

Calibration and Validation of Strategic Freight Transportation Planning Models with Limited Information

Bart Jourquin

Université catholique de Louvain, ILSM, Center for Supply Chain Management, Mons, Belgium

Email: bart.jourquin@uclouvain.be

How to cite this paper: Jourquin, B. (2016) Calibration and Validation of Strategic Freight Transportation Planning Models with Limited Information. *Journal of Transportation Technologies*, 6, 239-256.

<http://dx.doi.org/10.4236/jtts.2016.65023>

Received: July 13, 2016

Accepted: September 23, 2016

Published: September 26, 2016

Copyright © 2016 by author and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Strategic transportation network models are often used as support tools in the framework of decisions to be taken at the policy level, such as the Trans-European Network projects. These models are mostly setup using aggregated or limited data. If their calibration is regularly mentioned in the literature, their validation is barely discussed. In this paper, several modal choice model specifications that make only use of explanatory variables available at the network level are described and applied to a large scale case. A validation exercise is performed at three levels of aggregation. The paper is designed from a strategic transport planning perspective, and does not present new modal choice formulations or assignment procedures. Its main added value is the focus on calibration and validation considerations. Despite the limited explanatory information used, the global performance of the best models can be considered as satisfactory. However, the quality of the models varies from mode to mode, the use of railway transport being the most difficult to predict without more specific input.

Keywords

Calibration, Validation, Network, Freight Transport, Modal Choice, Assignment

1. Introduction

Strategic multimodal freight transportation network models are often used in the framework of transport policy decisions, in order to estimate the impacts of a new large infrastructure on traffic or modal split for instance.

These models mostly cover several countries and are based on rather aggregated information, such as origin-destination (OD) matrixes containing data collected or esti-

mated at a regional level, such as the “NUTS 2” level commonly used in Europe. The calibration and validation of such models are difficult because, in most cases, only a few explanatory variables, such as the transportation cost or distance per trip and per mode are available. Moreover, only limited observed data can be used for validation purposes.

Calibration and validation are two different tasks. Calibration involves estimating the values of various constants and parameters in the model structure, while validation is the application of the calibrated models and the comparison of the results against observed data.

Even if not always explicitly cited as such, the calibration step is described in most papers, as it corresponds to the estimation of the parameters of a modal choice model for instance: the values of these parameters are often published, along with some performance measures such the Log-Likelihood or the Akaike Information Criterion (AIC).

The validation step in regional, national or international multimodal freight models is, however, considerably less documented, while it has a very important impact on the level of confidence one can have in a model (de Jong *et al.* [1]). An interesting discussion about the reasons why calibration and validation of such models is difficult can be found in Zhang [2]. She points out the large number of (elements in each) variable(s) or the lack of availability of reference data. She cites two papers in which the authors have put an effort in validating their model. The first, Jourquin [3], performs a validation by comparing the modelled modal shares of road, rail, and inland waterway transport to the observed ones, per category of commodities. If this indicates that each modality in the network bears the right amount of freight flows, it does not guarantee that the flows are appropriately assigned to the right routes. The second paper, by Yamada *et al.* [4], also presents a model with three modes (road, rail, and sea) and two types of users (freight and passenger). The modal split estimated for this model was validated by comparing the modelled link flows with the actual link traffic counts, but the node flows were not calibrated or validated. As a result, when the model is used for node flow estimation, it is difficult to assess the validity or reliability of the results. Therefore, Zhang proposes her own freight model for road, rail and inland waterways in the Netherlands, calibrated and validated at the mode, route and node levels.

The present paper examines, in a systematic way, several modal choice model specifications applied to a model that covers a very large area (Europe), and for which only very limited or aggregated data is available for calibration and validation. This empirical analysis uses regional OD matrixes and the tested modal choice model utility functions only use explanatory variables (costs, lengths, durations) that are available from a network model. The objective of the exercise is to compare the quality of several multimodal traffic assignments, each one being based on a different multinomial logit modal-choice model specification.

In a comparable manner to what was done by Zhang [2], the validation of the models is performed at three levels of aggregation:

- A highly aggregated level, for which the global shares for each transportation mode

- and each group of commodities are estimated and compared to the observed ones;
- A first disaggregated level, from the node point of view, for which the estimated modal split is measured and compared to the actual one for each origin-destination pair;
 - A second disaggregated level, from a link perspective, for which the assigned flow on the network is compared to the “observed” one.

Beyond this introduction, the next section gives a brief presentation of some modal-choice model specifications. Section 3 describes the real world large scale multimodal freight transport network model under analysis. Several model specifications are estimated with classical econometric tools. For benchmarking purpose, an alternative and rather crude iterative methodology is applied to some of the previously tested specifications. All these results are compared and discussed. The conclusive section opens some perspectives.

2. Multinomial Logit Modal Choice Formulations

Most econometric choice models are based on the theory of utility maximization. The simplest random utility maximizing model used to analyze choices is the binary Logit model which can be extended to the multinomial Logit (Equation (0)) with more than two choices. Although this model has some limits set by its basic assumptions (Ben-Akiva and Bierlaire [5]), it is well-known and widely used.

$$\Pr_i = \frac{\exp(V_i)}{\sum_{j=1}^k \exp(V_j)} \quad (0)$$

where \Pr_i is the probability to choose alternative i among k alternatives, and V_i is the utility of alternative i .

The utility V_i is most often defined as a linear combination of one or more explanatory variables, each one being weighted by a parameter that needs to be estimated. In the simplest cases, with the total cost of transport TC_i for the mode i as the only variable we could have:

$$V_i = -\alpha TC_i \quad \text{or} \quad (1)$$

$$V_i = -\alpha_i TC_i \quad \text{and} \quad (2)$$

$$V_i = -\alpha TC_i + \delta_i \quad (3)$$

$$V_i = -\alpha_i TC_i + \delta_i. \quad (4)$$

Equations (1) and (3) are generic (or conditional) logits, involving a coefficient α common to each mode. Equations (1) and (2), without an estimated intercept, are seldom or never used. Equation (2) will, however, be tested in this paper, because its specification is very close to an alternative one, presented later, and that is used for benchmarking purposes.

The data available for this case study allows to introduce an additional variable T_i for the time duration of each transportation task. The software package that is used also permits to separately compute, for each route, the fixed and variable costs, respectively

FC_i and VC_i . Hence, additional Equations (5) and (6) can be considered:

$$V_i = -\alpha_i TC_i - \beta_i T_i + \delta_i, \tag{5}$$

and

$$V_i = -\alpha_i FC_i - \gamma_i VC_i - \beta_i T_i + \delta_i. \tag{6}$$

Beside the widely used Logit model, the literature sometimes mentions the so-called “Law of Abraham” (Abraham and Coquand [6]). Originally, this formulation, represented by Equation (7), was designed to spread traffic over two alternative routes of total costs TC_1 and TC_2 .

$$\frac{f_1}{f_2} = \left(\frac{TC_1}{TC_2} \right)^{-n} \tag{7}$$

where f_1 and f_2 represent the flow assigned on routes 1 and 2 and n is a positive coefficient. The latest is often fixed to a value between 8 and 10, or even 14 (Bonnel [7]). As illustrated by Equation (8), this formulation can be generalized to the multinomial case so that the market shares for each alternative i appear as:

$$Pr_i = \frac{TC_i^{-n}}{\sum_{j=1}^k TC_j^{-n}} \tag{8}$$

This formulation is rather similar to the logit formula, and it can be seen as equivalent to a logarithmic logit model (Leurent [8]), at least for the univariate linear case (Gaudry *et al.* [9]). Equations (9) and (10) illustrate how the utility function can be written in such a case.

$$V_i = -n \ln(TC_i), \text{ or} \tag{9}$$

$$V_i = -n_i \ln(TC_i), \text{ with a mode specific } n_i \tag{10}$$

This model can then be simply extended as in Equations (11) and (12). These formulations allow for a classical estimation of n and δ or n_i and δ_i .

$$V_i = -n \ln(\alpha_i TC_i) = -n \ln(\alpha_i) - n \ln(TC_i) = \delta_i - n \ln(TC_i) \tag{11}$$

$$V_i = -n_i \ln(\alpha_i TC_i) = -n_i \ln(\alpha_i) - n_i \ln(TC_i) = \delta_i - n_i \ln(TC_i). \tag{12}$$

Note that, in the case of a common $n = 1$, the estimated share for each alternative is simply inversely proportional to the TC_i 's.

Finally, the logarithmic Logit formulation can be further generalized to other explanatory variables, as for instance in Equation (13) in which duration is included, or Equation (14) with an additional distinction between fixed and variable costs:

$$V_i = \ln(\delta_i TC_i^{-n} T_i^{-p}) = -n_i \ln(TC_i) - p_i \ln(T_i) + \ln(\delta_i) \tag{13}$$

$$V_i = -n_i \ln(FC_i) - m_i \ln(VC_i) - p_i \ln(T_i) + \ln(\delta_i). \tag{14}$$

For benchmarking purposes, it is also interesting to measure how all these models perform better than a very simple method, not based on the well-known maximum likelihood.

Using the same dataset, it is possible to estimate the parameters of some modal split

models using a simple iterative procedure. The rather crude algorithm that is presented here is, by nature, able to estimate only one coefficient per mode, so that only univariate utility functions can be used, such as $U_i = -\alpha_i TC_i$:

- 1) **Set** the initial values of the coefficients to estimate, α_i to 1.0.
 - 2) **Compute** the $-\alpha_i TC_i$ cost of each modal route.
 - 3) **Apply** the modal choice function to the OD matrix.
 - 4) **Compare** the global estimated modal share of each mode to the observed ones.
 - a) **If** the estimated global market share of all the modes is close enough (<0.001% for instance) to the observed ones, **stop** the algorithm.
 - b) **Else, adjust** the α_i coefficients: if the estimated global modal share of a mode is higher than the observed one, increase α_i so that the use of the mode i becomes more expensive. Decrease the value of α_i if the estimated modal share is lower than expected¹.
- Return** to step 3.

This algorithm is rather straightforward, and it stops once the difference between the observed and estimated global market shares of each mode is lower than a given threshold. Such an “objective function” is very different from the one used to solve the econometric models presented earlier, based on the well-known maximum likelihood. The performances of the models that make use of the parameters estimated by this iterative method can be considered as lower bound benchmarks.

This procedure, which doesn't need a particular software such as SAS or R, will be applied to the linear multinomial logit model and the logarithmic multinomial logit (aka Abraham) specifications. For the latest, a way to find out which value(s) of n to use will be discussed.

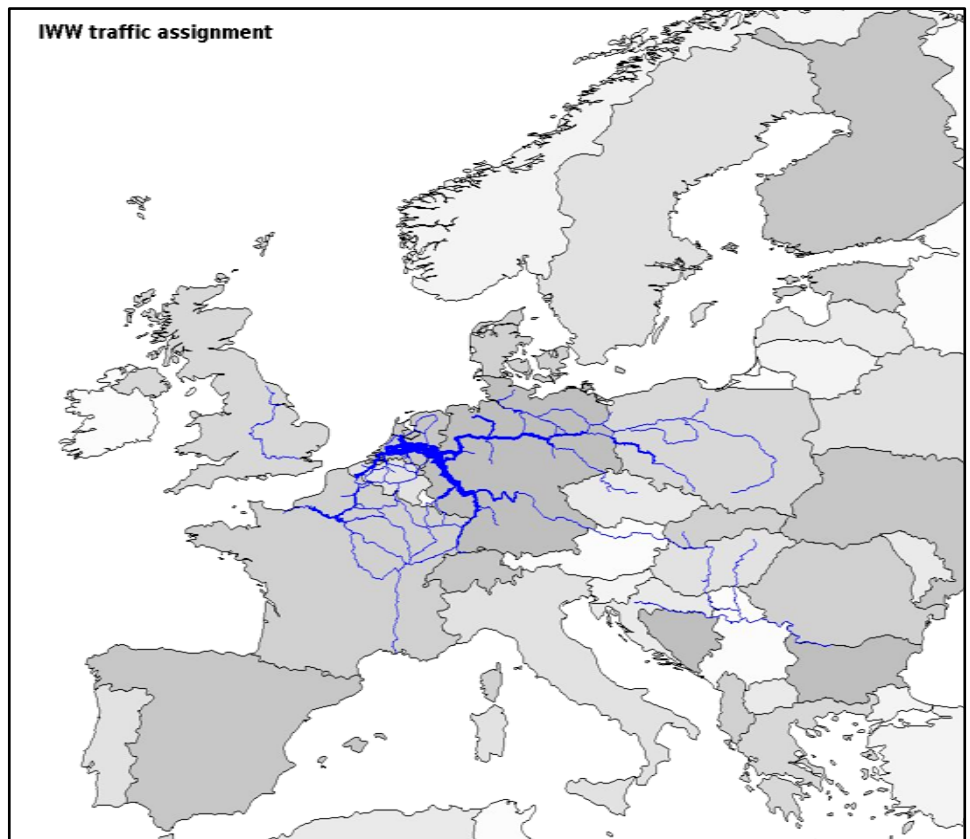
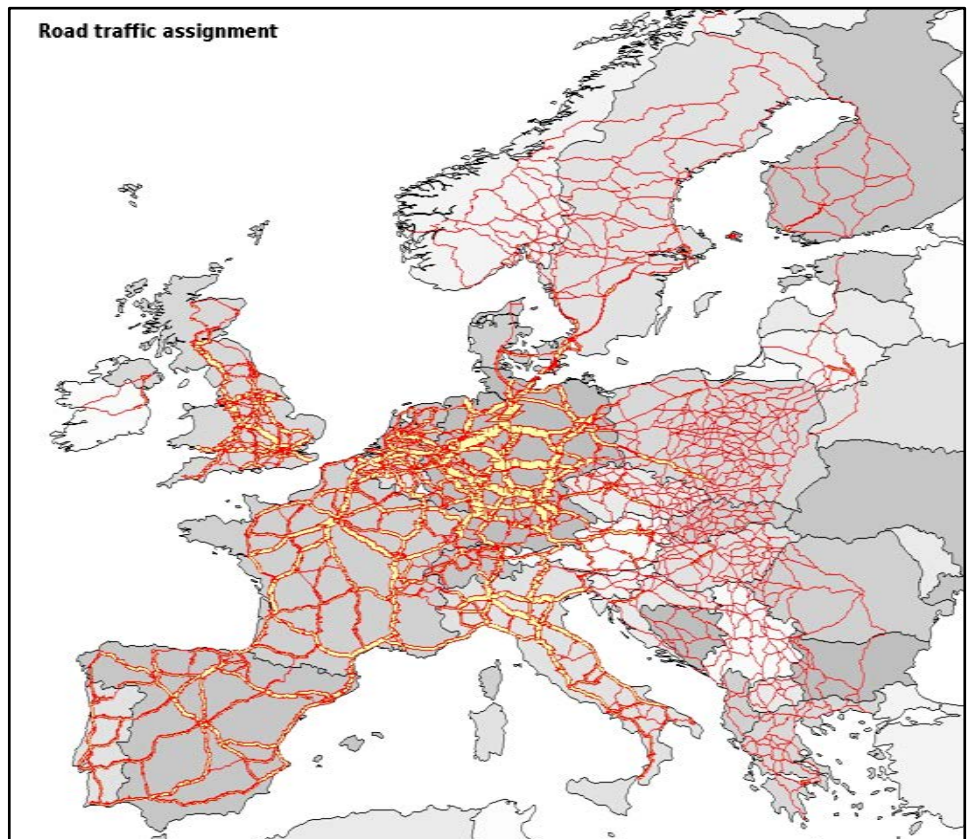
3. Application to the Trans-European Networks

The real-world case presented in this paper covers the European countries. The case was chosen for its complexity. Indeed, beside its large size (the network contains about 68,000 road links, 40,000 railway links and 1200 waterway links), the coverage of the networks is also heavily heterogeneous, as all the transportation modes cannot be used between all origins and destinations. As illustrated by **Figure 1**, it is particularly the case for inland waterways transport, which is mainly available in the north of France, the Benelux countries, the Rhine, Danube and Rhône rivers.

The demand data is provided by Panteia-NEA (Nea *et al.* [10]) for the year 2005. This dataset contains OD matrixes, at the NUTS2 European regional level (251 regions) and for 10 groups of commodities (the classical NST-R chapters 0 - 9). These matrixes are available for road, rail and inland waterway (IWW) transport. Obviously, these ten 251×251 matrixes contain OD relations between which no flow is observed for some modes and/or groups of commodities. **Table 1** gives a global idea of the content of the matrixes.

The first three maps presented in **Figure 1** illustrate the flows relative to each mode. They were obtained assigning the demand of a mode to its own network. The last map

¹In order to avoid an infinite set of possible combinations, one mode is considered as the reference mode, and its initial α_i remains unchanged.



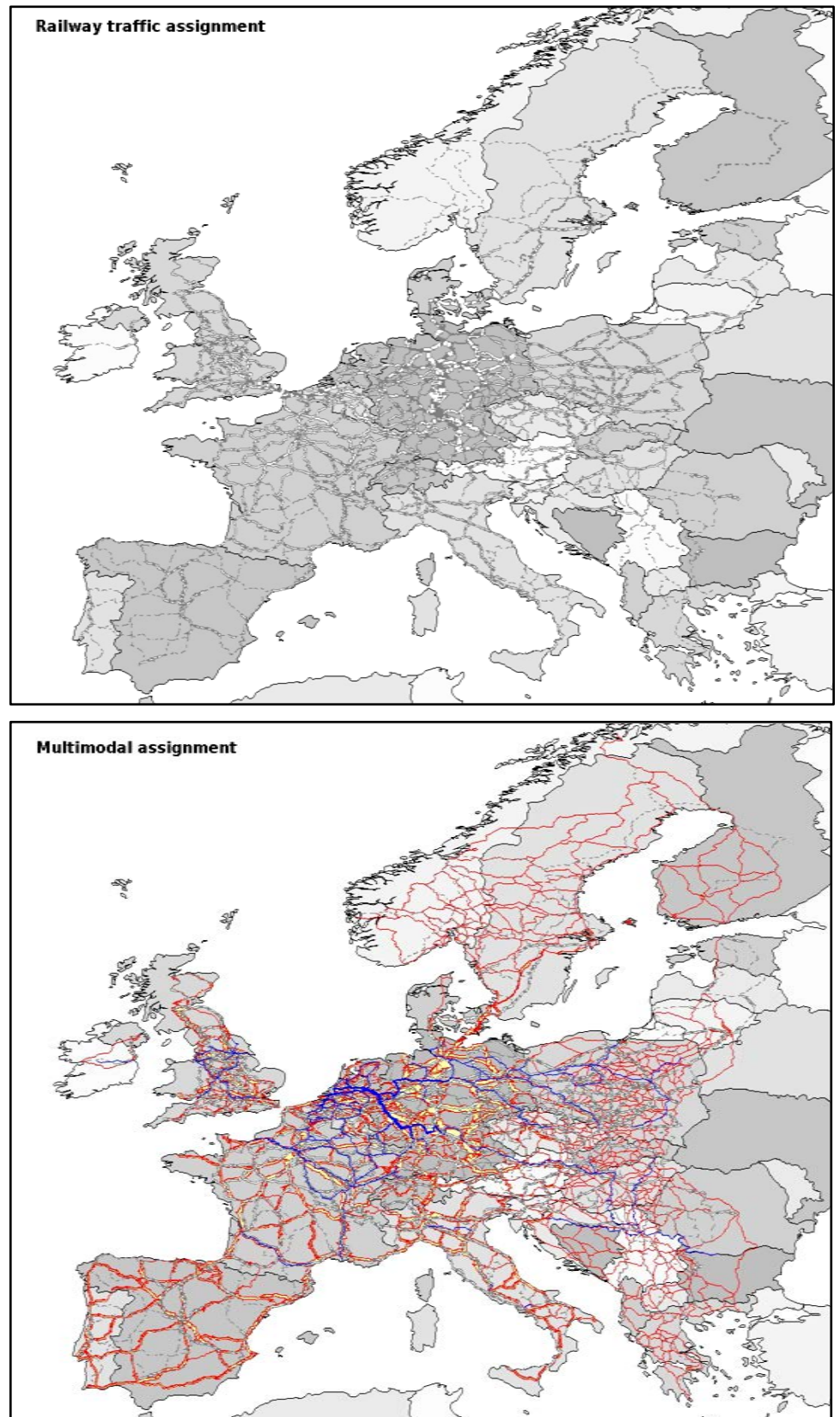


Figure 1. Road, IWW, rail and multimodal assignments.

Table 1. Modal OD matrixes.

	OD relations	Total qty (10 ³ t)	Average qty (t)	Median qty (t)	Market share
Road	124,966	5,048,257	40,397	6351	85.8%
Rail	45,681	532,517	11,657	3381	9.0%
IWW	13,245	314,137	23,717	4697	5.2%

is the result of a multimodal assignment using a single OD matrix, corresponding to the merge of the three modal matrixes. The resulting multimodal assignments vary with the applied modal choice model, and the objective of the exercise is to identify which modal choice model specification gives the best results when only limited explanatory data is available.

In the context of this paper, the only explanatory variables that are used are those that can be retrieved using a transport network analysis software with detailed costs information. For that task, we used the Nodus network model (Jourquin and Beuthe [11]). The software allows to retrieve, for each OD pair, for each mode and for each NST-R group of commodities, the total transportation cost (including loading, unloading, transit and transshipment costs), the travel duration and the length of the haul. The most recent release of Nodus (release 7.0β, 2016) also allows to separately compute the variable (related to distance) and fixed (handling, transit...) costs for each trip.

Nodus has the particularity to perform modal split and assignment in a single step: once a set of modal routes is computed between an OD relation, the corresponding demand is spread over these paths, using a (calibrated) modal choice model. The figures that are presented in the following tables, used for validation of the proposed models, are all gathered from the outputs of the assignments.

All the tested model specifications use exactly the same dataset, *i.e.* the same demand matrixes, the same networks and the same costs as defined in Beuthe *et al.* [12]. These cost functions are used in the Nodus network model for computing the cheapest route between each OD pair for each mode.

If a route exists for a given mode, the fixed cost FC , the variable cost VC , the length L and the duration T of the trip are retrieved. As a result, for each OD pair and each group of commodities, the dataset contains a record with the following fields: FC_{road} , VC_{road} , L_{road} , T_{road} , Q_{road} , FC_{rail} , VC_{rail} , L_{rail} , T_{rail} , Q_{rail} , FC_{iww} , VC_{iww} , L_{iww} , T_{iww} , Q_{iww} . If no route is found for a mode, the related fields are left empty. The total cost for a mode TC_{mode} on a route can be computed as $FC_{mode} + VC_{mode}$. Altogether, the dataset contains almost 160,000 records.

3.1. Validation at the Aggregated Level

Several combinations of these variables are tested in order to estimate their coefficients in multinomial Logit models, using the “mnLogit” R package (Hasan *et al.* [13]), a faster version of the Logit R package (Croissant [14]). The mnLogit package provides time and memory efficient estimation of multinomial logit models using maximum likelih-

ood method. Numerical optimization is performed by the Newton-Raphson method, using an optimized parallel C++ library to achieve fast computation of Hessian matrices.

Linear additive and logarithmic additive utility functions, following the formulations described in section 2, are tested. The L_{mode} variable is not retained, as it appears to be barely discriminant from mode to mode. Actually, it is highly correlated with the TC_{mode} variable and even perfectly correlated with VC_{mode} as the variable costs are defined by unit of length (km). Thus, the following additive combinations of explanatory variables are tested: TC_{mode} , $TC_{mode} + T_{mode}$ and $FC_{mode} + VC_{mode} + T_{mode}$. The log-linear additive combinations of the same variables are also tested. For each model specification, the coefficients are separately estimated for each group of commodities. Road transport is considered as the reference mode, the other two modes being estimated relatively to trucking². The estimated parameters of all these specifications are used to perform multimodal assignments in Nodus, which results are summarized in **Table 2**. The last three columns give the estimated global modal market shares.

As expected, models (1), (2) and (7), which don't estimate an intercept, give the poorest results. All the others estimated modal shares are close to the observed ones.

Next to these models estimated using the "mnLogit" R package, the "benchmark" iterative procedure explained in section 2 is also applied. **Table 3** describes the two specifications that were tested. Model (12) is a classical univariate multinomial logit, while model (13) is the law of Abraham.

It is important to note that, despite the similarity of these specifications with those of the univariate logit models (2) and (7), the nature of the α_i coefficients are different:

- The multiplicative parameter α_i estimated by the iterative procedure is not equivalent to the α_i parameters of the econometric models as the objective functions of both approaches are different.
- The estimation of the logarithmic, *aka* Abraham, model (13) is also different. While the econometric model (7) provides estimations for n_p , the iterative procedure estimates α_p , while the values of n must be given. This will be discussed later.

The objective function of the iterative procedure being the convergence of the estimated global market shares to the observed ones, the estimations are, by nature, very close to the observed ones.

If one except the models without an estimated intercept, *i.e.* (1) (2) and (7), which will be abandoned from now, all the tested models properly estimate the global modal shares. However, these figures are average values, and correct global market shares could very well be estimated while, at the per OD level, huge estimation errors are observed.

3.2. Validation at the OD Relation Level

The estimated coefficients for each commodity are applied to the explanatory variables of every OD relation of the multimodal multi-commodity demand matrix in order to

²The number of estimated coefficient is too large to be published here, but can be provided on request.

Table 2. Estimated market shares for the max likelihood linear and log-linear models.

ID	Var(s)	Cond.	Interc.	Utility V_i (target)	Market share		
					Road (85.8%)	Rail (9.0%)	IWW (5.2%)
1	TC	•		$-\alpha TC_i$	45.9%	32.3%	21.8%
2	TC			$-\alpha_i TC_i$	80.8%	13.1%	6.1%
3	TC	•	•	$-\alpha TC_i + \delta_i$	84.3%	10.5%	5.2%
4	TC		•	$-\alpha_i TC_i + \delta_i$	84.0%	10.8%	5.2%
5	TC, T		•	$-\alpha_i TC_i - \beta T_i + \delta_i$	84.5%	10.3%	5.2%
6	FC, VC, T		•	$-\alpha_i FC_i - \gamma_i VC_i - \beta T_i + \delta_i$	84.6%	10.2%	5.2%
7	TC			$-n_i \ln(TC_i)$	83.7	11.0%	5.3%
8	TC	•	•	$-n \ln(TC_i) + \delta_i$	84.2%	10.6%	5.2%
9	TC		•	$-n_i \ln(TC_i) + \delta_i$	83.8%	11.0%	5.2%
10	TC, T		•	$-n_i \ln(TC_i) - p_i \ln(T_i) + \ln \delta_i$	84.1%	10.7%	5.2%
11	FC, VC, T		•	$-n_i \ln(FC_i) - m_i \ln(TC_i) - p_i \ln(T_i) + \ln \delta_i$	85.4%	9.4%	5.2%

Table 3. Iterative models and estimated market shares.

ID	Modal choice specification	Market shares		
		Road	Rail	IWW
12	$\frac{\exp(\alpha_i TC_i)}{\sum_{j=1}^k \exp(\alpha_j TC_j)}$	85.9%	8.9%	5.2%
13	$\frac{\alpha_i TC_i^{-n}}{\sum_{j=1}^k \alpha_j TC_j^{-n}}$	85.8%	9.0%	5.2%

spread the total demand (tons) of each OD cell over the available transportation modes. As a result, a set of more than 480 000 records is generated, each one containing

- The origin o of the flow (NUTS2 region),
- The destination d of the flow (NUTS2 region),
- The group g of commodities (NST-R),
- The mode m ,
- The observed quantity for mode m between o and d for group g ,
- The estimated quantity for mode m between o and d for group g .

Beside the classical R^2 , a few other accuracy measures are used. An interesting overview of the common used (forecast) error measurements can be found in Shcherbakov *et al.* [15]. In the present case, we must take into account that some data corresponds to a zero market share, as all the transportation modes are not (or cannot) be used between all OD pairs. Thus, most statistical measures that involve ratios can lead to divisions by zero and are therefore useless in our context. Equations (15) to (18) are re-

tained from this review: two absolute forecasting errors (the Mean Absolute Error (MAE) and the Median Absolute Error (MdAE)), as well as two measures based on squared errors, the Mean Square Error (MSE) and the Root Mean Square Error (RMSE).

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |e_i| = \text{mean}_{i=1,n} |e_i| \quad (15)$$

where e_i is the error (observed modal share minus estimated one) with respect to observation i among n observations. Similarly,

$$\text{MdAE} = \text{median}_{i=1,n} |e_i|. \quad (16)$$

RMSE is often preferred to MSE, as it has a similar scale as the one of MAE and MdAE.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (e_i^2) = \text{mean}_{i=1,n} (e_i^2) \quad (17)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i^2)} = \sqrt{\text{mean}_{i=1,n} (e_i^2)}. \quad (18)$$

We choose to keep MAE, MdAE and RMSE plus the classical R^2 . In order to keep comparable values among modes, the first three indicators are computed using the observed and estimated market shares instead of the observed and estimated quantities. Hence, all the figures are in the same $[0, 1]$ scale.

The iterative model 13 (Abraham) was run with different values of n in order to identify the “best” ones. **Table 4** gives the MAE values obtained for a series of values of n , for each group of commodities. The bold values identify the lowest MAE for each NST-R, giving the corresponding “optimal” value of n . Although values of n set to 8 or 10 are often cited in the literature (Bonnell [7]), all the values of n that are higher than 6 result in a higher MAE for all groups of commodities. It is interesting to note that, for five out of the ten groups of commodities, the optimal value of n appears to be 1. In other words, for these groups, the market shares estimated for each alternative are simply proportional to $-TC_r$.

Table 5 gives the performance indicators of the tested models. Conditional models (3) and (8) perform less well and will not be discussed further in this paper.

The values of the performance indicators also vary from mode to mode, as illustrated in **Table 6**. Altogether, all the retained specifications produce comparable performances. The modal share for railway transport appears to be systematically less well estimated.

The two iterative methods (12) and (13), used for benchmarking purposes and which performed best at the aggregated level presented in the previous subsection are now clearly outweighed by the others. Starting from **Table 5** the values of these “benchmark” iterative procedures are presented in italic. As illustrated by the bold figures in **Table 6**, the utility functions of models (6, linear) and (11, log-linear) give the best results, but none is better than the other on all indicators.

Table 4. MAE values for each NST-R/n combination.

n	1	2	3	4	5	6
NST-R 0	0.195	0.201	0.210	0.218	0.225	0.230
NST-R 1	0.064	0.042	0.024	0.018	0.018	0.019
NST-R 2	0.334	0.332	0.331	0.332	0.333	0.334
NST-R 3	0.313	0.317	0.321	0.325	0.328	0.330
NST-R 4	0.338	0.337	0.337	0.338	0.340	0.342
NST-R 5	0.280	0.287	0.294	0.299	0.304	0.308
NST-R 6	0.165	0.173	0.183	0.193	0.203	0.212
NST-R 7	0.285	0.288	0.291	0.294	0.298	0.301
NST-R 8	0.130	0.088	0.069	0.067	0.068	0.070
NST-R 9	0.099	0.064	0.036	0.028	0.027	0.028

Table 5. Global validation at the OD level.

Model	ID	R ²	MAE	MdAE	RMSE
Linear	3	0.909	0.193	0.032	0.340
	4	0.905	0.185	0.014	0.344
	5	0.918	0.176	0.016	0.332
	6	0.932	0.180	0.018	0.338
Log	8	0.908	0.193	0.030	0.339
	9	0.909	0.189	0.015	0.345
	10	0.913	0.182	0.012	0.338
	11	0.934	0.169	0.015	0.319
Iter	12	<i>0.824</i>	<i>0.237</i>	<i>0.000</i>	<i>0.452</i>
	13	<i>0.915</i>	<i>0.201</i>	<i>0.083</i>	<i>0.336</i>

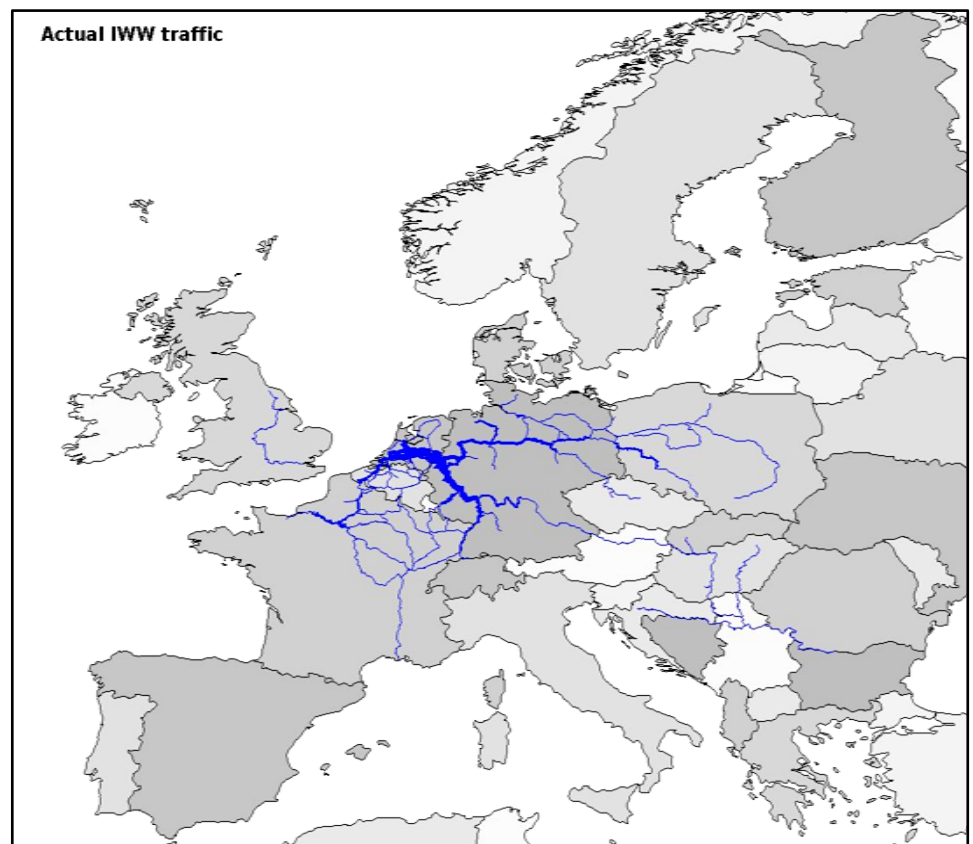
Table 6. Validation at the OD level per mode.

	ID	R ²			MAE			MdAE			RMSE		
		Road	Rail	IWW	Road	Rail	IWW	Road	Rail	IWW	Road	Rail	IWW
Linear	4	0.949	0.234	0.573	0.258	0.245	0.051	0.095	0.069	0.000	0.411	0.406	0.145
	5	0.956	0.254	0.636	0.244	0.234	0.049	0.090	0.064	0.000	0.396	0.391	0.142
	6	0.963	0.306	0.674	0.252	0.242	0.047	0.101	0.073	0.000	0.404	0.399	0.140
Log	9	0.951	0.251	0.611	0.265	0.253	0.050	0.111	0.079	0.000	0.413	0.409	0.142
	10	0.953	0.259	0.618	0.255	0.243	0.048	0.106	0.072	0.000	0.404	0.399	0.140
	11	0.965	0.321	0.676	0.235	0.225	0.048	0.096	0.063	0.000	0.380	0.374	0.141
Iter	12	<i>0.899</i>	<i>0.047</i>	<i>0.415</i>	<i>0.380</i>	<i>0.319</i>	<i>0.047</i>	<i>0.038</i>	<i>0.014</i>	<i>0.000</i>	<i>0.447</i>	<i>0.530</i>	<i>0.182</i>
	13	<i>0.955</i>	<i>0.199</i>	<i>0.519</i>	<i>0.285</i>	<i>0.265</i>	<i>0.053</i>	<i>0.176</i>	<i>0.140</i>	<i>0.000</i>	<i>0.401</i>	<i>0.394</i>	<i>0.146</i>

3.3. Validation at the Link Level of the Networks

In the two previous subsections, the validation of the models is performed at the demand (node) level, but the network topologies are completely ignored. As no (observed) count data is available along the segments of the networks for such a large area. One could consider the modal assignments presented in **Figure 1** as some kind of reference. Indeed, the best modal choice model specifications would result in multimodal assignments for which the flow on each modal network would closely correspond to those illustrated by the first three maps of **Figure 1**. However, the modal choice models are unable to perfectly predict the modal split for each OD relation, and this impacts the results of the assignments, as illustrated by **Figure 2**. In this figure, the first map (identical to the one represented in **Figure 1**) corresponds to the “actual” flow on the IWW network. The second one represents the flow resulting from a multimodal assignment, using modal choice model (6). At a first glance, both are identical. However, when both flows are compared (last map of **Figure 2**), it comes out that differences exist, particularly on the North Rhine River.

It is thus interesting to compare the flow assigned to each link of the networks by the multimodal assignment procedure to the values retrieved from the three modal assignments. Therefore, “counting points” are placed on all the segments of the network that are connected to at least two other segments. The flows are separately measured for both directions. This represents about 27,500 counting points on the road network,



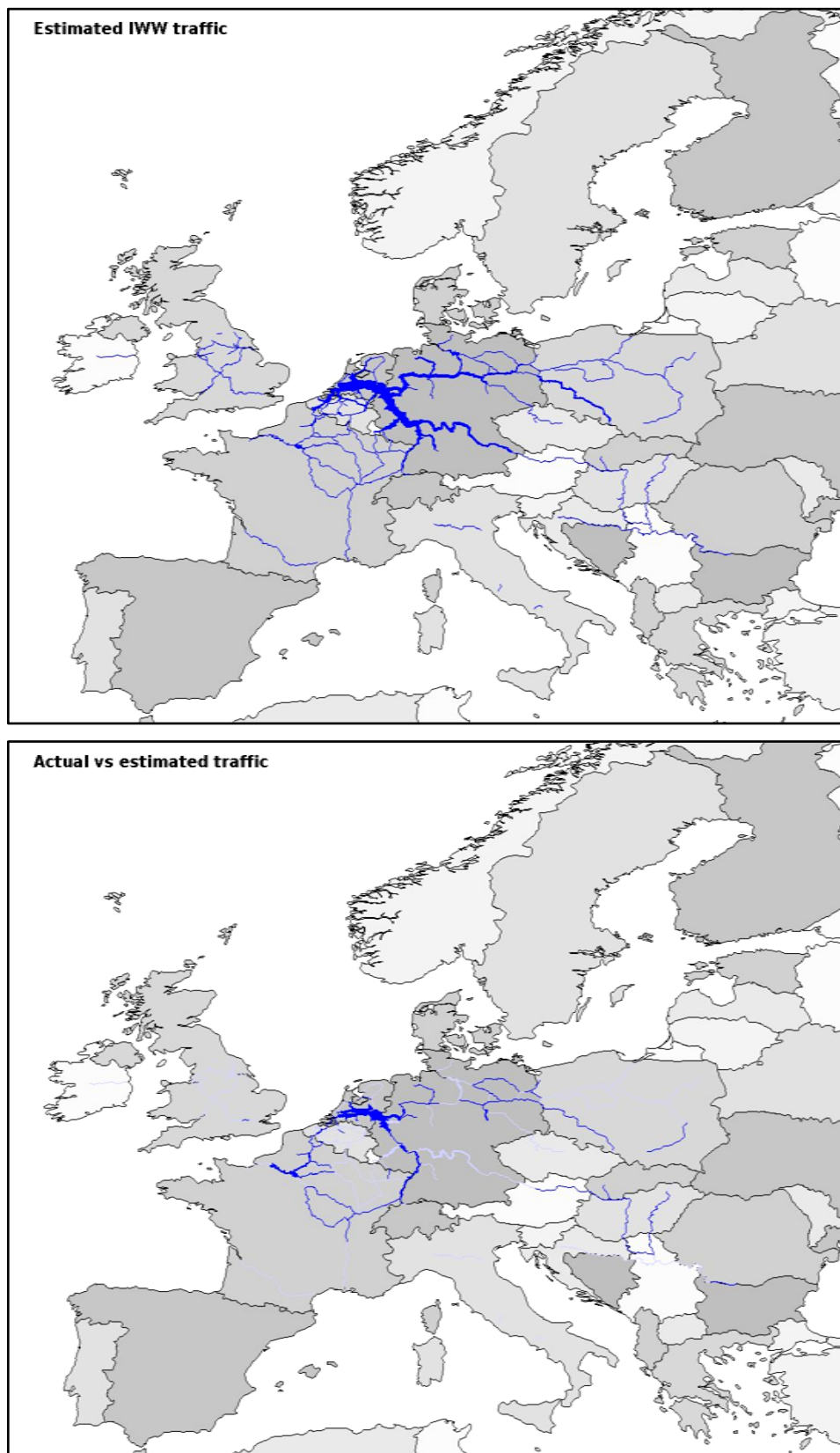


Figure 2. Difference between actual and estimated flows on the IWW network.

17,200 on the railways network and 850 for the inland waterways. **Table 7** presents the correlation between the “actual” and estimated flows measured at the counting points. Again, the two iterative procedures used for benchmarking appear to be clearly less performant. It is especially the case for rail transport for which the logit formulation gives very poor results ($R^2 = 0.061$). Here also, there is no big difference between the linear and log-linear models and the modal choice models perform less well for railway transport, with a maximum R^2 of 0.769.

3.4. Summary of Main Results

If the global performances of the models could be considered as acceptable at the three levels of aggregation (see **Table 2** for the aggregated level, **Table 5** for the OD level and column “All” of **Table 7** for the link level), the quality of the results largely varies from mode to mode. This is illustrated by **Table 8**, which clearly shows that the use of rail

Table 7. Validation at the flow per link level.

	ID	R ²			
		All	Road	Rail	IWW
Linear	4	0.960	0.980	0.629	0.914
	5	0.970	0.985	0.736	0.939
	6	0.975	0.987	0.772	0.953
Log	9	0.958	0.979	0.607	0.924
	10	0.964	0.982	0.655	0.938
	11	0.974	0.987	0.769	0.951
Iter	12	<i>0.891</i>	<i>0.953</i>	<i>0.061</i>	<i>0.863</i>
	13	<i>0.945</i>	<i>0.979</i>	<i>0.383</i>	<i>0.898</i>

Table 8. Summary of the performances at the three levels of aggregation.

Attributes	Aggregated shares			Disaggregated (OD)			Disaggregated (flow)			
	Market shares			R ²			R ²			
	Road	Rail	IWW	Road	Rail	IWW	Road	Rail	IWW	
Linear	$TC_p \delta_i$	84.0%	10.8%	5.2%	0.949	0.234	0.573	0.980	0.629	0.914
	$TC_p T_p \delta_i$	84.5%	10.3%	5.2%	0.956	0.254	0.636	0.985	0.736	0.939
	$FC_p VC_p T_p \delta_i$	84.6%	10.2%	5.2%	0.963	0.306	0.674	0.987	0.772	0.953
Log	$TC_p \delta_i$	83.8%	11.0%	5.2%	0.951	0.251	0.611	0.979	0.607	0.924
	$TC_p T_p \delta_i$	84.1%	10.7%	5.2%	0.953	0.259	0.618	0.982	0.655	0.938
	$FC_p VC_p T_p \delta_i$	85.4%	9.4%	5.2%	0.965	0.321	0.676	0.987	0.769	0.951
Iter.	$TC_i \text{ logit}$	<i>85.9%</i>	<i>8.9%</i>	<i>5.2%</i>	<i>0.899</i>	<i>0.047</i>	<i>0.415</i>	<i>0.953</i>	<i>0.061</i>	<i>0.863</i>
	$TC_i \text{ Abraham}$	<i>85.8%</i>	<i>9.0%</i>	<i>5.2%</i>	<i>0.955</i>	<i>0.199</i>	<i>0.519</i>	<i>0.979</i>	<i>0.383</i>	<i>0.898</i>
	(actual)	(85.8%)	(9.0%)	(5.2%)						

transport is the most difficult to predict. This is probably due to the fact that the railway networks are available almost everywhere in Europe, and that competition with road transport is very often observed. Therefore, additional attributes would be needed in the modal choice utility functions.

4. Conclusions and Perspectives

Strategic multimodal freight transportation network models are regularly used in the framework of important transport policy decisions to be taken, but are often setup using coarse origin-destination matrixes, defined at a regional level. In addition, only a few modal choice attributes are mostly available. Unfortunately, if numerous examples of such models can be found in the scientific literature, the published papers are almost silent on their validation.

Validation is precisely the main objective of this paper, which examines, in a systematic way, several modal choice models specifications that make only use of explanatory variables available at the network level (mainly costs and durations per mode/route). The parameters for a series of linear and log-linear multinomial logit specifications are estimated using R, and the results obtained by an alternative iterative estimation procedure are used as a lower bound benchmark.

The validation of the models is performed at three levels of aggregation:

- A highly aggregated level, for which the global shares for each transportation mode are estimated and compared to the observed ones;
- A first disaggregated level, from a “node” viewpoint, for which the estimated and actual shares are compared for each origin-destination pair;
- A second disaggregated level, from a “link” perspective, which compares the assigned flows to the actual ones on a per segment basis.

A multimodal freight network model, covering the whole European area, is presented. Inland waterways, rail and road transport compete where possible, and 10 categories of commodities are embedded in the origin-destination matrixes.

If the global performances of the models could be considered as acceptable, the quality of the results largely varies from mode to mode, the use of rail transport being the most difficult to predict.

But there is no miracle: having limited information clearly impacts the quality of the models.

Nowadays, most transport network models are embedded in geographic information systems, which can also give some information about the areas where demand and supply are located, such as the number of inhabitants, employment, gross regional product, accessibility to the modal networks... These variables can partially explain the attractiveness between the OD pairs and could be used, in some future work, as additional explanatory variables into the utility functions.

As pointed out by de Jong *et al.* [1], it would also be interesting to go a step further in the validation process, performing some “backcasting” (use the model to calculate a “forecast” for a year in the past) or validity testing (asking industry experts and regional

planners for their opinion on whether the model behavior and the model results look reasonable). Unfortunately, and especially for models covering international areas, there is a glaring lack of data and available expertise covering all the countries included in large strategic models.

Acknowledgements

I gratefully acknowledge the support of Michel Beuthe of the Université catholique de Louvain in Mons. Our numerous discussions on the way to present a structured list of utility function specifications, including those related to the “Law of Abraham” were stimulating.

This work has received financial support from the research fund of the Université catholique de Louvain in Mons.

References

- [1] de Jong, G., Tavasszy, L., Bates, J., Grønland, S.E., Huber, S., Kleven, O., Lange, P., Ottemöller, O. and Schmorak, N. (2016) The Issues in Modelling Freight Transport at the National Level. *Case Studies on Transport Policy*, **4**, 13-21.
<http://dx.doi.org/10.1016/j.cstp.2015.08.002>
- [2] Zhang, M. (2013) A Freight Transport Model for Integrated Network, Service, and Policy Design. PhD Thesis, TU Delft.
- [3] Jourquin, B. (2005) A Multi-Flow Multi-Modal Assignment Procedure on Large Freight Transportation Networks. *Studies in Regional Science*, **35**, 929-946.
<http://dx.doi.org/10.2457/srs.35.929>
- [4] Yamada, T., Russ, B.F., Castro, J. and Taniguchi, E. (2009) Designing Multimodal Freight Transport Networks: A Heuristic Approach and Applications. *Transportation Science*, **43**, 129-143. <http://dx.doi.org/10.1287/trsc.1080.0250>
- [5] Ben-Akiva, M. and Bierlaire, M. (1999) Discrete Choice Methods and Their Applications to Short Term Travel Decisions. In: Hall, M., Ed., *Handbook of Transportation Science*, Kluwer Academic Publishers, USA, 5-33.
- [6] Abraham, C. and Coquand, R. (1961) La répartition du trafic entre itineraries concurrents. *Revue Générale des Routes et Aérodrômes*, **357**, 57-76.
- [7] Bonnel, P. (2001) Prévision de la demande transport. Rapport présenté en vue de l'obtention du diplôme d'habilitation à diriger des recherches, ENTPE, Université Lumière Lyon 2.
- [8] Leurent, F. (1995) Constitution d'une boîte-à-outils pour opérer diverses affectations statiques du trafic routier. Note de travail 95.2, INRETS, Arcueil.
- [9] Gaudry, M., Briand, A., Paulmyer, I. and Tran, C.-L. (2006) Choix modal transpyrénéen ferroviaire, intermodal et routier: Un modèle Logit Universel de forme Box-Cox Généralisée. Working paper DEST n° 7, Institut National de la Recherche sur les Transports et leur Sécurité, INRETS, France.
- [10] NEA, CE Delft, Planco, MDS Transmodal and Via Donau (2011) Medium and Long-Term Perspectives of IWT in the European Union. European Commission, DG-MOVE.
http://www.ce.nl/art/uploads/file/4330_IWT_EU_main_report.pdf
- [11] Jourquin, B. and Beuthe, M. (1996) Transportation Policy Analysis with a GIS: The Virtual Network of Freight Transportation in Europe. *Transportation Research C*, **4**, 359-371.
[http://dx.doi.org/10.1016/S0968-090X\(96\)00019-8](http://dx.doi.org/10.1016/S0968-090X(96)00019-8)

- [12] Beuthe, M., Jourquin, B., Urbain, N., Bruinsma, F., Lingemann, I., Ubbels, B. and Van Heumen, E. (2012) Estimating the Impacts of Water Depth and New Infrastructures on Transport by Inland Navigation: A Multi Modal Approach for the Rhine Corridor. *Procedia—Social and Behavioral Sciences*, **54**, 387-401.
<http://dx.doi.org/10.1016/j.sbspro.2012.09.758>
- [13] Hasan, A., Zhiyu, W. and Mahani, A.S. (2015) mnlogit: Multinomial Logit Model. R Package Version 1.2.4. <https://cran.r-project.org/web/packages/mnlogit>
- [14] Croissant, Y. (2013) mlogit: Multinomial Logit Model. R Package Version 0.2-4.
<http://CRAN.R-project.org/package=mlogit>
- [15] Shcherbakov, M., Brebels, A., Shcherbakova, N., Tyukov, A., Janovsky, T. and Kamaev, V. (2013) A Survey of Forecast Error Measures. *World Applied Sciences Journal (Information Technologies in Modern Industry, Education & Society)*, **24**, 171-176.



Scientific Research Publishing

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact jts@scirp.org

