logit parameters are a set of coefficients for the set of observed characteristics.

The selected configuration of the new modal-split model is such a multinomial logit model and the parameters of the model have been calibrated separately for every segment. The multinomial logit model is used and the choice probabilities of the available modes per commodity group for every OD relation are determined by the following formula:

$$P_{m|cij} = \frac{e^{V_{m|cij}}}{\sum_{l \in M} e^{V_{l|cij}}}$$

$$with: V_{m|cij} = \beta_{m0} + \sum_k \beta_{mk} x_{cijmk}$$

Where:

M: Set of available modes.

$P_{m|cij}$: Choice probability of mode *m* given commodity group *c* and OD relation *ij*.

$V_{m|cij}$: Systematic utility of mode *m* given commodity group *c* and OD relation *ij*.

x_{cijmk}: Explanatory variable *k* for mode *m* given commodity group *c* and OD relation *ij*.

β_{mk}: Logit parameter for mode *m* and level of service *k*.

In the modal-spit model the market shares of the different modes of transport are estimated for every OD relation (*ij*) and commodity group (*c*). The explanatory variables used in the model are the alternative specific constant, the relative total cost per ton, the relative total time, the distance, the total annual transport volume, border resistance dummies for road and rail transport, domestic and intercontinental transport dummies and a main port region dummy. The dependent variable in the model is the market share of each available transport mode. The choice probabilities are calculated with the forecast year level of services (explanatory variables) and the shift between the base year and forecast year choice probabilities is applied to the base year modal-split with forecast year tonnages in the forecasted OD matrix resulting from the NEAC trade model.

5 DATA AND SCENARIOS

The database for the base year is the ETIS[4] Freight matrix for 2000. The database for the forecast year is the ETIS Freight matrix for 2020 before the modal-split (thus after trade model run, still with base year modal-split). Model runs were carried out with the old and new modal-split models. Five scenarios have been used:

Scenario 1: 10% decrease of road generalized costs;
Scenario 2: 10% decrease of road travel times;
Scenario 3: 10% decrease of rail generalized costs;
Scenario 4: 10% decrease of rail travel times;
Scenario 5: TEN STAC TrendPlus scenario: Modal Split input
- Specific developments per mode and NSTR chapter,
- Basic policy actions,
- Infrastructure assumptions,
- Accompanying measures.

Scenario 5 is a TEN-STAC[5] scenario with normal transport developments, which follows the European transport policy, which is mentioned in the White Paper[6]. The scenario's available are specifically made for the old model and are incorporated into the new model.

6 RESULTS

The comparison of the results of the two models for the five scenarios is somewhat difficult to make. The scenarios had to be suitable for both the old and new NEAC modal-split model. Since the old model has no input for costs and times for mode sea (the old model does not includes modal shifts for sea) the input for the new model is set to equal the base year level of service for sea. The new model also includes the EU border as resistance, by changing the value of the dummy variable for road transport for OD relations involving countries that have acceded to the EU between base year and scenario year (according to the scenario). The new model also is calibrated and based on more recent data with more regional detail which makes it impossible to state something on what the influence of the methods used for the models and the different datasets used for calibration is on the resulting modal-split.

The tables 1, 2 and 3 illustrate the results of the old and new model for the five scenarios compared to the input data. Flows with sea transport are excluded since the old model and the scenarios do not include sea transport.

The analysis of the output data shows that the new model in general has smaller shifts between modes. Also road transport is at a slightly higher base level given the trends in logistics, like supply chain management and just in time delivery, and the change in modal-split for transport in relation with the new member states of the EU. The new model is calibrated on more recent data and takes these factors into account.

Table 1: Mode at origin by mode at destination; NST/R commodity group 9 (Machinery, transport equipment, manufactured articles and miscellaneous articles).

		road	rail	inland waterways
		% of total	% of total	% of total
road	input	93.8%	.1%	.0%
	scen 1 old	94.5%	.1%	.0%
	scen 1 new	94.1%	.1%	.0%
	scen 2 old	93.9%	.1%	.0%
	scen 2 new	93.9%	.1%	.0%
	scen 3 old	93.1%	.1%	.0%
	scen 3 new	93.6%	.1%	.0%
	scen 4 old	93.5%	.1%	.0%
	scen 4 new	93.8%	.1%	.0%
	scen 5 old	92.7%	.1%	.0%
	scen 5 new	93.6%	.1%	.0%
rail	input	.2%	5.5%	.0%
	scen 1 old	.2%	4.9%	.0%
	scen 1 new	.2%	5.2%	.0%
	scen 2 old	.2%	5.4%	.0%
	scen 2 new	.2%	5.4%	.0%
	scen 3 old	.2%	6.2%	.0%
	scen 3 new	.2%	5.7%	.0%
	scen 4 old	.2%	5.9%	.0%
	scen 4 new	.2%	5.5%	.0%
	scen 5 old	.2%	6.5%	.0%
	scen 5 new	.2%	5.7%	.0%
inland waterways	input	.0%	.0%	.4%
	scen 1 old	.0%	.0%	.3%
	scen 1 new	.0%	.0%	.4%
	scen 2 old	.0%	.0%	.4%
	scen 2 new	.0%	.0%	.4%
	scen 3 old	.0%	.0%	.4%
	scen 3 new	.0%	.0%	.4%
	scen 4 old	.0%	.0%	.4%
	scen 4 new	.0%	.0%	.4%
	scen 5 old	.0%	.0%	.5%
	scen 5 new	.0%	.0%	.4%

Table 2: Mode at origin by mode at destination; domestic freight flows in Germany.

		road	rail	inland waterways
		% of total	% of total	% of total
road	input	91.6%		
	scen 1 old	92.7%		
	scen 1 new	91.9%		
	scen 2 old	91.9%		
	scen 2 new	91.7%		
	scen 3 old	90.7%		
	scen 3 new	91.5%		
	scen 4 old	91.2%		
	scen 4 new	91.6%		
	scen 5 old	89.3%		
	scen 5 new	91.5%		
rail	input		6.7%	
	scen 1 old		5.9%	
	scen 1 new		6.4%	
	scen 2 old		6.4%	
	scen 2 new		6.6%	
	scen 3 old		7.6%	
	scen 3 new		6.8%	
	scen 4 old		7.1%	
	scen 4 new		6.6%	
	scen 5 old		8.2%	
	scen 5 new		6.8%	
inland waterways	input			1.8%
	scen 1 old			1.4%
	scen 1 new			1.7%
	scen 2 old			1.7%
	scen 2 new			1.7%
	scen 3 old			1.7%
	scen 3 new			1.7%
	scen 4 old			1.8%
	scen 4 new			1.7%
	scen 5 old			2.5%
	scen 5 new			1.8%

Table 3: Mode at origin by mode at destination; export from France.

		road	rail	inland waterways
		% of total	% of total	% of total
road	input	75.3%	1.5%	1.1%
	scen 1 old	78.9%	1.3%	1.1%
	scen 1 new	75.4%	1.5%	1.1%
	scen 2 old	75.5%	1.4%	1.1%
	scen 2 new	75.3%	1.5%	1.1%
	scen 3 old	72.4%	1.5%	1.0%
	scen 3 new	75.0%	1.5%	1.1%
	scen 4 old	74.0%	1.5%	1.1%
	scen 4 new	75.1%	1.5%	1.1%
	scen 5 old	72.0%	1.5%	1.1%
	scen 5 new	75.0%	1.5%	1.1%
rail	input	1.8%	10.5%	.2%
	scen 1 old	1.8%	9.0%	.1%
	scen 1 new	1.8%	10.5%	.2%
	scen 2 old	1.8%	10.7%	.2%
	scen 2 new	1.8%	10.6%	.2%
	scen 3 old	1.8%	12.9%	.2%
	scen 3 new	1.9%	10.8%	.2%
	scen 4 old	1.8%	11.5%	.2%
	scen 4 new	1.9%	10.7%	.2%
	scen 5 old	1.8%	11.7%	.2%
	scen 5 new	1.9%	10.7%	.2%
inland waterways	input	.5%	.0%	9.1%
	scen 1 old	.4%	.0%	7.4%
	scen 1 new	.5%	.0%	9.0%
	scen 2 old	.6%	.0%	8.7%
	scen 2 new	.5%	.0%	9.0%
	scen 3 old	.5%	.0%	9.6%
	scen 3 new	.5%	.0%	9.0%
	scen 4 old	.5%	.0%	9.4%
	scen 4 new	.5%	.0%	9.1%
	scen 5 old	.6%	.0%	11.2%
	scen 5 new	.5%	.0%	9.1%

7 CONCLUSION

The old and the new NEAC modal-split model are based on the same idea about freight mode choice. The used models differ considerably however. The new logit model is better able to handle the more detailed regional division in Eastern Europe and includes modal shift of the mode sea. The model benefits from the calibration with more detailed and recent data. For the used scenarios the new model generally has less extreme shifts between modes. The new model also adjusts for the changing modal-split due to the accession to the EU of new member states and trends in logistics such as just in time deliveries and supply chain management which lead to an increase in road transport.

BIBLIOGRAPHY

Ben-Akiva, M., and Lerman S.R. (1985) *Discrete choice analysis: Theory and Application to Travel Demand*, The MIT Press, London.

Chen, T.M., et al. (2005) *Deliverable 3: Report on model specification and calibration results TRANS-TOOLS (TOOLS for TRansport forecasting ANd Scenario testing) Deliverable3. Funded by 6th Framework RTD Programme*, TNO Inro, Delft.

European Commission (2001) *WHITE PAPER – European transport policy for 2010: time to decide*, Office for Official Publications of the European Communities, Luxembourg.

Leest, E.E.G.A. van der (2005) *Development of a New Modal-Split Model for the NEAC system; the Modal-Split in European Freight Transport* (Thesis), Tilburg University, Tilburg.

NEA, IWW, COWI, NESTEAR, PWC Italy, TINA, IVT, HERRY, and Mkmetric (2003) *Description of the base year 2000 and forecasts 2020. Deliverable D3 of the TEN-STAC project funded by the European Commission (DG TREN)*, NEA, Rijswijk.

NEA, IWW, NESTEAR, MKmetric, ISIS, MDS, IVT and VTT (2005) *ETIS-BASE D7 ETIS database methodology application by data set: freight and passenger demand, external effects*, NEA, Rijswijk.

Notes

[1]For more information on TRANS-TOOLS and the implementation of the modal-split model in the TRANS-TOOLS model see: Chen, T.M., et al. (2005).
[2]For an extended discussion on literature on and practical applications of modal-split models see: Leest, E.E.G.A. van der (2005).
[2]See for instance Ben-Akiva, M., and Lerman, S.R. (1985).
[4]For more information on ETIS and the ETIS freight demand matrix see: NEA, IWW, NESTEAR, MKmetric, ISIS, MDS, IVT, and VTT (2005).
[5]For more information on TEN STAC and a description of the scenario see: NEA, IWW, COWI, NESTEAR, PWC Italy, TINA, IVT, HERRY, and Mkmetric (2003).
[6]See European Commision (2001).