TØI report
699/2004

Construction of basis OD matrices for 1999 freight flows between municipalities in Norway

Arild Vold
Inger Beate Hovi

ISSN 0802-0175
ISBN 82-480-0405-8

Oslo, January 2004

Title: Construction of basis OD matrices for 1999 freight flows between municipalities in Norway

Author(s): *Arild Vold; Inger Beate Hovi*

TØI report 699/2004
Oslo, 2004-01
18 pages
ISBN 82-480-0405-8
ISSN 0802-0175

Financed by:
Public roads Administration

Project: 2857 A more precise NEMO

Project manager: Inger Beate Hovi

Quality manager: Kjell Werner Johansen

Key words:
Freight counts; industry statistics; logistic chains

Summary:
The objective of this paper is to describe how data from production and commodity statistics was used to construct basis OD matrices for freight transport between all pairs of municipalities. OD matrices were constructed for 11 commodity groups that together cover all freight transport along road, by train and by boat. A model was built where the basic principle is to balance data for transport in, out and within the municipalities such that the tonnes sent equals the tonnes received (i.e., mass balance). Doubly constrained gravity models were implemented with the balanced data and additional count statistic data to put constraints on the OD pattern at 15x15 aggregate zoning level. The gravity model was applied and the balanced OD matrices were subsequently evaluation by comparison with OD patterns at 39x39 aggregate zoning level based on the count statistics.

Tittel: Etablering av basis OD matriser for godsstrømmer mellom kommuner i Norge i 1999

Forfatter(e) *Arild Vold; Inger Beate Hovi*

TØI rapport 699/2004
Oslo: 2004-01
18 sider
ISBN 82-480-0405-8
ISSN 0802-0175

Finansieringskilde:
Statens vegvesen Vegdirektoratet

Prosjekt: 2857 En mer presis NEMO

Prosjektleder: Inger Beate Hovi

Kvalitetsansvarlig: Kjell Werner Johansen

Emneord:
Godstelling; næringsstatistikk; logistikkjeder

Sammendrag:
Målsetningen med dette dokumentet er å beskrive hvordan data fra produksjon og varehandelsstatistikk ble brukt for å konstruere basis OD matriser for godstransport mellom alle par av Norske kommuner. OD matriser ble for 11 varegrupper som samlet dekker all godstransport transport langs vei, med tog og båt. Vi har bygget en modell der det basale prinsippet er å balansere data for transport inn, ut og innen kommunene slik at sendte godsstrømmer (tonn) blir lik mottatte strømmer (dvs. massebalanse). En gravitasjonsmodell med doble føringer ble implementert med balanserte data og der ekstra statistikk basert på utvalgstellinger ble brukt for å legge føringer på OD mønsteret for et 15x15 soners aggregeringsnivå. Gravitasjonsmodellen ble anvendt og de balanserte OD matrisene ble deretter evaluert ved sammenligning med tellingene på 39x39 soners aggregeringsnivå.

Language of report: English

The report can be ordered from:
Institute of Transport Economics, the library
P O Box 6110 Etterstad N-0602 Oslo, Norway
Telephone +47 22 57 38 00 Telefax +47 22 57 02 90
Price NOK 150

Rapporten kan bestilles fra:
Transportøkonomisk institutt, biblioteket
Postboks 6110 Etterstad, 0602 Oslo
Telefon 22 57 38 00 - Telefaks 22 57 02 90
Pris 150

Preface

TØI accomplished an internal project autumn 2003 as part of the work to maintain and develop NEMO (Network model for freight transport within Norway and between Norway and foreign countries, Version 2, Vold et al. 2002) with Inger Beate Hovi as the project leader. This report describes the work performed to improve the OD matrix for domestic freight transport between municipalities in Norway. This includes description of the models and use of data to construct the OD matrices and the practical assumptions required to calibrate and evaluate the models and resulting OD matrices, respectively.

The two authors jointly developed the principles of the logistic pathway model for construction of basis OD matrices on the basis of the work already performed by Vold et al. (2002). Arild Vold formalised and typed the model in mathematical terms and Inger Beate processed raw data to provide processed data for freight transport flows in and out of the Norwegian municipalities. Arild Vold specified, calibrated and evaluated the three-dimensional gravity model and was the main responsible for writing this report.

We are grateful to Jardar Andersen who performed calibration of parameters in freight link and transfer cost functions on the basis of output from the three-dimensional gravity model, and Viggo Jean-Hansen for good discussions. We would also like to thank Head of Department Kjell Werner Johansen, who has been responsible for quality assurance and Laila Aastorp Andersen who were responsible for the final layout of the report.

Oslo, January 2004
Institute of Transport Economics

Sønneve Ølnes	*Kjell Werner Johansen*
Acting Managing Director	Head of Department

Contents

TØI report 699/2004
Authors: Arild Vold
Inger Beate Hovi
Oslo 2004, 18 pages English language

Summary:

Construction of basis OD matrices for 1999 freight flows between municipalities in Norway

Basis origin-destination (OD) matrices for freight transport constitute a convenient grouping of information that can be used directly in statistical analysis and as input in freight transport models. The objective of this report is to describe how data from production and commodity statistics for the 435 Norwegian municipalities and the Norwegian foreign trade statistics was used to construct basis OD matrices for freight transport between all pairs of municipalities. OD matrices were constructed for 11 commodity groups that together cover all freight transport along road, by train and by boat. The main challenge of this task was to construct a model that can be used with these statistics for representation of the logistic pathway. A model was built where the basic principle is to balance data for transport in, out and within the municipalities such that the tonnes sent equals the tonnes received (i.e., mass balance). Doubly constrained gravity models were implemented with the balanced data and additional count statistic data to put constraints on the OD pattern at 15x15 aggregate zoning level. The gravity model was applied for generation of OD-patterns for the commodity groups, where the cost elements of the gravity models were adjusted to improve the resulting transport distance distribution in coherence with the counts statistics data. The balanced OD matrices were subsequently evaluation by comparison with OD patterns at 39x39 aggregate zoning level based on the count statistics.

The report can be ordered from:
Institute of Transport Economics, PO Box 6110 Etterstad, N-0602 Oslo, Norway
Telephone: +47 22 57 38 00   Telefax: +47 22 57 02 90

1 Introduction

Basis OD matrices for freight transport constitute a convenient grouping of information that can be used directly in statistical analysis and as input in freight transport models. There are two major data sources that we consider suitable as a foundation for construction of basis OD matrix for domestic transport in Norway. One source is the national count statistics for freight transport by boat, train and truck (Statistics Norway). These counts[1] statistics represents transport within and between all Norwegian municipalities represented in tonnes per commodity group (grouped according to the NSTR catalogue). The other sources are register statistics for national production and national commodity trade in the 435 Norwegian municipalities plus the Norwegian foreign trade statistics (Statistic Norway). The statistics can be applied to determine the total purchase and sale (Norwegian kroner) between the commercial links in each municipality for each commodity group.

The counts statistics are mode specific and include information of the actual pattern of the transport. A drawback with the counts is that we do not know where the seller and buyer are located. This is because the mode specific transports can be just a part of the transport chain where earlier and later transport previous to load and after reloads may take place.

The register statistic includes the total amount of goods sold and purchased within business links in the municipalities, but no information of where the goods are transported.

The objective of this report is to describe how the register- and count statistics were processed and combined for construction of basis freight transport OD matrices. OD matrices for 11 commodity groups were constructed that together covers all freight transport along road, by train and by boat. The main challenge of this task was to construct a model that can be used with these statistics for representation of the logistic pathway. Basic principles of the model are to balance data for transport in, out and within the municipalities such that the tonnes sent equals the tonnes received (i.e., mass balance). A three-dimensional balancing method was applied for generation of OD-patterns for the 11 commodity groups, with balanced data and additional count statistic data to put constraints on the OD pattern at 15x15 aggregate zoning level. The method includes a doubly constrained gravity model with cost elements that are adjusted to improve the resulting transport distance distribution in coherence with the counts statistics data. The balanced OD matrices are subsequently evaluation by comparison with OD patterns at 39x39 aggregate zoning level based on the count statistics.

The overall idea of using the three-dimensional balancing method this way is to let the data for total transport in and out of the municipalities to govern the absolute level of tonnes transported within and between the municipalities, whereas the count matrices are used as the indicator that govern the pattern of the freight transport flows at the aggregated level to ensure that the pattern of the balanced OD matrices resemble the pattern of the counts.

[1] The data for counts constitute OD matrices for each commodity for transport within and between municipalities in Norway, based on counts for the three modes: rail, boat and truck.

Chapter 2 describe the logistic pathway model, which was established to exploit and synthesise the register and count statistic data while capturing the underlying driving forces of freight transport. The three-dimensional balancing procedure is also described. Register and count data are aggregated to 11 commodity groups, and register data are organised according to trade links, and processed to ensure mass balance. Chapter 3 explains how the three-dimensional balancing procedure was calibrated and applied to generate OD-matrices for each the 11 commodity groups. This is followed by a simple evaluation of the balanced OD matrices.

The Appendix gives an overview of the steps in the construction of the OD matrices with the aid of the STAN software package, and an overview and description of the STAN macros applied.

2 The logistic pathway model for construction of basis OD matrices

The total commodity flows obtained from the register and count statistics were aggregated into 11 commodity groups:

1. Food
2. Fish (frozen)
3. Thermo
4. Vehicles/machinery
5. General cargo
6. Timber and wood ware
7. Minerals and stone products
8. Chemical products
9. Metals and ore
10. Petroleum products (liquid)
11. Fish (fresh)

The composition of the commodity groups were based on the requirements: (1) that commodity groups represents output from corresponding businesses, which makes them interesting to analyse from model users point of view, (2) that the collection of commodities within each commodity group should have approximately the same requirements for transport quality (and thus transport costs), (3) that available data are sufficient to construct OD-matrices for the commodity groups, and (4) that the shares of the commodity groups that are produced should vary little among the municipalities. We consider the grouping to be convenient for many purposes. Food, fish, thermo (food that require cooling or freezing while transported), and petroleum products are all commodities with special requirements for transport quality. Chemical products and petroleum products are both commodities that are classified as dangerous goods.

2.1 Processing of register data

At the municipality level, empirical register data for transport into municipalities are available for the trade links:

- Industrial input
- Industrial purchase
- Wholesalers purchase
- Direct purchase by retailers

and empirical register data for transport out of municipalities are available for the trade links:

- Industrial production

- Sale from industry
- Direct sale from wholesalers
- Sale from retailers
- Delivery from agriculture, forestry and fishery

These data constitute all logistic links for sale and purchase between producers, wholesalers and retailers, with the exception of data for internal sales and purchase within the trade sector, i.e., *Sale from wholesalers to retailers* and *purchase from wholesalers to retailers*. We represented *sales from wholesalers to retailers* by *purchase to wholesaler* minus *direct sales from wholesalers to industry or households*. *Retailer's purchase from wholesalers* was approximated by *sales from retailers*.

The register data are reported in value (i.e., Norwegian kroner). Thus it is necessary to transform the data to corresponding tonnes transported in order to apply them for construction of basis OD matrices, which was achieved by exploitation of data from the production statistic (Statistic Norway) for the value and tonnes of production by enterprises consisting of single firms (There are about 1750 such firms among a total of 4500 firms, but the value of production from the 1750 firms constitute a much smaller share). Although the emprical foundation is small, we have used these data to assess the national industrial product prices for the 11 commodity groups, with the exception of a few occations where we used the local price per tonne to exploit the data to a greater extent (Vold et al., 2002). Prices on the commoditis when they are used as input were set at 0.7 times the price on the produced commodity in the municipalities. Commodity prices on firm's purchace of commoditities that are not part of their production are set at 1.0 times the national production price, and we derive a commodity price of 1.019 times the national production price for sale of commoditities that are not part of their production. For agriculture, forest and fish production, we have data (Statistics Norway) in tonnes for each commodity. Wholesalers and retailers prices are set at industrial prices, but we have added a fixed factor of 15% on retailers prices if the purchased amount to grossist and retailes (in tonnes) differs in an implausible way.

The assessed total tonnes of commodity transport into a municipality *i* are represented by

$$A_i = (\text{industrial input})_i + (\text{industrial purchase})_i + (\text{wholesalers purchase})_i + (\text{retailers purchase from wholesalers})_i + (\text{to zones abroad})_i + (\text{retailers direct purchase})_i$$

and the transport of a commodity out of a municipality can be represented by

$$B_i = (\text{industrial production})_i + (\text{sale from industry})_i + (\text{direct sale from wholesalers})_i + (\text{sale from wholesalers to retailers})_i + (\text{delivery from agriculture, forestry and fisheries})_i + (\text{from zones abroad})_i$$

A complete and balanced representation of the transport flows should satisfy:

$$\sum_i A_i + \sum_i \left[(\text{residual})_i\right] = \sum_i B_i \ ,$$

where *n = 436* and the residual component represents mainly consumption in the production and service sectors and private consumption not sold by retailers. $\sum_i B_i$ is greater than $\sum_i A_i$ for commodity groups 2, 3, 4, 6, 7, 8, 9 and 10, whereas

$Sum\_in = \sum_i A_i$ is greater than $Sum\_out = \sum_i B_i$ for commodity groups 1 and 5 (see Table 1).

Table 1. The sum of counts data, and in- and out register data for the commodity groups (million tonnes). Domestic trade, only. TØI report 699/2004.

		2+11	3	4	5	6		8	9	10	Sum
Counts	12,8	1,3	9,9	7,4	71,8	12,4	123,4	6,2	6,3	15,3	266,8
Sum_in ($\sum_i A_i$)	6,3	2,6	4,2	3,7	95,1	6,7	3,8	1,7	2,6	14,2	140,9
Sum_out ($\sum_i B_i$)	6,2	4,6	10,6	5,5	74,3	28,9	41,3	6,2	10,9	19,6	208,1

In the cases where $\sum_i B_i > \sum_i A_i$, the residual can be interpreted as *direct purchase from wholesalers for consumption* but also other components on the purchase side. The commodities represented by the *residual link* are transported within the municipalities and from other municipalities. In lack of register data for this trade, we assumed that the transport of commodities represented by the residuals are distributed to the municipalities in proportion to the size of an activity indicator in the municipalities. Thus

$$(residual)_i = \frac{(\sum_j B_j - \sum_j A_j)}{\sum_j IA_j \cdot} \cdot IA_i ,$$

where the activity indicator is $IA_i = (sale\ \ from\ whole\ salers\ to\ retailers)_i +$ $(purchase\ by\ wholesalers)_i$ for all the commodity groups except 9 (see discussion below). An activity indicator based on information about employment by industry related to our commodity groups in different municipalities would probably improve the redistribution. This is currently postponed for future research, however.

In cases where *Sum_in* is greater than *Sum_out*, the *activity indicator* becomes negative, which can be interpreted as an indication that the inputs are overestimated or that the output are underestimated. Currently we don't have a very good interpretation of how to deal with such situations (commodity groups 1 and 5). It is also a bit problematic to interpret the situation where *Sum_out* is very much greater than *Sum_in* (commodity group 6, 7, 8 and 9).

The current way of handling commodity group 1 was to increase (1) industrial production, (2) sale from industry, and (3) delivery from soil, forestry and fishery by 5%.

The current way of handling commodity group 5 was to adjust down by 25% the sales volumes purchased and sold by wholesalers and the volumes purchased by retailers. By adjusting all these three trade links, we maintain the trade balance. A by-effect, is however, that the $Sum\_in = \sum_i A_i$ decreases twice that of $Sum\_out = \sum_i B_i$, which pushes the balance in the right direction. We also increased the (1) industrial production, (2) sale from industry, and (3) delivery from soil, forestry and fishery by 35%, to make *Sum_out* greater than *Sum_inn*.

The problem with commodity group 6 was currently solved by decreasing all production in this group by 50% (i.e., *Sum_out* was decreased from 28.9 mill tonn to 16.8 mill. tonnes). Similarly for commodity group 9, production was decreased by 50%. The residual for commodity group 9 was distribution unlike the residual for the other commodity groups - according to the distribution of industrial input among the municipalities.

With the corrections for commodity groups 1, 5, 6 and 9 the residual is consistently calculated as prescribed above with new figures for *Sum_in* and *Sum_out* (Table 3).

Commodity group 7 is special because a large amount of transports take place without the change of ownership. This is for instance the case for transport of mass from a construction area to a place for deposit. It is assumed that the transports without change of ownership are located within the zones. Hence, to establish balanced OD-matrices with these transports, we simply first apply the residual to balance the data as described above, apply matrix balancing as described below and then exchange the diagonal of the balanced matrix by the diagonal of the counts.
Commodity groups 2, fish, were also treated in a special way, since it was required that this commodity should be divided in fresh and frozen subgroups (commodities 2 and 11). We made this subdivision in the out-marginals A_i according to the export figures in the foreign trade statistics and in the in-marginals B_j according to the total shares of production of fresh and frozen fish distributed from Norwegian municipalities. The division of th etotal in-marginals in fresh and frozen fish are obtained by using the national export shares of fresh and frozen fish.

Finally, it was also recognised a pattern in the register data that doesn't harmonise with expectations for commodity group 10, especially too much transport to Oslo. New data for commodity group 10 was established based on information about quantities of produced petroleum products from the Norwegian Petroleum institute (NP) of 15.55 million tonnes, where 7.61 million tonnes were distributed in the domestic market. The logistic supply chain for petroleum products is directed through 17 main terminals in 17 municipalities. It was necessary to guess which counties the different terminals are serving (Table 2). Thereafter, the domestic data for retailer sale, reported by county, were approximately disaggregated to municipality level by applying information from the commodity trade statistic. The *Sum_in* data for commodity 10 includes: (1) purchase by retailer, (2) purchace by wholesaler (i.e., to main terminal), (3) purchase by industry, (4) export from Norway. The *Sum_out* includes: (1) Refinery production 15.55, (2) sale from wholesaler, (3) sale from industry and (4) import to Norway.

The adjustments of *Sum_in* and *Sum_out* for each commodity group and the residual component resulted in processed register data (Table 3) that satisfy the mass balance that equal amounts are sent and received. The sum of *Sum_out* for all commodity groups (i.e., 188,8 million tonnes) are approximately 50 million tonnes greater than the total national production reported in TØ/515/2001, which is a reasonable since *Sum_out* includes the transport from the wholesalers in addition to the transport from producers.

Table 2. Overview of main terminals and guessed assumptions of which counties they serve. TØI report 699/2004.

Municipality location	

of main terminals	Counties served by main terminal
Fredrikstad	Østfold
Oslo	Akershus, Oslo, Hedmark, Oppland
Tønsberg	Vestfold, Buskerud, Telemark
Kristiansand S	Aust- og Vest-Agder
Stavanger	Rogaland
Haugesund	Rogaland
Bergen	Hordaland
Førde	Sogn og Fjordane
Ålesund	Romsdalen
Kristiansund N	Nord-Møre
Trondheim	Sør og Nord-Trøndelag
Mo i Rana	Nordland
Bodø	Nordland
Harstad	Troms
Tromsø	Troms
Hammerfest	Finnmark
Kirkenes	Finnmark

Table 3. The sum of counts data, and processed in- and out registers data (million tonnes). Domestic trade, only. TØI report 699/2004.

	1	2	3	4	5	6	7	8	9	10	11	Sum
Counts	12,8	2,7	9,9	7,4	71,8	12,4	123,4	6,2	6,3	15,3	0,52	269,1
Sum In	6,3	1,56	4,2	3,7	72,6	6,7	3,8	2,4	2,6	11,1	1,02	115,9
Sum Out	6,3	2,95	10,6	5,5	72,8	16,8	41,2	5,4	6,3	19,3	1,69	188,8

2.2 Three-dimensional matrix balancing

The elements in the basis OD matrices shall represent the transport from seller to purchaser.

The processed register data represents the flow of traded commodities in and out of the business links in the municipalities. They include no information, however, of the quantity of trade between pairs of the municipalities. Thus a balancing method is needed to determine an OD pattern for the trade between pairs of municipalities. A basic doubly constrained gravity model gives an OD pattern for the 11 commodity groups that deviates largely from the pattern in the count data and also largely from independent knowledge of the freight flows.

To improve the reliability of the OD pattern, it is necessary to extend the basic doubly constrained gravity model with constraints that forces the trade pattern to be sufficiently close to knowledge of the freight flows. Constraints that ensure trade pattern coherence with count data can be added. Care is required, however, while adding such constraints, since the counts data only represents transport between load, reload and unload, but no information of whether the transport originates from the sender or from some reload terminal which could be located in a different zone. There is neither information of whether the destination zone of transport and the receiver are located in the same zone, or whether the destination zone is just for reload. This implies that the count matrices cannot be used to assess transport flows between seller and purchaser, since the destination zone may just be a stop to reload. Most of the reload takes place after a relatively short trip by truck. Hence, there is probably an overrepresentation of short trips and under representation of the longest trips in the counts relative to the expected trip pattern transport between business links. We don't have data that tell us the extent of auxiliary trips and intermediate reload for the different commodity groups. Hence it is difficult to design a method that corrects the count matrices to eliminate the effects of auxiliary trips and reload.

Hence, to include constraints for the OD pattern based on the counts, we decided to make a 15x15 aggregate[2] (Table 4) of the count matrices and then scale down the diagonal by: 10%, 0%, 10%, 10%, 10%, 10%, 0%, 10%, 0%, 0% and 0%. Although this is very rough, the aggregation and the scaling have the effect that the count matrices become better approximations of the flow pattern of trade between municipalities.

Table 4. Count matrices aggregation from NTP-zones to a 15x15 zoning that was applied in the matrix balancing procedure. TØI report 699/2004.

	NTP-zones
1	1. Oslo
2	2. Ytre Østfold, 3.Indre Østfold
3	4. Hedemarken, 5.Glåmdal
4	6. Østerdalen, 7.Lilleh./Gjøvik, 8.Gudbrandsdalen, 9.Valdres
5	10. Hadeland, 11.Drammen, 12.Ringerike
6	13. Hallingdal, 16.Øvre Telemark, 22.Voss, 23.Hardanger, 25.Indre Sogn
7	14. Vestfold
8	15. Grenland, 17.Aust-Agder
9	18. Vest-Agder, 19.Jæren, 20.Nord-Rogaland
10	21. Bergen, 24.Sunnhordaland
11	26. Ytre Sogn/Fjordane, 27.Sunnmøre, 28.Romsdal
12	29. Nordmore, 30.Trondheim
13	31. Fosen, 32.Nord-Trøndelag, 33.Helgeland
14	34. Salten, 35.Ofoten/Lofoten/Vesterålen
15	36. Sør-Troms 37.Tromsø og Nord-Troms, 38.Vest-Finnmark 39.Øst-Finnmark

The aggregated and scaled counts were included by extending the doubly constrained model to a three-dimensional matrix balancing approach as stated in the STAN User's Manual (INRO, 1998), which is defined as the system of equations:

$$t_{ij} = \alpha_i \cdot \beta_j \cdot \gamma_{k_{ij}} \cdot f(c_{ij}(\mathbf{p})) \quad \forall i, j$$

$$\sum_{i=1}^{n} t_{ij} = A_j \quad \forall j$$

$$\sum_{j=1}^{n} t_{ij} = B_i \quad \forall i$$

$$\sum_{(i,j)|k_{ij}=k} t_{ij} = F_k \quad \forall k ,$$

where F_k, $k = 1,\ldots, 225$ are based on aggregated and scaled counts data for the tonnes transported between 15 aggregate regions, which are proportionally scaled such that

[2] The 435 municipalities was first aggregated into the 39 NTP-zones. Thereafter, the NTP-zones were aggregated into a 15x15 aggregation.

$\sum_j A_j = \sum_i B_i = \sum_k F_k$.

The system of equations consist of $m+n+n+n*n$ equations with the same number of unknowns, i.e., balancing parameters α_i, $i = 1,...,n$ and β_i, $i = 1,...,n$ and elements (transported tonnes) in the OD- matrix t_{ij}, $ij = 1,...,n$, where $n = 435$ (the number of municipalities) and $\gamma_{k_{ij}}$, k_{ij} = 1,…,225. As deterrence function we use

$$f(c_{ij}(\mathbf{p})) = \exp(\gamma \cdot c_{ij}(\mathbf{p})),$$

where **p** are parameters in the cost functions for the transport links. We needed to calibrate the transport costs $c_{ij}(\mathbf{p})$ between zones *i* and *j*, and the parameters $\gamma_{k_{ij}}$, $k_{ij} = 1,...,m$ and γ, for each commodity group.

3 Calibration and evaluation

3.1 Calibration

For each commodity group, the three-dimensional matrix balancing procedure, which is part of the logistic pathway model includes the calibration parameter γ and $\gamma_{k_{ij}}$ in the deterrence function and the parameters **p** in the real network cost functions that are unique for each commodity group[3], where the parameters $\gamma_{k_{ij}}$ were determined as part of the three-dimensional balancing procedure. The transport costs were adjusted within and obtained from the STAN based real network representation of the Norwegian transport network (TØ/1582/2004).

For commodity group 1, we have that the sum of the unscaled counts are about twice as high as the processed register data. It is not believed that the whole of this difference is due to reload. It is also believed that the difference is due to uncertainty in the data sources, and possible erroneous counts. Hence, while the difference in the tonnes distributed in the two data sources was so large, it made no sense to calibrate in order to obtain equal tonnekm in the balanced and unscaled counts OD matrices.

Instead, we determined γ and **p** for each commodity group iteratively such that shares transported per distance intervals (tonnes) and modal shares (tonne-km) in the final OD matrices becomes coherent with the count matrices (Table 5, 6, 7 and 7)[4]. This approach was in particular consistent with the aim of letting the register data determine the total flows in and out of the municipalities, while letting the scaled counts determine the OD patterns at a 15x15 zone representation. Herein the modal share of the balanced matrices was obtained by assignment of the balanced OD matrix in the STAN based real network representation. Satisfactory convergence was achieved for all commodity groups after 5-6 iterations.

[3] It is noticed that this balancing approach have the characteristics that the resistance represented by $\gamma_{k_{ij}} \cdot \exp(\gamma \cdot c_{ij}(\mathbf{p}))$ is discontinuous along borders where $\gamma_{k_{ij}}$ changes, i.e., within the 15x15 regionalization.

[4] If transport internal to the municipalities is not accounted for, then the total tonnes transported are reduced if γ is reduced, since this reduces the share or transport internal to the municipalities.

Table 5. The share of tonnes transported in different distance intervals in the balanced OD matrices. TØI report 699/2004.

Distance	1	2	3	4	5	6	7	8	9	10	11	Sum
0-50 km	37 %	24 %	41 %	52 %	40 %	44 %	58 %	27 %	34 %	29 %	19 %	41 %
50-100 km	24 %	26 %	29 %	25 %	24 %	23 %	16 %	30 %	21 %	21 %	24 %	22 %
100-200 km	19 %	18 %	17 %	14 %	16 %	15 %	13 %	21 %	22 %	12 %	16 %	16 %
200-300 km	8 %	8 %	5 %	4 %	6 %	6 %	5 %	7 %	9 %	4 %	7 %	6 %
300-400 km	4 %	9 %	3 %	1 %	3 %	4 %	4 %	2 %	6 %	4 %	3 %	3 %
400-500 km	5 %	9 %	3 %	1 %	5 %	3 %	2 %	3 %	3 %	8 %	9 %	4 %
500-750 km	2 %	5 %	1 %	2 %	4 %	4 %	1 %	4 %	2 %	9 %	10 %	3 %
750-1000 km	0 %	1 %	0 %	0 %	1 %	0 %	0 %	1 %	2 %	1 %	1 %	1 %
1000-1500 km	1 %	1 %	1 %	0 %	2 %	0 %	0 %	2 %	2 %	8 %	7 %	2 %
1500-2000 km	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	3 %	2 %	0 %
>2000 km	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	1 %	0 %	0 %
Sum	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**

Table 6. The share of tonnes transported in different distance intervals in the count matrices[1]. TØI report 699/2004.

Distance	1	2	3	4	5	6	7	8	9	10	11	Sum
0-50 km	39 %	24 %	43 %	54 %	42 %	45 %	65 %	29 %	33 %	30 %	21 %	46 %
50-100 km	27 %	24 %	27 %	22 %	21 %	23 %	15 %	31 %	23 %	20 %	18 %	21 %
100-200 km	17 %	18 %	15 %	14 %	13 %	19 %	12 %	17 %	16 %	16 %	11 %	14 %
200-300 km	6 %	8 %	5 %	3 %	5 %	7 %	3 %	5 %	5 %	4 %	5 %	5 %
300-400 km	2 %	9 %	2 %	2 %	2 %	3 %	2 %	3 %	5 %	3 %	2 %	2 %
400-500 km	5 %	10 %	5 %	2 %	9 %	1 %	1 %	5 %	6 %	7 %	16 %	5 %
500-750 km	2 %	6 %	2 %	2 %	3 %	3 %	1 %	6 %	3 %	8 %	12 %	3 %
750-1000 km	0 %	1 %	0 %	0 %	1 %	0 %	0 %	1 %	3 %	1 %	3 %	1 %
1000-1500 km	1 %	1 %	1 %	0 %	2 %	0 %	1 %	3 %	5 %	7 %	11 %	2 %
1500-2000 km	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	2 %	2 %	0 %
>2000 km	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %
Sum	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**

[1] Based on unscaled count matrices

Table 7. Modal shares of tonne-km in the balanced OD matrices (%), except tonne-km for transport within the municipalities which is not included in the STAN assignment procedures. TØI report 699/2004.

	1	2	3	4	5	6	7	8	9	10	11	Total
Truck	85	47	95	91	59	45	31	58	57	17	91	49
Rail	13	8	4	0	22	10	0	2	0	0	2	11
Waterborne	2	44	0	9	19	45	69	40	43	83	7	40

Table 8. Modal shares of tonne-km based on the mode specific count OD matrices (%). TØI report 699/2004.

	1	2	3	4	5	6	7	8	9	10	11	Total
Truck	84 %	56 %	95 %	91 %	59 %	45 %	32 %	58 %	58 %	17 %	56 %	50 %
Rail	14 %	9 %	5 %	0 %	22 %	13 %	0 %	2 %	0 %	0 %	9 %	10 %
Waterborne	2 %	36 %	0 %	9 %	19 %	42 %	68 %	40 %	42 %	83 %	36 %	40 %

3.2 Evaluation

Examination and comparison with data sources of high reliance on specific relations can be used to improve the balanced OD matrices by simply replacing the OD flows for these relations. There are no independent data sets available, however, for evaluation of the complete balanced OD matrices.

A perfect match was obtained between balanced matrices and scaled count matrices at the 15x15 zone aggregation, where the scaling accounted for the under-representation of long trips and over-representation of short trips rather subjectively (i.e., scaled diagonal and scaled total tonnes to equal total tonnes in the marginal).

The appropriate level of aggregation of freight flow OD matrices, depend on the application purpose. The required resolution is often higher than 15x15, e.g., 19x19 to represent the counties, 39x39 to represent the NTP-zones (zones applied for the National Transport Plan) or 435x435 to represent the municipalities. Due to the presumed overrepresentation of short trips in the unscaled count matrices and due to the separate representation of auxiliary trips and division of multimodal trips into single modal trips, it is our belief that it is of limited value to compare at a more disaggregated level, i.e., the balanced OD matrices represent complete transport whereas the count matrices represents the sum of transport per mode. Also there is no statistical measure for the uncertainty in the balanced OD matrices.

Although this implies that comparison of count and balanced matrices is of limited value, we have compared the balanced and count matrices (unscaled diagonal but overall scaling to make the total tonnes equal) at NTP-level (39x39 zones) by assuming that the trips between OD pairs follow a Poisson distribution, which means that expected number of trips equals the variation which is equal to the size of the observation. We used the counts as the observation and applied a statistical test to check the probability that the balanced matrices could represent the scaled count matrices (total scaling and scaling of the diagonal). Our general picture was that a great number of the elements were close to the counts, but expectedly large deviations were also found.

3.3 Proposed future work

It is our belief that the best way to reduce the uncertainty in the balanced OD matrices is to apply a more disaggregated commodity grouping. A more disaggregate commodity grouping makes the commodity groups smaller and more homogeneous which is advantageous, since this reduces the municipalities where trade takes place per commodity group. This is advantageous since fewer municipalities reduce the pairs of municipalities where trade possibly takes place. Hence by later aggregation of the disaggregate commodity groups, the trade pattern can be determined with greater certainty at the aggregated level for the current 11 commodity groups.

Referanser

Andersen, Jardar (2003). *Kostnadsfunksjoner i sammenslått versjon av NEMO*. TØI arbeidsdokument TØ1582/04. Oslo: Transportøkonomisk institutt.

INRO (1998) STAN User's Manual, Release 5, Les Conseillers Consultants Inc.

Rideng, Arne (2001). *Transportytelser i Norge 1946-2000*. TØI-rapport 515/01. Oslo: Transportøkonomisk institutt.

Vold, Arild, Inger Beate Hovi, Jardar Andersen, Olga Ivanova, and Viggo Jean-Hansen (2002). *NEMO: nettverksmodell for godstransport innen Norge or mellom Norge og utlandet.* TØI rapport 581/2002. Oslo: Transportøkonomisk institutt.

Appendix

Steps in construction, calibration and evaluation of the basis OD matrices

Construction of the basis OD matrices is based on a series of steps:

1. Application of EXCEL spread sheet for processing of register data to produce the data for total transport in and out of the municipalities.
2. A C-program to transform data from the EXCEL spread sheet to data files that can be read by the STAN computer software into a STAN databank with the OD matrices based on counts and a real network transport representation with link based cost functions.
3. A set of STAN macros for reading processed register data into the STAN databank, three-dimensional matrix balancing, and for writing balanced OD matrices and results from real network assignment of the OD matrices. Ensambles *ge* and *gf* were established to represent OD matrices at 15x15 (applied in balancing procedure) and 39x39 NTP-zoning level of aggregation.

A1 STAN macros

STAN matrices can be referred to according to a numbering system and according to names. Deletion and subsequent reestablishment of the matrices occurs frequently, which may changes their numbering. Thus, the use of unique names was primarily used to avoid problem with inconsistencies caused by altered numbering of the matrices.

To make the balanced OD matrices as part of the STAN databank, the following sequence of STAN macros (with short explanations) need to be run, where the subsequence 9-11 need to be run iteratively for calibration:

1. margin.mac

Reads files md12.inn,…, md22.inn and mo12.inn,…,mo22.inn with processed register data for domestic transport in and out of the municipalities into the STAN databank. The files are represented as matrices (omag1-omag11 and dmag1-dmag11) in the STAN databank.

2. Lag1ermatrise.mac

This macro makes the matrix mf47 with one in all elements.

3. Ass1.mac

Runs assignment with mf47 as demand. The assignment generates the transport costs per tonnekm for each commodity group. The transport costs are currently capacity independent. Hence, the volumes on the links do not affect the transport costs per tonnekm.

4. Slettkostmatriser.mac

This macro is run to delete existing cost matrices.

5. Initms10_225.mac

This macro calls *initms.mac* that initialises *non-existing* scalar matrices that are needed in the constraints in the three-dimensional matrix balancing procedure.

6. Styrkostmat.mac

This macro calls *lagkostmatriser.mac* for all commodity groups.

lagkostmatriser.mac

This macro creates a cost matrix for a commodity group (mfkost1,…,mfkost11). Element in these matrices are used in the deterrence function in the three-dimensional matrix balancing procedure.

7. Fromat.mac

This macro initialises matrices for the commodity groups that are needed to store the deterrence function for each of the commodity groups for each of the 438x438 OD pairs, $\exp(\gamma_1 \cdot mfkost1_{ij}),\ldots,\exp(\gamma_{11} \cdot mfkost11_{ij})$.

8. StyrKpq.mac

This macro calls *GenKpq.mac* for all the constraints that are needed to obtain coherence with the 15x15 count matrices.

GenKpq.mac

This macro generates the matrix $[k_{pq}]$ that specifies which elements of the count data matrices that should be aggregated (i.e., summed) and assigned to the third-dimension balancing elements F_k.

9. Styrgamma.mac

Put new values for $\gamma_1,\ldots,\gamma_{11}$ for the commodity groups (ms01, gamma1 – ms11, gamma11).

10. Styrbalance.mac

This macro establishes elements in the matrices for the deterrence function for each commodity group and saves the matrices (mfseed1-mfseed11). For each commodity group, then, there is a sequential call to the macros *Gen3tell.mac, Gen3komp.mac* and *bal3dim.mac*.

Gen3tell.mac

This macro takes as input a count data matrix for a commodity group (V1-tel, tel-2, V3-tel,…, V10tel, tel-11) and a factor that is used to scale the diagonal elements of the aggregated 15x15 count matrices to reduce the effect of reload. The elements in the 15x15 count matrices are saved in tell1,…,tell225.

Gen3komp.mac

This macro scale the 15x15 count matrices such that the total tonnes transported become the same as the total tonnes transported according to the processed register data.

Bal3dim.mac

This macro makes three-dimensional matrix balancing for the commodity groups. Input parameters are origin and destination vectors (omag1-omag11 and dmag1-dmag11), matrices for storage of balanced OD-matrices (nemo1-nemo11) and matrices representing the deterrence function. The macro applied the elements in the 15x15 count matrices tell1,…,tell225.

AddDiag7.mac
This macro substitutes the diagonal of nemo7 by the diagonal from the corresponding count matrix V7-tel.

11. Styrtonndist.mac and Styrtonndisttell.mac

These macros both call *tonndist.mac*

tonndist.mac
This macro creates **tonndist.out** and **tonndisttell.out** for the balanced and counts matrices, respectively, with tonnes transported in different distance intervals of each commodity group. t**onndist.out** and **tonndisttell.out** do not include distance internal to the zones. Needs pre-run of the macro: *lagkortestevei.mac.*

12. Styr3mat.mac and Styr3mattell.mac

This macro call *3mat.mac* for each commodity group to generate 3x3 aggregated matrices for either counts data matrices or balanced matrices.

3mat.mac
Generate the outputfile **3matriser.out** and **3matrisertell.out,** respectively. The macro is set up to include trips internal to the municipalities. The diagonal elements of **3matriser.out** are scaled (see gen3tell.mac).

13. assign_innenriks.mac

This macro makes assignment based on balanced matrices. This generates total tonnekm for the balanced OD matrices in the transport network. These tonnekm do not include tonnekm inside municipality zones, since STAN does not calculate tonnekm for trips inside zones. The macro calls *skrivres.mac.*

Skrivres.mac
This macro writes out assignment results (tonnekm) for balanced matrices to **assres.out**.

14. assign_innenrikstell.mac

This macro makes assignment based on the sum of mode specific count matrices for truck, rail and boat. The macro calls *skrivrestell.mac*.

Skrivrestell.mac
This macro writes out assignment results (tonnekm) for count matrices to **assrestell.out**.

Other STAN macros:

*15. Annot**

Corresponds to "demarcation"

16. Addermatriser.mac

This macro sums OD matrices for domestic and international transport.

17. Styrzgr_out.mac

May either call *zgroups_out.mac* or *zgr15_out.mac*, for generation of balanced and count OD matrices for 39x39 and 15x15 level of aggregation.

Zgroups_out.mac and *zgr15_out.mac*
These macros take the name of a 438x438 commodity OD matrix as input and generate 39x39 and 15x15 OD matrices, respectively.

18. Styrtrspm.mac

This macro call *trspm1.mac* sequentially for each commodity groups, then *stykk.mac*, *tommer.mac*, *torrbulk.mac* and *flyt.mac, and* finally *trspdel.mac* sequentially for each commodity group. The macros subdivides the balanced OD matrices for total transports into mode specific balanced OD matrices for the 11 NEMO commodity groups and the 4 old NEMO commodity groups. The macros write the matrices aggregated to 39x39 level to file. Finally *trspdel.mac* delete these matrices from the STAN databank.

A2 Evaluation

Evaluation of the basis OD matrices is performed with the aid of the EXCEL file KONVERT.xls (remember to use "." as the decimal point), where one needs to answer yes on the initial question of whether to activate the initial macros.

A BASIC program is called from the sheet "Innenriks_resultater", which reads in result files from the STAN-macros (see AI.1). The BASIC program can be changed by starting Visual BASIC from the Excel menu and then choosing either Forms: DialogFiledef or Modules: Hent_filB, Hovedprogram.