TOI report
731/2004

Basis data requirements for developing the freight model system in Norway

Inger Beate Hovi
Arild Vold
Viggo Jean-Hansen
Steinar Johansen
Anne Madslien

ISSN 0802-0175
ISBN 82-480-0441-4

Oslo, November 2004

Title: Basis data requirements for developing the freight model system in Norway

Author(s): *Inger Beate Hovi; Arild Vold; Viggo Jean-Hansen; Steinar Johansen; Anne Madslien*

TØI report 731/2004
Oslo, 2004-11
49 pages
ISBN 82-480-0441-4
ISSN 0802-0175

Financed by:
Joint workgroup for transport analyses

Project: 2997 Basis data requirements for developing the freight model system

Project manager: Inger Beate Hovi

Quality manager: Kjell Werner Johansen

Key words:
Freight transport; Statistics; Network model; Basis Data

Summary:
The aim of this project has been to go through available data for further developing the freight model in Norway on national and regional level, and identify needs for future data collections.
Register data from SSB are available at more detailed level than recommended used in regional freight models. The main missing data is information about delivery pattern. This can be obtained for use in regional models from extractions from transport operators data systems. The main problem is to get a covering picture, caused by the high amount of small enterprises, not expected to have their own data systems. Another problem is to catch transports on own account. Related to a locistic module there are missing data about consignment size and volumes through internal warehouses. This information in addition to delivery pattern (at national level) and improved information about transport costs can be obtained from a commodity flow survey. The main drawback of such a survey is the high level of costs related to carry it through with a sufficient level of quality.

Tittel: Behov for grunnlagsdata for videreutvikling av godsmodellsystemet i Norge

Forfatter(e) *Inger Beate Hovi; Arild Vold; Viggo Jean-Hansen; Steinar Johansen; Anne Madslien*

TØI rapport 731/2004
Oslo: 2004-11
49 sider
ISBN 82-480-0441-4
ISSN 0802-0175

Finansieringskilde:
NTP Arbeidsgruppe for transportanalyser

Prosjekt: 2997 Grunnlagsdata for tverretatlige godsmodeller

Prosjektleder: Inger Beate Hovi

Kvalitetsansvarlig: Kjell Werner Johansen

Emneord:
Godstransport; Statistikk; Nettverksmodell; Datagrunnlag

Sammendrag:
Formålet med dette prosjektet har vært å gå gjennom det som finnes av tilgjengelige bakgrunnsdata for godstransportmodeller på nasjonalt og regionalt nivå, og hva som eventuelt trengs av nye data.
Det finnes tilgjengelige registerdata i SSB på et mer detaljert nivå enn det vi her anbefaler å bruke i regionale godstransportmodeller. Den viktigste manglende datakilden er informasjon om leveransemønster. For regionale modeller kan dette hentes fra transportørenes datasystemer, men man har et problem med å få et dekkende bilde på grunn av at transportnæringen består av mange små foretak som ikke har egne datasystemer. Det er også en utfordring å få med egentransportene. Relatert til logistikkmodulprosjektet er det manglende data om sendingsstørrelse og godsstrømmer gjennom internlagre. Denne informasjonen i tillegg til informasjon om leveringsmønster (på nasjonalt nivå) og forbedret informasjon om transportkostnader kan bli hentet fra en varestrømsundersøkelse. Ulempen med en slik undersøkelse er at den blir svært dyr å gjennomføre hvis den skal få et omfang som gir god datakvalitet.

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

This report has been worked out on demand for the NTP working group for transport analyses. The main purpose has been to identify basis data requirements for developing the freight model system in Norway, namely the NEMO and PINGO models. The approach national versus regional models is discussed and recommendations for zonal level and data availability for regional models are given. Finally, recommendations for additional data surveys are given.

There have been two project meetings with the NTP working group during the project. The working group consist of Oskar Kleven (who is the project leader in the working group), Linda Alfheim (Norwegian Public Roads Administration), Harald Thune-Larsen (Avinor AS), Anita Vingan (Norwegian National Rail Administration) and Erik Ørbeck (Coastal Administration). We are grateful to useful comments given by the working group.

The project leader at TØI has been cand oecon Inger Beate Hovi who has written the report in co-operation with cand oecon Viggo Jean-Hansen, cand oecon Steinar Johansen, M.Sc. Eng Anne Madslien and dr scient Arild Vold. Head of department Kjell Werner Johansen has been responsible for the quality assurance and secretary Laila Aastorp Andersen has been responsible for the final layout.

Oslo, November 2004
Institute of Transport Economics

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

Contents

TØI-rapport 731/2004
Forfatter(e): Inger Beate Hovi, Arild Vold, Viggo Jean-Hansen,
Steinar Johansen og Anne Madslien
Oslo 2004, 49 sider

Sammendrag:

Behov for grunnlagsdata for videreutvikling av godsmodellsystemet i Norge

Innledning

Formålet med foreliggende rapport har vært å beskrive dagens nasjonale modellsystem for godstransport og peke på metodiske svakheter og manglende data. Et annet mål har vært å beskrive fremtidig behov for grunnlagsdata knyttet til videreutvikling av den nasjonale godstransportmodellen og prinsipper for å utvikle modeller til analyser på regionalt nivå. Bakgrunnen er at man ønsker å oppnå et bedre bilde av transportmønsteret i den nasjonale godsmodellen NEMO. Rapporten er utarbeidet på oppdrag for Nasjonal Transportplans (NTP) arbeidsgruppe for transportanalyser.

I rapporten konkluderes det med at PINGO (prognosemodellen) and NEMO (nettverksmodellen) generelt er metodisk gode nok og at en logistikkmodul vil passe godt inn i modellsystemet. Imidlertid finnes det metodiske svakheter, grunnet mangel på data som kan forbedres ved gjennomføring av en vareströmsundersøkelse (VSU) i Norge. Forbedring kan også oppnås ved bedre utnytting av eksisterende datakilder, blant annet ved å øke antall soner og/eller varegrupper i modellen.

Logistikk

Det gis en kortfattet diskusjon av et foreslått opplegg for å implementering av en logistikkmodul i hhv den svenske og norske godstransportmodellen (de Jong et al, 2004). Vi er i hovedsak enige i de foreslåtte prinsipper, men stiller størsmål til detaljeringsnivå, siden dette kan bli tungvint både med hensyn til behovet for nye data og at kompleksiteten vil lede til økte utviklings- og vedlikeholdskostnader for modellsystemet. Vi er også kritiske til planene om å gå fra matriser for produksjon, engros og konsum (PWC-matriser) til matriser for produksjon og konsum (PC-matriser), siden dette vil være i konflikt med økonomiske modeller generelt og PINGO spesielt.

TØI har nylig gjennomført en terminalundersøkelse som vil kunne forbedre datagrunnlaget for utvikling av en logistikkmodul, men informasjon om lagerbeholdning og størrelse på varesendingene mangler fortsatt.

Aggregeringsnivå

Et viktig prinsipp med hensyn til valg av soneinndeling i godsmodeller på nasjonal versus regionalt nivå, er at det må være konsistens i varestrømmene mellom de forskjellige aggregeringsnivå. Det er viktig at varestrømmene blir representert uten overlapping (f eks må det tas stilling til om servicetransporter skal inngå i gods- eller persontransportmodellene på regionalt nivå). Valg av sonestørrelse i regionale godsmodeller bør vurderes med tanke på behovet for sammenheng mellom gods- og persontransportmodeller, analyser og datatilgang. I dagens nasjonale godstransportmodell er sonestørrelsen enkeltkommuner, noe som er tilfredsstillende for godstransportanalyser på overordnet og nasjonalt nivå. For analyser av problemstillinger på regionalt nivå bør antall soner i nettverksmodellen økes, og det konkluderes i rapporten med å benytte en soneinndeling som tilsvarer det som er benyttet i den nasjonale persontransportmodellen (NTM5). Dette muliggjør at en vil kunne foreta felles analyser for person- og godstransport i modellsystemet.

En svensk godstransportmodell for analyser på regionalt nivå (DISTRA) er beskrevet for å illustrere omfanget til en relevant regional modell. Mulig bruk av den eksisterende, regionaløkonomiske modellen PANDA, er skissert. PANDA består av to delmoduler, en for prediksjon av produksjon og sysselsetting og en for befolkningsvekst i regioner. Regionene kan spesifiseres til å bestå av fra 1 til 435 kommuner. Kryssløpstabeller og regionale innkjøpskoeffisienter (RPC) vil kunne gi informasjon på kommunenivå til å trekke soneinterne varestrømmer ut av marginalene i NEMO før kjøring av gravitasjonsmodeller for å etablere et leveransemønster. Informasjon om produksjonsverdier er tilgjengelig fra PANDA på kommunenivå og kan brukes for å utvikle PINGO til en prognosemodell for regionale modeller.

Tilgjengelige datakilder og behov for nye data

Ved utvikling av regionale modeller trengs et mer detaljert datamateriale enn det som er benyttet til utvikling av dagens nasjonale modell. Det må derfor fremheves at det foreligger betydelige muligheter for å frambringe et mer detaljert grunnlag fra SSBs registerdata, både når det gjelder industri, varehandels- og utenrikshandelsstatistikk. Spesielt fremheves muligheten for å hente ut mer detaljert informasjon for industri- og varehandelsstatistikk på postnummernivå, som kan aggregeres opp NTM5-nivå. Muligheten for mer detaljert informasjon i eksisterende statistikk gjelder både mht geografi, men også mht aggregeringsnivået for varegrupperingen. Å øke antall varegrupper i modellen, vil føre til at en kan utnytte mer informasjon mht om varen brukes som innsatsfaktor (går til industri) eller om det er ferdigvarer som går til varehandel. En vil derved kunne etablere et bedre leveransemønster enn ved dagens gravitasjonsmodeller uten å gjennomføre undersøkelser av leveransemønsteret.

Den viktigste manglende datakilde er opplysninger om leveringsmønster, både mht geografi, men også mht hvilke forretningsledd transportene går gjennom. Det konkluderes derfor med at skal en komme vesentlig videre i arbeidet med å frambringe gode matriser for varestrømmene, bør det tas sikte på å gjennomføre en varestrømsundersøkelse (VSU) etter tilsvarende mønster som det er gjort i Sverige. Ulempen med slike undersøkelser er at de er svært dyre å gjennomføre. Kostnadene kan imidlertid begrenses dersom undersøkelsen avgrenses til å omfatte bestemte regioner eller noen utvalgte næringer.

Alternativt til en VSU kan en for regionale modeller framskaffe opplysninger om forsendelsesmønster på grunnlag av elektronisk informasjon fra transportørenes egne dataregistre. En vil trolig kunne få tilgang til informasjon på postnummernivå, men det vil være problematisk å framskaffe et materiale som gir et dekkende bilde, pga at det er svært mange enmannsforetak i transportbransjen, og disse har neppe egne datasystemer. Et annet problem vil være å få oversikt over egentransportene.

Konklusjon

Formålet med dette prosjektet har vært å gå gjennom det som finnes av tilgjengelige bakgrunnsdata for godstransportmodeller på nasjonalt og regionalt nivå, og hva som eventuelt trengs av nye data.

Det finnes tilgjengelige registerdata i SSB på et mer detaljert geografisk nivå enn det vi her anbefaler å benytte i regionale godstransportmodeller. Den viktigste manglende datakilden er informasjon om leveransemønster. For regionale modeller kan dette delvis hentes fra transportørenes datasystemer, men man har et problem med å få et dekkende bilde på grunn av at transportnæringen består av mange små foretak som ikke har egne datasystemer. Det er også en utfordring å få med egentransportene. Relatert til logistikkmodulprosjektet er det manglende data om sendingsstørrelse og godsstrømmer gjennom internlagre. Denne informasjonen i tillegg til informasjon om leveringsmønster (på nasjonalt nivå) og forbedret informasjon om transportkostnader kan bli hentet fra en varestrømsundersøkelse. Ulempen med en slik undersøkelse er at den blir svært dyr å gjennomføre hvis den skal få et omfang som gir god datakvalitet.

TØI report 731/2004
Author(s): Inger Beate Hovi, Arild Vold, Viggo Jean-Hansen
Steinar Johansen and Anne Madslien
Oslo 2004, 49 pages, English language

Summary:

Basis data requirements for developing the freight model system in Norway

Introduction

This report was produced for the National Transport Plan (NTP) working group for transport analyses. The objective of the report is to describe the current national freight model system and pinpoint its methodological insufficiencies and the lack of data. The objective of the report is also to describe the desired data collection and model development to improve the national model system and to suggest principles for development of freight models and the implied data requirements at the regional and local scale.

We conclude that in general PINGO (the forecast model) and NEMO (the network model) are methodologically sound, and that a logistic module will fit well into the model system as a whole. There are methodological insufficiencies though that are grounded in lack of data, and it is concluded that insufficiencies can to a large extent be corrected for if data from a commodity flow survey (CFS) becomes available. However, it is realized that model improvements can also be obtained by better utilization of existing data sources to increase the number of commodity groups.

Logistics

We discuss briefly the intermediate logistic project report by de Jong et al. (2004). We agree about the most of the plans but we question the level of detail and disaggregation since this can be cumbersome both with regard to the fact that new data is required and because the complexity will lead to increased development and maintenance costs of the model system. We are critical about the plans to move from Production-Wholesalers-Consumers (PWC) matrices to Production-Consumers (PC) matrices since this will be in conflict with economic models in general and PINGO in particular. It is explained that TOIs ongoing terminal survey can improve the data availability for development of the logistic module, but information about internal inventories and consignment size are missing data.

Levels of aggregation

An important issue with regard to the National versus regional modelling levels is the possibility of obtaining consistency of the freight flows in different modelling levels. It is important that all the different trip purposes are represented and that there is no overlap (e.g., we need to decide whether service trips should be part of the regional freight or passenger models). The degree of disaggregation should be considered with regard to the need for coherency between freight and passenger models, analysis requirements and with regard to data availability. The municipality level is considered sufficient for freight analysis, but expansion to the NTM5[1] level (with the aid of specific indicators) is mandatory in order to carry out joint passenger and freight analyses, and also introduce congestion in the model system.

The DISTRA model is briefly described to illustrate the scope of a relevant regional model. More in depth considerations about the possible use of the existing PANDA model are then outlined. The PANDA model consists of two main modules, an input-output module for prediction of production/employment and a demographic module for population growth. A region is defined as a number of municipalities aggregated together (the number of municipalities in a region can, in principle, range between 1 and 435). It is considered viable to use the Input-output tables and the Regional Purchasing Coeffecients (RPC) in PANDA at the municipality level to subtract zonal internal flows from the NEMO marginals before running the gravity model to establish the delivery pattern. Information about production values is available from PANDA at municipality level and could thereby be used to

[1] NTM5 – The Norwegian **N**ational **T**ransport **M**odel, i. e the national model for Passenger transport in Norway.

develop PINGO to be a forecast model for regional models as well.

Available data sources and needs for further data

Development of regional models requires identification of data requirements, which should be followed by exploitation of existing data sources and collection of new data. This report includes a general overview of data about freight transport at the regional level. It explains how to possibly apply the available trade statistics at the postal code zoning level. The manufacturing statistics and the external trade statistics could in principle be disaggregated in the same pattern. For other kinds of production, were information is not available at the zonal level of postal codes, techniques are available for this purpose including use of indicators to spread data to the desired zoning level.

Changing in existing data sources

Some existing statistical sources that are regularly updated have gone trough some changes that lead to less quality in the continuation of the statistics. The lorry survey is one example where the vehicles between 1 and 3,5 tons have been removed. Moreover the survey among coastal vessels has not been carried out since 1993, and the external trade statistics suffer from the fact that the number of custom offices will be halved during 2004 with the effect of a more coarse external trade statistic.

Recommendations for new data

It is a challenge to follow up the intentions about improving and developing the national freight model as well as the new regional freight models in a situation where the need for new data emerges and the quality of some existing data sources are reduced. It is concluded that a commodity flow survey (CFS) will fill many needs. Other means of collecting data is also of interest and can give both supplementary information and overlapping but more detailed information that can give us in-depth insight. Until now there have not been any CFS for Norway[2], and they are expensive to conduct. Despite the cost, we think that this could be the way to proceed. New data in terms of a commodity flow survey would also become useful, in that this would make it possible to apply the CFS data in the matrix balancing instead of the mode specific counts and then save the mode specific counts for evaluation purpose.

Main conclusion

Improved information of freight flows can be achieved from use of more detailed statistic than already used in the development of the freight model system for Norway. Register data from SSB are available at more detailed level than recommended used even in regional models in this report.

The main missing data is information about delivery pattern (both with respect to geography but also with respect to witch business link the delivery goes through). Information about delivery pattern between detailed zonal sizes can be obtained for truck transports for use in regional models from data extraction from transport operators data systems. The main problem is however how to get a covering picture, because the transport marked consists of many small enterprises (some with only one employer) were the small transport operators are not expected to have any electronic system for their transports at all. Another problem is how to catch transports on own account.

With respect to the logistic module, there is a lack of information about consignment size and volumes through internal warehouses (inventories). Both this information and the information about delivery pattern and transport costs can be obtained from a commodity flow survey (CFS). There are experiences from such surveys from both USA and Sweden. The main drawback with a CFS is however the high level of costs related to such a survey, but a survey can be limited to either a geographic area or a limited number of industries.

[2] Except for an old study by Statistics Norway for some traded commodities (Strøm K F: Varestrømmer i engros- og detaljhandel. Rapporter 83/31. SSb (1983) This study was carried out in order to split the trade sector in the National Accounts in retail- and wholesale trade.

1 A short presentation of the National freight model system

The national model system for freight transport in Norway consists of both the demand model PINGO and the network model NEMO. In addition, a logistic module is planned implemented in the network model. Multimodal freight models on regional level are not yet established.

PINGO and NEMO together represent a network model with elastic demand. Capacity restraints and congestion are not implemented in the network model, therefore it is not necessary to do any interactive iterations between PINGO and NEMO to establish equilibrium.

1.1 Background and approach

New data can be included for estimation and calibration to improve the reliability of the national freight model system. But new data can also make it possible to estimate new and more sophisticated versions of the model. A more sophisticated model can open the opportunity for more advanced model analysis.

Related to the work of improving the national model system for freight transport in Norway and establishing regional models for freight transport, the NTP working group for transport analysis has asked for a pre study to investigate the desired data collection and model development to improve the national model system and to suggest principles for development of freight models and the implied data requirements at the regional and local scale. This includes theoretical and methodological considerations for introducing a more disaggregated zonal level in the model, suggestions and recommendations for a way to improve the PC and OD matrices in the model and the relation between these, and finally a way to improve the cost functions in the model and representation of logistic pathways. The reason is a request for a more proper view of the transport pattern in the National model, but also the need to investigate possibilities and obstacles to development of freight models for the regional scale.

An important feature of freight models is to have whole transport chains represented in the model where stock holding, distribution centres, terminals, etc are implemented in addition to localisation of production and consumption of commodities. A common terminology for this is logistics.

It is important, though, that we have a clear view of what we want to use new data for. This means that data collection and model development should be connected with the objective of achieving a model that both satisfies our need for mechanisms and the requirement that model results should be within acceptable errors. Fulfilment of the needs of the NTP analysis group, which is both to improve the reliability of the model and to do more advanced analysis, requires that the model system is further developed and that new data are collected. Our view is that the current model system is a good starting point - both with regard to further improvement at the national level and also with regard to development of freight transport models at the regional scale.

Chapter 1 contains a short presentation of the current national model system. Further model development and new data to improve the national model is discussed in chapter 2. Chapter 3 contains a discussion about national versus regional models. The discussion addresses the possible increase in the number of zones, which modelling principles from national models that can be adapted to the regional scale and new modelling principles including data requirements. Chapter 4 gives an overview of data availability for regional models, hereby what is available but unused data, new data sources caused out of new technology and a presentation of facilities in Panda and some relevant statistical sources that can be exploited in order to derive the required data to establish models at the regional level. Chapter 5 gives an overview of changes in the existing data sources. Finally Chapter 6 outlines possible plans for data collection based on the requirements set forth in chapter 2 to 5.

1.2 Principles of NEMO

The NEMO model is a national freight network model that determine transport solutions for freight flows between the 435 municipalities in Norway, import, export, transport to and from the continental shelf and transit transport using Norwegian infrastructure (Hovi and Vold 2003; Vold and Hovi, 2004). The model include (1) basic freight flow matrices that represent the transport pattern subdivided by 13 commodity groups (2) cost functions for the transport of the 13 commodity groups[1] by truck, rail, boat, plane and pipelines and (3) a real network representation of the infrastructure for each mode in Norway and the main infrastructure net in Europe. The OD matrices and cost functions are calibrated for 1999. NEMO is represented in the STAN computer software, which includes algorithms that assign the OD flows to uni-modal and intermodal transport solutions. The assignment selects one transport solution for each commodity group between each pair of zones (municipalities). The selected transport solution minimizes the total cost of transporting the commodities, i.e., the price paid to the transport operators plus commodity owner's non-paid costs. The transport operators cost include distance dependent and time dependent cost, and loading and unloading costs. Non-paid costs include the costs of degradation, time value for commodities during leading time, risk of damaging on goods during transporting and expectations about delays.

NEMO can be used to assess how changes in the OD-flows and in the transport cost components affects the transport solutions, and hence the mode choice distribution. NEMO is suitable for analysing short-term effects of changes in infrastructure or transport costs. The absence of capacity restraints and congestion in the model implies that there are not transport costs and invariant with respect to time.

1.3 Principles of PINGO

PINGO (Ivanova, Vold and Jean-Hansen, 2002; 2003) is an economic model (Spatial Computable General Equilibrium model) that is based on a so-called Spatial Accounting Matrix that represents input and output of production and consumption per county for 11 commodity groups. In each county PINGO represents nine production sectors producing the 11 commodities. There are also sectors for production of service and physical capital (investments). Representative households earn income by selling their work to the

[1] The groups are 1) Food, 2) Fresh fish, 3) Thermo products, 4) Vehicles /machinery, 5) General cargo, 6) Timber and wood ware, 7) Minerals and stone products, 8) Chemical products, 9) Metal waste and ore, 10) Petroleum products, liquid, 11) Frozen fish, 12) High value commodities and 13) Petroleum products, gas

production and service sectors and from social security services etc. Retailers in each county are trading with retailers in own and other counties and foreign countries. This is represented in the SAM matrix in terms of transport flows between counties per commodity group and the cost of transport are also represented in the matrix. The SAM matrix is established on the basis of the National Accounts by County, the External trade statistics and OD flows and costs from the base year representation in NEMO.

PINGO ca be used to assess how external factors (e.g., price changes on commodities or transport, population growth, increasing exports, increasing incomes, infrastructure investments, governmental transmissions and technological innovations) affect the freight flow pattern for each commodity group in and between 19 counties and between each county and each of the foreign zones. These changes can be used in the NEMO model to calculate effects on transport work and hence external environmental costs etc. PINGO can also be used as a top-down model to spread the growth of national forecast scenarios determined by MSG-5[2]. Since PINGO is more aggregated than NEMO, data transmissions between the two models have to be aggregated or disaggregated. PINGO is suitable for estimating changes both in a short and a long-term perspective.

1.4 Data used in NEMO and PINGO

1.4.1 Base matrices, NEMO

Domestic transport contains transport between all pairs of municipalities in Norway. This includes transport of imported commodities from the place where toll is paid and further into the Norwegian transport network. Base matrices for international transports contain export from the municipalities where the commodities are produced and import from other countries to the municipality in Norway where the toll is paid.

Base matrices were based on basic data from Agricultural -, Forestry-, and Fishery statistics, Manufacturing statistics (including mining) and Trade statistics (comprising both wholesale and retailers). All the statistics are produced by Statistics Norway for the year 1999. The statistics contain data for production and trade of commodities. Most of the data are given in terms of production values (basic values excluding indirect taxes) for the 13 NEMO commodities.

The values are transformed to tons in a way that gives an unambiguous picture of the freight flows for each commodity group out of and in to each of the 435 municipalities in Norway. The Manufacturing statistics, includes data used to determine a conversion factor from value to tons, and contains:

- Production value for industrial products
- Sales value for industrial products
- The cost of input used in industrial production
- Sales value of products sold by industries
- Purchase value of traded products sold by industries

The volumes in tons of each commodity group are calculated from single enterprise establishments in the Manufacturing statistics. (These enterprises constitute around 25 % of all establishments in the statistics.) For larger enterprises with many establishments in different manufacturing industries included, the value/volume ratio is not given for each

[2] MSG – Multi Sectored Growth model. The Ministry of Finance in Norway uses the model in long term planning for spanning out possible solutions for the Norwegian economy.

establishment in the enterprise. This problem can be improved by further work; e.g. to compare the value/volume ratio for larger products from the External Trade statistics. The Statistics for the Trade industry contain data for the sales values from establishments in the traded products in 1999. The establishments in the trade industry are divided in wholesale, agency and retailers of specific products.

Delivery pattern are not given from the industry statistics, only sum out of and into each municipality. A tree dimensional gravity model had to be used for ensuring the delivery pattern, were transport costs and information about delivery pattern from the transport counts in Norway (more detailed explained in chapter 1.3.4) were used as explanatory variables

Base matrices for foreign transport (tons) between counties in Norway and other countries are obtained from the External Trade statistics for 1999 (Statistics Norway). The elements in the base matrices represents transport from the county where the producers are located in Norway to destinations in other countries and from origins in other countries to the first place of custom declaration in Norway. Subdivision from counties to municipalities is done on basis of indicators established from the production data (export) and input and wholesale statistics (import). Also some corrugations have to be done for imports to the Østfold and Oslo counties because of a squewness in the reporting method in the External Trade.

1.4.2 Cost functions, NEMO

Operative costs, truck

Distance dependent cost for truck is based on estimates on energy use by lorry type (Rypdal et al, 1997). Time depended costs are based on the Estimated lorry cost index and Survey of lorry transports in Norway from 1998-2000 (Statistics Norway).

Operative costs, boat

Distance dependent cost for coastal transport is also based on estimates on energy use by ship type (Rypdal et. al, 1997). In order to calculate commodity specific distance- and time dependent costs for domestic transport by boat, maritime statistics for 1993 (Statistics Norway) was used to determine the shares (in tonne km) of each commodity group, transported with different types of ships, under assumptions of unchanged shares from 1993 until 1999 (in lack of data).

Time dependent cost for coastal transport is based on yearly accounts for domestic hired transport in Norway (published yearly in Maritime statistics until 1996, Statistics Norway. After 1996 information is only available from the cost components from National accounts).

Data for costs for international transports by ships are scarce. The observed prices for such transports depend upon the demand for the type of product the ship is designed to transport. (E.g. the value of new contracts for hire of panamasize dry cargo ships depends upon the steep increase for the demand of fertilisers created by modernisation of the agriculture in China. This will might be a cost for the hiring company, but the real costs for the shipping service involved are unchanged.)

Construction of cost functions for international ferries was based on transport prices for 1999 from official time and route tables.

Operative costs, train

Distance dependent costs for train are calculated on the basis of total energy- and diesel consumption for freight train in 1999 from the energy accounts of Norwegian Railway Company (NSB). There is no available data, however, for capacity utilization related to different commodities[3]. To circumvent this problem, data from Survey of lorry transport in Norway from 1998-2000 (Statistics Norway) was used for capacity utilization by commodities for transports longer than 300 km. The cost components were adjusted according to this. An infrastructure charge comes in addition.

Time dependent cost for train is calculated on basis of transport prices by subtracting the distance dependent cost from the transport prices minus the largest discount. This is not a very satisfactory method, since the profit in this case becomes part of the time dependent cost component. However, according to company accounts for NSB Freight it is not unreasonable to assume zero profit.

Reload costs

Operative costs for reload are based on a terminal cost survey accomplished by Lervåg et al (2001).

Calculations of the operator's time dependent costs for vehicles in the terminal are based on assumption that loading and unloading are equal in time.

Quality cost (non-operative costs)

Quality costs describe characteristics of a transport service that may affect mode choice, which includes information about risk of delay (on links, at the border and in terminals), obtained from the Swedish model, and estimates on transport frequencies (per week) for available transport modes.

Quality costs also includes capital costs for commodities related to transport time, where commodity value per tonne where obtained from the Foreign Trade Statistics 1999. Degradation costs are added for 'fresh fish', 'thermo products' and 'vehicles and machinery' representing the value loss with respect to time for these commodities.

1.4.3 Network, NEMO

The road and rail networks were based on the EMME/2-network applied in an earlier version of the national passenger transport model. In cooperation with the Norwegian State Railways we decided upon approximate 40 railway stations where transfers between modes are possible. The sea network was developed by SINTEF/MARINTEK and includes altogether 169 ports.

When it comes to connecting the networks for each mode and connecting the networks to the centroïds, all freight transported on rail and sea are assumed to be part of a chain involving truck transport in the beginning and the end of the chain. Since this is an oversimplification the picture are corrected by collecting information on how the freight really is handled in each port and on each railway station, to establish direct connection links.

Network outside Norway is based on the STEMM network, covering main roads, rail and sea-borne corridors in Europe. This network is rather simple and could be further developed. Airfreight to and from Norway is implemented in the model as direct lines between two airports. A network for pipelines is implemented in the model in 2003.

---

[3] The capacity is measured as commodity weight per trip as the share of the containers carrying capacity.

1.4.4 Calibration data, NEMO

To ensure that given percentages of the freight flows are going between some main aggregated zones in Norway, information from the Lorry counts in Norway from 1993 to 2000, the survey among coastal vessels from 1993 (adjusted up to 1999-level) and statistics from CargoNet about delivery pattern and volume on railways in Norway, were used to establish mode specific OD matrices. The main problem by using these surveys directly to establish base matrices for the model was that they do not have information about where the transport at the specific mode started and ended. The patterns are only equal for goods being transported directly without any change of mode.

1.4.5 SAM matrices, PINGO

The base year SAM matrix represents equilibrium in the Norwegian economy. Columns of the matrix represent the economic agents' accounts while its rows represent equilibrium on the markets for goods and factors of production. Transport of each commodity within each county and between all pairs of counties is represented on the off-diagonal sub matrices of the SAM. Positive elements in the columns are output of goods or endowments of factors of production, while negative are input or demands. This implies that in the case when the SAM matrix represents equilibrium in the economy, its rows and columns sum up to zero.

National Accounts Statistics report the gross production and the import in terms of producer prices[4] (18 values), whereas the demand is valued in market prices (18+19 value). The supplies and output part of the economy is valuated according to the basic value (10 value) which means that VAT, profit and taxes/subsidies and trade margins are kept out, whereas the demands and input part of the economy is valued in market prices (18+19). Hence, the two parts of the economy are calculated in different value set.

The different value sets have the consequence that rows in the SAM matrix for the economy do not sum to zero. Moreover, according to statistics Norway, industrial enterprises receive positive profits from their production, which have the consequence that columns in the SAM matrix for the economy do not sum up to zero.

In order to be able to use the SAM matrix received from statistics Norway to calibrate the PINGO model, which is based on the assumption of perfect competition and full balance on the markets for goods and production factors, there is a need to introduce additional goods and agents allowing for balancing the SAM matrix. In this case balancing means to ensure that all rows and columns of the base year SAM matrix sum up to zero (these additional goods consist consequently of trade margins, services, transfers, VAT and other indirect taxes).

Export and *import* sectors are used in PINGO in order to balance commodity flows in the economy i.e. rows of the SAM matrix. The export sector is responsible for consuming part of regional commodities exported abroad, while the import sector is responsible for

---

[4] One of the purposes of working out National Accounts estimates is to use the estimates in model. In order to do so, one has to calculate all the estimates as "basic prizes". That means taxes and trade margins are removed to calculate the production or consumption in basic values.

Value set in NA	**Interpretation of the value set**
18 +19	Market value (observed value)
19 = 14 + 15 +16 +17	Trade margin valued in market prices
18 = 10 + 11+12 +13	Producer value or imports value, valued cif (cost, insurance, freight)
10	Value in basic prises after all components are deducted

producing commodities imported to Norway. No exported goods are delivered by the trade sector in Norway as differently to other larger economies. All the exports leave the establishment that actually has produced the commodity.

An *operating surplus commodity* is used to balance the sectors accounts (columns of the SAM matrix). The operating surplus commodity is county specific and is either produced or consumed by the sectors. Operating surplus is interpreted as input to production when the producers receive profit and as output when they face losses. In addition to profits/losses of the sectors, the operating surplus also represents monetary investments, taxes/subsidies from the government, services and transfers, which are not taken explicitly into account in PINGO.

A *governmental sector* is used to balance the distribution of the operative surplus commodity in the economy (the corresponding rows of SAM matrix). In case when supply exceeds demand, the governmental sector consumes the rest of operative surplus commodity, and visa versa in case when demand exceeds supply. It is also responsible for balancing export and import activities. The governmental sector in PINGO is used just to balance the SAM matrix and does not have any policy instruments in its disposal.

2. Needs for developments and new data to improve the freight model system

In order to obtain data that are valuable with regard to improvement of the freight model systems, it is always important to understand the methodology applied in the models and also which new data that are needed and whether the new data can be efficiently used with existing data.

Possible changes in the current model system should address current weaknesses and support the need for analyses. According to the needs specified in *Work programme 2004 – 2008 for development of multimodal freight models* (NTP Transportanalyser, 2004), we need models both to predict effects on transport demand, i.e., changed need for freight transport, and effects that affect the logistic pathways. The planned logistic module fit well into today's model system PINGO/NEMO, where PINGO is the demand model and where the new logistic module and NEMO handles the logistics and output of important transport indicators.

Experiences with construction of freight models in USA are expressed by Sorrantini (2000) who says that although many states use trend analysis to estimate and forecast statewide truck travel based on traffic counts and Annual Average Daily Traffic (AADT) to estimate future traffic growth factors, it is believed that the demand for freight is better explained when derived from economic activities rather than from traffic counts and projections. In addition, interaction between links is ignored when a trend line approach is applied to statewide truck travel estimation. Sorrantini (2000) further says that: "The national Cooperative Highway Research Program (NCHRP 1983) addressed the need for a freight-oriented planning process with one of the main requirements being the preparation of the freight components of statewide master plans. Also the resulting freight model should use vehicle or commodity flow data as major input rather than vehicle count and or frequency data alone. Thus, freight traffic projections should be based on economic activity instead of trend extrapolation".

The US experience is quite in favour of continuing the maintenance and development of the PINGO model which is used to forecast economic activity and thereof the future freight flows. The US experience also indicate the favour of CFS. As a general principle, we share the US experiences and we also have the philosophy that all business links that make use of transport services should be represented in the base matrices. By business links we meant each link in the delivery chain from producer to consumer (end user of a product) were there are an added-value function and were products are traded. In addition there is a need to represent the transport system were the costs of different logistic solutions between business links can be assessed and were transport chains and the use of different types of terminals are represented. Then we make a difference between firm logistics and transport logistics were the first can be represented in the base matrices, because the turnover is represented in the Industry or Trade statistics (Statistics Norway). The transport logistics are not available in any official statistics. This is because transport statistics only is available for each mode, were transport links not are registered.

2.1 Current principles for representing freight flows

The current base matrices in NEMO are based on register data, for economic activity in the municipalities. The register data comprise different data sources for commodity flows in and out of different business links (see Vold and Hovi, 2004). The base matrices in NEMO are referred to as PWC matrices (producer-wholesaler-consumer)[5]. This is because the current base matrices do not represent solely the transport flow between producer and consumer. The current matrices includes the representation of the transport through the wholesaler link. Thus we have already represented some logistics in the base matrices that are used in NEMO (i.e., we consider the wholesaler link as a part of the logistic chain).

The three-dimensional balancing that spread data for the commodity flows in and out of the business links to a PWC pattern do however not ensure that the total freight flows between pairs of specific business links becomes correct (i.e., producer-producer, producer-wholesaler, wholesaler-retailer, producer-consumer and so on). This means that there can be transports between zones where the agents are wholesalers instead of transports between producers and wholesalers.

2.2 Proposed improvement of principles and new data for representing freight flows

This section discusses possible actions to correct for the fact that the freight flows in the current version of NEMO can be erroneous both in volume on a disaggregated level but also with regard to which business links the transport go through. This section does also discuss the possible gains of applying a more detailed or disaggregated commodity grouping.

2.2.1 Justify freight flows at disaggregated zonal level

The present NEMO model represent freight flows that are calibrated at a relatively coarse level. It is relatively uncertain at the municipality level. Introduction of more detailed zonal level and/or development of regional models on the basis of the national model makes it necessary to check base matrices at municipality level (or in sum for smaller groups of municipalities) on the basis of other and reliable statistics. The main problem is that during the development of the freight model most of the available statistical sources are already used, and the lack of evaluation data are often considerably.

2.2.2 Improved methodology for transport between business links

To eliminate the type of error with regard to which types of business link the transport go between (explained in section 2.2), register data can be split in the way that deliveries between alternative pairs of business links are represented in different sub-matrices, such that we obtain separate and non-overlapping datasets for each pair of business links (i.e., from producer to wholesaler, from producer directly to retailer, from wholesaler to producer etc.). The problem is, however, that we do only know the share of trade between

[5] C in PWC is for domestic consumption (C=consumption, I=investments, V=input) and export (consumption abroad). Earlier for import, every commodity group had a CIV distribution in the external trade statistics. The reason was to get knowledge about which parts of the import were increasing.

the business links within the trade sector (i.e., wholesaler-retailer). One lead is for instance that larger input volumes of a specific commodity to a producer will most probably go directly from the input producer. Also, commodities for final consumption are usually subject to more frequent use of wholesalers than bulk commodities and fresh food.

A commodity flow survey can be a source were information about delivery pattern between different business links can be collected. Eventually the shares between other pairs of business links can be based on qualified guesses and/or simply asking a selection of companies about who they send and receive shipments to and from we may obtain rough estimates for the shares.

A split of the register data into sets per pair of business links can be followed by two-dimensional balancing techniques for every data set to obtain separate OD matrices for each pair of business links. The balanced matrices for each pair of business links can then be added to a complete PWC matrix.

Ideally three-dimensional balancing should be used also on the split register data, but because the calibration data (based on mode specific transport counts) cannot be split in transport per business link, we need supplementary information for this. According to section 2.2.2, however, data from commodity flow surveys may replace the role of mode specific counts in matrix balancing techniques. And if the CFS data include information about the pair of business links the deliveries goes through, then these data can be split and applied in three-dimensional balancing for the pair of business links.

2.2.3 More detailed commodity grouping

The commodity grouping in the operative freight model system are very aggregated. The choice of commodity groups were first of all done under the limitation from STAN of maximum 15 commodity groups per assignment. A more detailed commodity grouping can improve the delivery pattern in the PC matrices. The reason is that by increasing the commodity groups from 13 to say 130 the level of aggregation is more detailed and the commodity groups is more clear defined. Per example; some of today's commodity groups include both typical inputs and final products. These products have quite different delivery patterns, since inputs are mainly delivered directly to industry while final products are delivered for consumption. While the inputs mainly consists of bigger shipments, final products usually consists of smaller shipments with a more complex delivery pattern and thereby more use of wholesalers and distribution terminals. The main gain is two-fold; on the one hand more detailed commodity groups give more clear value to volume ratios, and on the other hand more specific commodity groups, there would give less zones with production or economic activity[6], connected to both delivery and receiving. The uncertainty related to establishing the delivery pattern would thereby decrease.

A more detailed commodity grouping should ideally be accompanied by cost functions that are re-calibrated to fit detailed commodity level. If this is too costly, aggregated cost functions can still be used.

Limitations in STAN of maximum 15 commodity groups per assignment would either be met with an aggregation of commodity groups before assignment or more than one assignment per scenario. If commodity groups are aggregated before running the assignments this will reduce the needs for revising the cost functions.

[6] This would maybe not be the case for consumable goods that are expected to be traded in each municipalities (per example food, clothes, electrical articles and fuel).

2.3 A logistic module

A group consisting of Rand Europe, Solving International, Solving Bohlin & Strömberg and INRO are currently doing a project for the NTP working group and the SAMGODS group, where the objective is to specify a logistic module that can be used in their respective national freight transport models.

The group has delivered a report (de Jong et al., 2004) where they outline preliminary ideas on the structure for the logistic model. The report says that the logistic module should take as input PC matrices and give as output OD matrices that are subsequently assigned in a multimodal real network model. In addition to assessing the logistic supply chains from producers to consumers, the logistic model will also provide matrices with transport costs that can be applied in the base matrix project for Sweden and headed by Inregia.

The interim report considers the development of the logistic model in the context of the current Swedish and Norwegian base matrices. The considerations include the opinions:

- *"Both in Norway and Sweden, the PC matrices do not include ports and airports. They give the producing or consuming county for import or export….The logistic project should therefore handle ports and airports to include these in the OD matrices"*. This is of course desirable, but requires data for the turnover in the ports and airports. There is an ongoing work to improve port statistics in Norway which should make this possible for ports. For airports the data sources can be given from Avinor AS.
- *"Intra zone transport should be generated by the logistic module and included in the OD matrices"*. This can probably be taken directly from the mode specific counts, since there is little chance that internal transport is multimodal. One challenge is however the quality in the Lorry survey at municipality level, because the survey sample is drawn to give unbiased estimates on county level, based on a tree years period.
- *The Norwegian client does not regard the W information in the PWC as very reliable, but if this is requested by the logistic model project, it can be used for development of a logistics model.* The group don't like the idea of developing logistic modules for Norway and Sweden in "separate ways". They consider the wholesale sector an integral part of the logistics decision making in the supply chain. We have commented on this in two bullet point at the end of this section.

The group sets fourth some key options for combining logistics and network model. The A option says that warehouses should be identified as intermediate destinations and added as centroids (i.e., that warehouses will be represented in the OD matrices from the logistic module). The option B says the logistic model could split the tonnes by consignment size, and then let STAN carry out the assignment per consignment size class to the available modes, vehicle sizes and routes. Option C include option A and says the logistic module should cover the decisions on modes and transfers between modes. STAN will cover route choice for the "legs" including empty transport. All the options A, B and C considers assignment of aggregated OD matrices which is one of the groups overall premises that output from the logistic module should be aggregated before assignment to the intermodal network model. One challenge is to get enough data, specially related to internal storage and consignment size.

The group suggest that the logistic module should work on a disaggregated level, but also says "But data availability is the key to determining the unit of observation. This will be investigated in Phase 2 and recommendations will be given". It is possible that this disaggregation is required to build a trustworthy behavioural model. It is apparent, though, that the suggestion of using a distribution of shipment sizes per firm and commodity group is to a large extent based on the assumption that data can be obtained from a CFS. Since a CFS is still not available for Norway, it is probably not realistic to disaggregate this much. A very disaggregated approach also implies that the model complexity increases and thus the cost of development and maintenance of the model system. We think that the gains of disaggregating to the level proposed by the group should be seen in relation to the complexity of other parts of the model system and the model system as a whole and the burden of establishing a very disaggregate logistic module. We realise, however, that some disaggregation is needed which for instance would make it possible to develop a discrete choice model for alternative logistic solutions per shipment size for each commodity group and OD pair (the current model use only the minimum cost solution). Herein, data for the distribution of shipment sizes could be scaled up according to total transport demand (obtained from the PC or PWC matrices), and the distribution could be exogenously altered for forecasts to account for technological and economic development.

The logistic model project group advise that the wholesaler W is removed from the PWC matrices and treated as part of the logistic module and eventually included again in the output OD matrices from the logistic module. The group are aware that the commodities are value added when they goes through the wholesaler business link. They suggest that a constant can be added to the value of the commodities representing transaction costs, profit etc. We agree that the logistic module should be in a position to change the flow shares through the different wholesalers and the location of the wholesalers, but we think it is an unsatisfactory solution to remove the wholesaler link from the Norwegian base matrices for the following reasons:

- It is logical to include transport through all links where the commodities are value added, and to apply the logistic module to assess the logistics between these links. The wholesaler can be considered as a firm. This gives consistency with economic models (e.g., forecast models like PINGO) where economic agents are represented similarly by sectors with input demand and output supply.
- The base matrices are used in the PINGO model. Since wholesalers are economic agents, it is recommended to represent the wholesalers explicitly as separate sectors (today they are merged with producers). To fulfil these, base matrices need to be PWC.

2.4 Terminal activities

An important reason to include a logistic module is to better represent terminal and distance dependent freight transport costs and ensure right transport pattern and delivery links. We may differentiate between intermodal reload terminals (e.g. ports and terminals with railway connection), internal storages, consolidation and distribution centres and wholesalers warehouses. In the wholesalers warehouses the value of the commodities are changing and therefore this activity is represented in the economic register statistic. Other kind of warehouses (company internal) there are not any added value function and therefore not any available information about freight volumes at all[7]. For ports and airports terminals there are available statistics, while information about railway terminals seems to be confidential information within CargoNet AS (approximately freight volumes

[7] Information about size (in square meters) and location of warehouses can be given from GAB, but this source does not have any information if the warehouses are used or empty.

can be given), freight volume through transport operators' terminals[8] can be based on a survey among general cargo terminals in Norway, carried out by TØI this year. The survey would also give information about reload costs and give estimates about time used in terminals. The survey are a surplus to the survey done by SINTEF (Lervåg et al, 2001), were cost estimates for reload costs and time only were averages and not specific regional estimates.

[8] These terminals are similar to distribution and consolidation terminals.

3. National versus regional models

Consistency between the national model (PINGO/NEMO) and regional models can be fulfilled if the regional models are calibrated such that freight flows into and out of the region are consistent with the base matrices in NEMO. Another way of achieving consistency is to calibrate the regional model independently of the national model, and then correct the base matrices in the national model accordingly. The first way of achieving consistency is necessary if data from NEMO are used as exogenous data for transport flows in and out of the region, whereas the last method is suitable if we have greater belief in freight flows in and out of the region based on new data or alternative methods of exploiting the data already used in establishing NEMO. The choice of method of obtaining consistency will depend upon the type of mathematical model structure to be used for the regional models and whether there are new data that contribute significantly to improve the reliability of the transport flows. Crucial in this regard is the choice of zonal level for the national and regional models (see section 3.2). It is also of importance whether base matrices per vehicle type and commodity group directly in the network assignment, i.e., we could assume that there is always uni-modal transport between terminals and business links (tonnes and tonnes per vehicle type) in urban areas. If it is so, the freight matrices can be assigned together with matrices for passenger transport in a network model (e.g., EMME/2).

Service transports are missing in the NEMO model. These transports are mostly shorter transports (at regional level) and the small vehicles shares of business related transport are considerable: From the Lorry Survey in Norway 1993 to 1999[9] freight vehicles with maximum payload up to 3,5 tonnes constitute of nearly fifty per cent of the number of tours, almost forty per cent of the number of vehicle kilometres, but only five and less than tree per cent in number of tonnes and tonne kilometres, respectively. The need for including small freight vehicles is therefore considerable at regional level. Hereunder comes the challenge between the division into passenger and freight transport.

Transport of goods from retailer to consumers is covered as part of the shopping trips in models for passenger transport.

3.1 Model capabilities and data requirements in the DISTRA model

In 2003 there were carried out a pre study in Sweden were the purpose was to outline a feasible way ahead for the development of the Swedish national freight model system with focus on modelling local and regional distribution and collection traffic. The Distra approach implies an extension of the present SAMGODS[10] model in three dimensions: 1) detail, 2) kind and 3) scope. First, Distra will add transport data on a low spatial level. Secondly it will include service transports, thus transforming SAMGODS to something more than a freight model. Thirdly, the additions will make it possible to solve problems

---

[9] From 2000 vehicles with a payload under 3,5 tonnes are excluded from the survey.

[10] The SAMGODS model is the national freight model in Sweden, and have a quite similar structure as the NEMO model.

that can not be handled by the SAMGODS model. The aim of DISTRA is to get goods flows and other business transports not included in the regional passenger model in Sweden (SAMPERS) on the same zonal level as the SAMPERS model and assigned in the same detailed road network.

The input and output of the DISTRA model in a forecast situation could roughly be summarised as follows:

Input:

From SAMGODS:	Intramunicipal goods flows Intermunicipal goods flows on truck
From SAMS-data:	Socio-economic data on zonal level (disaggregated from RAPS data to smaller local units)
From SAMPERS:	Road network Car matrices for passenger transport

Output:
Goods and service transports on different types of vehicles to the road network giving traffic work and transport times and costs.

There are 81 Distra-regions in Sweden, based on labour-regions, and consists of in total 6000 zones, while Samgods have 288 zones.

The Distra model is intended to consist of tree main parts:

1) collection and distribution traffic
2) time-of-day module
3) rail transport module (terminal)

The main development of the Distra model are intended to start with the Samgods model, were intra municipal freight flows have to be disaggregated and assigned to the Sampers network. Freight flows at municipality level have to be validated and adjusted before disaggregating.

Data requirements for the DISTRA base module consist mainly of labour marked, data that already exists and data from the Swedish Commodity Flow Survey. There are drawn up a plan for collecting additional CFS data, which is to:

- Extend the CFS to cover local distribution of commodity flows in a better way.
- For service transports the local units in the service sectors must be handled specifically. This could be done with some smaller surveys at specific geographic areas or sectors.

3.2 Zonal level in the national and regional models

Domestic transports

The current zonal level in NEMO consists of the 435 Norwegian municipalities, one zone per country in Europe[11] while the rest of the world are represented by one to three zones per continent. If the model should be more disaggregated there are in principal five different levels of aggregation to start from, and were actual data are available:

1. A zonal level corresponding to basic statistical units, which is the smallest geographical unit, each with a separate number and name.
2. A zonal level corresponding to postal codes (or an aggregation of postal codes).

[11] Except for Sweden, Finland and Russia, were the countries are represented by respectively four, two and two zones to reach reasonable modal split at the Norwegian border.

3. A zonal level corresponding to the zones used in NTM5[12].
4. A zonal level corresponding to Norwegian municipalities.
5. A zonal level corresponding to Norwegian counties

Data availability rise with level of aggregation: On level 1, only GAB data are available. On level 2, aggregated GAB data are available in addition to trade and manufacturing statistics. Level 3 have availability to aggregated datas from level 1 and 2. Level 4 (municipalities) is the zonal level in the operative national freight network model, and is the level where the trade and manufacturing statistics are available at a lowest cost. In addition all primary statistics and mode specific statistics are available at municipality level. Level 5 (county level) is used in Pingo, where National account on county level are the primary statistics.

The number of zones corresponding to the different levels of aggregation is presented in table 3.1. A recommended level would be municipality level in National models and NTM5 level in regional models. To comply with the NTM5 level, data sources at the municipality level need to be disaggregated, whereas data sources for the basic statistical unit and postal code need to be aggregated. Both the basic statistical unit level and postal code level gives a zone structure that is needlessly detailed related to freight transport. But it is sometimes required to assign freight flows together with passenger flows in the same transport network representation to investigate the total capacity utilisation. This means that it is sometimes necessary to expand the zone level from municipality to NTM5 level.

Table 3.1. Number of municipalities, basic statistical unit and NTM5 zones per county.

		Number of municipalities (= number of zones in NEMO)	Number of basic statistical units	Number of Postal codes[13]	Zones in NTM5	Ratio between number of zones in NTM5 and NEMO
1	Østfold	18	689	229	40	2,22
2	Akershus	22	1 366	219	78	3,55
3	Oslo	1	552	661	26	26,00
4	Hedmark	22	799	153	50	2,27
5	Oppland	26	773	172	64	2,46
6	Buskerud	21	685	183	69	3,29
7	Vestfold	15	580	180	63	4,20
8	Telemark	18	595	186	49	2,72
9	Aust-Agder	15	366	114	38	2,53
10	Vest-Agder	15	446	141	52	3,47
11	Rogaland	26	940	274	110	4,23
12	Hordaland	34	1 106	455	137	4,03
14	Sogn og Fjordane	26	488	191	73	2,81
15	Møre og Romsdal	38	764	275	151	3,97
16	Sør-Trøndelag	25	829	310	104	4,16
17	Nord-Trøndelag	24	482	125	69	2,88
18	Nordland	45	1 065	364	138	3,07
19	Troms	25	585	231	78	3,12
20	Finnmark	19	310	121	39	2,05
	Sum	**435**	**13 420**	**4 584**	**1 428**	**3,28**

TØI-report 731/2004

[12] NTM5 – The Norwegian **N**ational **T**ransport **M**odel, i. e the national model for Passenger transport in Norway.

[13] The number of postal zones in the table also include post box numbers. These are not actual to use as zones in any model, because they don't have any information about location for enterprises.

Foreign transports
A suitable zonal level outside Norway in the national model is NUTS2 (corresponding to Norwegian provinces), were the source could be the database from the SCENES[14] model. An exception is Sweden which is the most important trade partner for Norway. It is our opinion that county (län) is a more appropriate zonal level for Sweden.

Norway is however only represented in the SCENES model by one external zone. Information from the SCENES database has therefore to be combined by the Norwegian external trade statistics, and entropy models can be used to ensure delivery pattern between the detailed zones.

3.3 Disaggregation principles

This section discuss the gains of -, and data and methods requirements of disaggregating the zone level.

3.3.1 Gains of disaggregating the zone level

Disaggregation of the zone level can make it possible to obtain base matrices with more detailed information of transport flows. This would improve the capability of NEMO as a tool for investigating the exploitation of the transport infrastructure. Moreover, disaggregation is necessary in order to make the freight model consistent with the national model for passenger transport. Consistency with the national model for passenger transport is required in order to assign OD matrices for freight and passenger transport traffic onto the same network. Simultaneous assignment would give the opportunity to assess the total traffic load on the transport infrastructure, and include congestion

3.3.2 Methods of disaggregating the zone level

The level for register statistic used in establishing the base matrices in NEMO are at municipality level. To obtain these data at an even more disaggregated level, it is often necessary to apply proxy indicators to spread the statistic to the more disaggregated zoning level. The use of indicators to spread data to a more detailed zone resolution is common. An example is the work by Sorrantini (2000) used the US CFS 1993 together with Input-Output (I-O) coefficients to generate truck flows for the state of Wisconsin. He derived heavy truck mode production and attraction rates at the county level based on employment for 28 economic sectors. The truck trips at the county level were disaggregated to the traffic analysis zones (TAZ's) using population as a disaggregation factor. Sorrantini (2000) emphasises that employment would be a more reliable factor for disaggregation, but that there is no information available on employment by economic sector at the TAZ level. Consequently population was used as the disaggregation factor.

Moreover, Huang and Smith (1999) developed a truck travel-demand model for Wisconsin using 1993 CFS data. The statewide travel demand model for this area is based on 624 traffic analysis zones (TAZ's), with additional 50 external stations on Wisconsin's border, whereas the basic US CFS 1993 were developed initially at the county level (72 freight zones). Also in this work the county-level trip tables were expanded to the TAZ level based on population.

[14] SCENES is an European network model for freight and passenger transport.

It is evident that the indicators applied for disaggregation of freight flows should be carefully selected to make sure that there is a good as possible correspondence between the magnitude of the flows and the indicators.

In the NEMO freight model we may start out with the marginals in the national model. Indicators can be established based on register data from Statistics Norway that can be available on postal zone level (see chapter 4.1 for the trade and manufacturing industry). For economic activities were register data is mot available at such a coarse level, these sources can be supplemented by GAB-data, giving information about floor space after main use (e.g. warehouse, shopping centre, business building, etc) on basic statistical unit level. Relevant indicators from the work of establishing the model for national passenger transport NTM5 (1428 zones) include:

- Number of enterprises and employers divided by main industry activity
- Size of basic statistic unit (in square km)
- Population by gender and age

These data can be supplemented by information from the register of Enterprises and Establishments (Statistics Norway) about sales value by main business for each of the 1428 zones. For spreading the flows to a lower zonal level. The appropriate indicator will depend on which freight flow we are looking at.

To derive delivery pattern from the more detailed data for goods production and attraction, marginals gravity models can be used in a manner equivalent as in today model.

3.3.3 Disaggregation by time

The freight flows in both NEMO and PING is annual values. In regional models congestion is a more actual approach. Per example the DISTRA model have a time of day module, were passenger traffic is collected from the SAMPERS model. Most data are available only on annual level, but traffic counts made continuous by the Norwegian road administration makes it possible to derive time of day possibilities for road traffic. The freight volumes have firstly to bee derived from annual to daily flows. If passenger and freight traffic are represented into the same network representation, congestion analysis can be done.

4 Data availability for regional models

In order to move the focus from long distance freight transport represented in the national models to smaller regions, often dominated by road traffic (distribution and collection traffic), we face the problem that data for this type of transport is scarce. To get more knowledge about shorter transports in general and the logistics in the larger cities in particular, there is a need for data about these transports at a geographically detailed level. At the moment we have matrices for 13 commodities at the level of municipalities, to make a model at a regional or local level we need such matrices for a more detailed zone structure.

4.1 The Trade Statistics and the Manufacturing Statistics

In Norway the economic statistics are quite good. In the national model (NEMO) we have among other statistics used the *Trade Statistics* (the wholesale and retail trade statistics, giving turnover in value terms) to establish marginals (input and output) for the "NEMO-commodities" at the municipality level. It is however also possible to get data from the Trade Statistics for the input and output of the trade sector by *postal zones* (for disaggregated NACE sectors[15]). In Oslo, with a population of 550 000, there are for instance more than 400 postal zones[16], so this is rather detailed information. In the county surrounding Oslo, Akershus, with almost the same population as Oslo, the postal codes are not that detailed, around 100 postal zones.

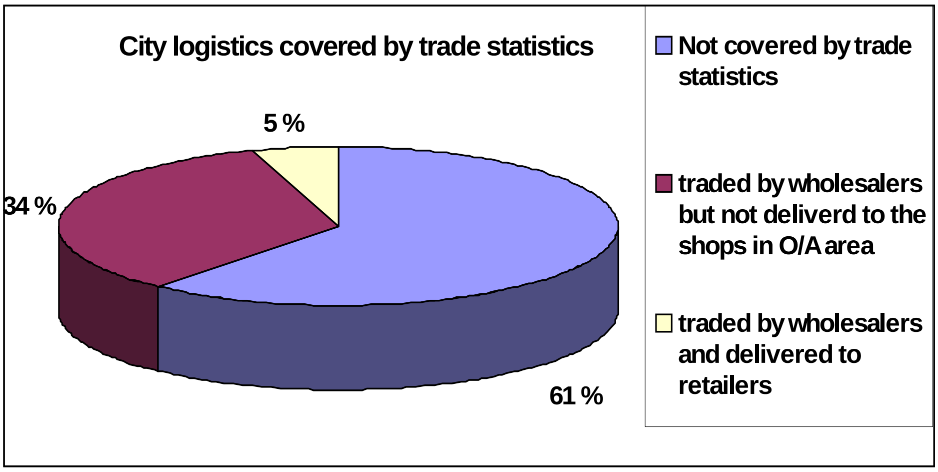
Of all the transported goods in Oslo and Akershus, the Trade Statistics cover approximately 39 % (5 % by retailers and 34 % by wholesalers). (Figure 4.1). These figures are deduced from Statistics Norway's Lorry Survey, which is conducted every year among the 41 000 registered lorries in Norway. The Lorry Survey gives information about origin and destination municipality for each trip, but not the more detailed localization of the trips (e.g. where the trips start and end within the municipality).

Based on the data from the Trade Statistics, the tons traded by the wholesalers and the demand from the retailers could be calculated. The conversion from the turnover in value terms to tons has to be done through the use of information from retailers' and wholesalers' margins, together with information from the External Trade Statistics for similar commodity groups as in the Trade Statistics (the External Trade Statistics have information on both weight and value of the goods, and can be used to calculate conversion factors).

[15] Nace-standard adapted for the Norwegian Economy based upon the EU standard NACE Rev.1. The Norwegian NACE classification consists of 131 groups divided in the given NACE division 50, 51, 52 at the lowest level (4- or 5- digit level).

[16] The difference from the level in table 3.1 is because post box codes are subtracted.

Figure 4.1. What is covered by use of the Trade Statistics of the transported goods reported by the Lorry Survey in Oslo and Akershus



TØI-report 731/2004

When we want to spread the data for the wholesalers' and retailers' trade (the input and output) to deliveries between them (e.g. spread out in a matrix), some assumptions have to be done. If we for instance are studying a single commodity group with 100 retailers and 20 wholesalers located all over Oslo, we first make the assumption that all of the wholesalers (20) have deliveries to every of the retailers (100). The amount of a commodity delivered from a wholesaler to a retailer is according to the wholesalers market share for the commodity. Afterwards we do a cut off (in terms of the value of the delivery) by assuming that the smallest of the 100 retailers do not get deliveries from all of the 20 wholesalers. This seems to be a realistic assumption. This method will however work best in areas having a certain amount of wholesalers, normally the urban areas. In other areas the assumptions may not be that reasonable.

Contrary to the Lorry Survey (as shown in figure 4.1 for the Oslo/Akershus area), the Trade Statistics does not cover the major part of the tons transported in an area. One way to cover a larger part of the freight transport is to work out similar data sets (input and output) for other sectors using freight transport. One important data source could be the *Manufacturing statistics* produced yearly by the Statistics Norway that also can be available at the level of postal zones. Unfortunately we will in some situations face a problem of finding the correct location of the establishment from this statistics. If the establishment is part of an enterprise with many establishments, only the address of the main office is given. This address is of minor interest in cases where none or only a small part of the production is done at the main office.

If an enterprise organizes more than one establishment, then the Manufacturing Statistics gives the establishments' turnover, but not the number of tons. The ratio between these is only given at the enterprise level, not for each establishment.

One way of solving this problem is to look at the *External Trade Statistics* (see next section), since this statistics in principle contain value and tons for every entry published in the statistics. There will however be substantially variation in the value/volume ratio even for the most detailed commodity level.

4.2 The External Trade Statistics

A data source for the geographically located freight distribution connected to foreign trade is the *External Trade Statistics* (assembled and processed by Statistics Norway). In the official statistics the information about the flows in import and export are given at the county level in Norway (production county in export, county of custom clearance in import), but in principle it is possible to get the data at a more detailed level (in NEMO only data at county level are used).

In the External Trade Statistics, all deliveries can be identified by codes that determine the type of commodity, the value of the delivery, the number of tons and, for import, the *organization number*. The organization number makes it possible to know the exact localization of the importing company. But, the same weakness as was explained for the Manufacturing Statistics also occurs in this case: When the enterprise consists of several establishments, we only get the address of the main office (which is often not the destination of the transport). In some cases we also face this problem when using the statistics at the county level (the main office may be located in another county than the establishment that receives the goods).

The method of using the organization numbers to locate the destinations of imported goods has been used twice for the imported commodities over the Port of Oslo. As mentioned earlier we are not confident that we always get the address for where the commodities are actually transported, it may be just an administrative address. Still, it is much better than not having this information. In any case we do not get the information on whether a terminal (e.g. Alnabru) is used or not. It will then be necessary to make some assumptions.

Larger manufacturing plants in Norway are most usually located by the sea in order to use sea transport, often without any local transports in the road network. If one gets export data from the External Trade Statistics broken down on aggregated commodity groups, the production county, and by the Regional Custom area were it is customs declared, it is often possible to trace the transport. It may also give some indication to look at the value/volume ratio of exports by sea from the regional office closest to the plant.

Imports are more difficult. But due to the open Norwegian economy and to the fact that most of both consumer goods and machinery and equipment are imported, it is important to address the problem seriously. Imports entering a major port will most probably be registered (customs cleared) in the port, but not always. If it is possible to trace the importer by the organization number of the company, we will know the localization and hence the destination of the transport. The External Trade Statistics unit in the Statistics Norway has developed statistical tools (programmes) to process such data. By use of these tools we can get a broad picture of where in Norway the customers of e.g. a port are located.

A problem with the imports arriving by road, is that Svinesund in Østfold county have a very high amount of the customs clearances. The user may well be a wholesaler located in e.g. Oslo or Akershus counties. This problem will in many cases be solved by getting the organization number, and hence the localization by postal code, of the importing company. Postal codes can be mapped to basic statistical units or more aggregated zonal levels. There is still the pot-hole, however, that the localization/address of the importing company is sometimes the main office which can be different from the localization/address to where the goods are actually delivered.

4.3 New technologies make new possibilities?

Tracking by use of satellites (GSM transmitters) and introduction of PDA (personal data adapters) and mobile phones based on GPRS technology (first generation of mobile wideband) have opened for use of electronic consignment notes without a too high investment costs and therefore also available for smaller transporters. Transport operators supply a higher level of logistic services. It is more and more common that commodities have to be paid at delivery. Mobile payment terminals make it possible to demand payment by delivering, and thereby purchase on credit is soon a transmitted area.

Tracking, electronic consignment notes and mobile payment terminals are opening for new possibilities for electronic statistic production.

4.3.1 Transport operators

Electronic consignments notes have opened for a new data source, namely electronic data extraction from the transport operators data systems. We have been in contact with two of the main transporters in Norway with national covering networks. Both confirm that their data system give availability to information about freight movements on consignment level.

The available information is covering consignments weight and value (net prices), consignor's and receiver's postal codes. Terminals used during the transport chain are also registered.

Information is available since approximately year 2000, and can be given for any time period (year, quarter, month, week). Uncertainties are rising with age of data.

There are some imperfections for making a covering system: Commodity groups used by the shippers are 1) parcels, 2) general cargo, 3) terminal considered consignments, 4) consignments not considered in terminals, and finally 5) hazardous goods, and is far from the economic activity where it is produced or traded.

If consignments are deposited or collected by the receiver; only delivery or collection terminal are registered. Mode of transport are not registered (there is no link between mode and consignment), but modal split between road and rail can approximately be identified through fixed transport routes. There is not any available information about truck size used and thereby number of trips from operators systems.

4.3.2 Seaborne transport

AIS (Automatic Identification system)

AIS is first of all an anti-collision system, but can also be used to give information about position and movements of vessels. Primary covering area are the Norwegian coast to the territorial limits. Range of applications of the AIS are:

1. Anti collisions system for the vessels themselves
2. Supervision of the seaborne traffic
3. Tools for statistic production about traffic movements in the fairways

Through the IMO number data and information via AIS can be connected with information from the port statistics, ship's registres (NIS, NOR and Lloyds) and other sources were IMO number is a key variable.

The progress of the AIS is:

- 35 base stations are planned build in Norway in 2004
- Within a period (not yet defined) all vessels should have AIS installed
- All tankers must install AIS by the first reconditioning of the safety certificate
- At least 31. December 2004 all other vessels in international traffic have to install AIS
- At least 1. July 2008 all vessels in national transport have to install AIS

Thereby it seems to be some years before AIS is a possible source for data production. The linkage between AIS and the Port statistics between the IMO number seems to be a key for information about freight flows and delivery pattern for seaborne transports. The port statistics is, however, not covering all seaborne transport in Norway, but a major part of the tons are covered. Commodity level in the Port statistics is approximately 25 groups for the 25 ports who are classified as PortWin harbours. A problem is related to container transports were commodity is unknown. For the rest of the public traffic harbours (45) not classified as PortWin harbours, commodity is not specified.

Tracking

Satellite tracking of all Norwegian fishing vessels larger than 24 foots were introduced 1st July 2000. This meant that all Norwegians fishing vessels larger than 24 foots should send position, course and speed once a hour without regard to wherever they are. This is automatic done by use of satellite communication system.

In addition there are entered mutual agreements with EU, Russia and Iceland about exchange of the tracking information. This source would not give any historical picture since the tracking only have been done since 1. July 2000.

Information about vessel movements would never have information about freight flow volumes. Therefore this information seems to best suited as a calibration foundation for vessel module for fishing vessel movements in NEMO.

4.3.3 Business sectors

Referring to chapter 2, where we concluded by the most satisfactory way to derive freight flow pattern is with point of departure in economic activity. Especially with respect to the purpose of forecasting, economic activity seems to be the best basis. Therefore a more satisfactory method is to establish base matrices seems too be with basis in business sectors.

The forest industry did early made a common system for logistics and invoicing. The system is operated by a company called Skog Data AS, and is based on a system were each driver have their own PDA in the trucks with availability to detailed information about route (illustrated in maps on the PDA) and consignments are given. Skog Data then have a quite detailed overview over all timber transports from the forest to the timber industry, were location (both for collection and delivery) are available at coordinate level. Since timber transports are limitated by the roads carrying capacity (permitted axle load), rote planning and load per trip are detailed planned in the matter of minimizing costs. The result is rather detailed information about transport volume and route choice.

The data are not free, but a data field can be ordered from Skog Data to a cost of approximately 30.000 NOK (in 2004).

As far as we know this is the only business sector were all transports are entered into a common system.

4.3.4 Conclusions

The main conclusion is that new technologies made new possibilities to availability of data. The maybe most exiting possibilities is related to the AIS for seaborne transport if a linkage can be done to the port statistics. The port statistics is however not covering all seaborne transport in Norway, but a major part of the tons are covered.

For timber transports there are also available data with a much higher level of quality than already used in the freight model. The drawback is that there are a relatively high cost connected to the availability. Timber transports are however important commodity groups because of the limitation related to the need for investments in the roads carrying capacity.

Statistics based on electronic data extraction from the transporters data systems, are more difficult to establish a covering picture. This is caused out of several reasons:

- There are a very high amount of transport firms in Norway, and therefore a challenge to draw a representative sample, preliminary only bigger transporters have tracking systems.
- Detailed information from transporters data systems are very close to what is confidential information within each transport operator, caused out of trade secrets.
- Information is not available at commodity level, and would only be available for hire transport and not transport on own account.

The main conclusion must be that in the first round these data are most interesting for building a regional model, because they give information about transport pattern between zones at a more coarse level than any other available source does.

In Statistics Norway it have become more usual to focus on reducing the task burden for the recipients. One step is to start with electronic outcomes from the transport operators systems to substitute for the heavy questionnaire in the truck surveys. One point of departure would be to start with the four main transport operators n Norway with national covering network (Linjegods, Tollpost Globe, NorCargo and DHL).

4.4 Possible uses of PANDA

A model for prediction of demand changes at the regional level can probably be based on the methodology that is used in PINGO. However, while data for PINGO can easily be obtained from the National Accounts by County (produced by Statistics Norway), it will be more challenging to establish a complete data set for a regional version of PINGO. The national accounts by county is an important source, but regional models would need a more disaggregated forecast tool, and hence more disaggregated data to estimate and calibrate the models. This section takes a look at the possibilities for exploiting existing modelling tools (PANDA) and data sources to obtain the required data for the development of regional models.

4.4.1 General about Panda - How does the system work? Distance and migration

The PANDA model consists of two main modules, an input-output module for prediction of production/employment and a demographic module for population growth. A region is defined as a number of municipalities aggregated together (the number of municipalities in a region can, in principle, range between 1 and 435). Employment and location of population, respectively, can also be predicted on the municipality level.

Distance is introduced into the model only as a part of the demographic module. Distance is also used for the people in an existing location's choice between commuting and migrating. When distance from their existing location and jobs increase, more and more people tend to move to the municipality of work or the neighbouring municipalities. Distance is, in other words, used for

1. dividing located people into migration and commuting, and
2. locating the migrating population into primarily the municipality of work, or into neighbouring municipalities according to the housing market and distance.

Distance in PANDA, in other words, influences the location of population. The system gives us figures, not only for the location of the population, but also for commuting.

An entropy model locates the population in municipalities according to the distance between place of work (which we know from the location of industries) and (the potentially nearest) place of living. The location of people in municipalities according to the entropy system is restricted by the housing market – households are located primarily to the municipality where they work, and then to the neighbouring municipalities according to distance as the houses are filled up (Retired people move according to exogenously given migration rates and unemployed people move according to the migration rates, and whether they are willing to move to get a job outside the region). As we know the localisation of people at the start of the simulation, this stabilizes the system.

Industrial development

The basic economic data in PANDA are collected from construction and use of matrices based on national account statistics by county. Production is driven by exogenously determined growth rates of final demand. These growth rates can be collected from different sources (like the Long Term Programme), or the model user can decide them. The growth rates of final demand (except private consumption and sector investments) give impulses into the regional economy, and an input-output system based on flows of commodities between sectors within the region, and on endogenously determined income and private consumption determines production levels of all production sectors and the commodity flow. Prediction of commodity flows in PANDA is independent of distance. Fixed input and output coefficients are estimated from the construction and use tables.

RPC coefficients

PANDA determine and use Regional Purchasing Coeffecients (RPC) to estimate the outflow and inflow of commodities in municipalities in a region. The RPC coefficient could potentially be applied with PINGO and/or NEMO to adapt these models to predictions at the regional level. The rest of this section explain how the RPC coefficient is constructed and make suggestions for how to exploit the RPC in order to make it possible to adapt PINGO to a more disaggregated level. The RPC coefficient could also be used as a complementary source when constructing the base matrices, by using the information about zone internal consumption share before running the gravity model.

The RPC[17] is the proportion of total demand or absorption, which has its origin from production industries within the region itself, or RPC = Pr/RAr (where Pr is the production in region r and RAr is regional demand). Production in the region (for each sector) is calculated by assuming constant labour productivity in all municipalities within

[17] Here: The supply-demand ratio is interpreted as RPC

the county. This means that the region's share of production is equal to the region's share of employment in the sector, which can be written as

Pr = nr * P   0 < nr < 1, sum over nr = 1

P is total production, while nr is the share of employment in region r (for all sectors).

Regional absorption, or demand, consists of several components (for each commodity):

RAr = Ir + FDr = Ir + PCr + SCr + KCr + Jr + Er

Ir is intermediate deliveries, which consists of deliveries to all other production sectors inside the region, and FDr is final demand. FD consists of private consumption (PC), government consumption (SC + KC), investments (J) and exports (E). When the footnote r denotes the whole county, information can be collected directly from the Regional National Account, and RA and all its components are observable. For lower regional levels, demand figures have to be approximated, using indicators to scale down the county absorption.

- Private consumption (PC) is calculated for each municipality, based on income figures from the tax statistics.
- Government consumption (SC and KC) is approximated by using employment in central (S) and local (K) government sectors, respectively, as indicators.
- Investments (J) are calculated using an investment matrix on county level (of delivering and purchasing investment sectors), by applying sum of employment in the purchasing investment sectors as an indicator.
- Exports (E) are calculated by using national share of exports on detailed commodity level (much more detailed than in PANDA) and aggregating together to PANDA sectors.
- Intermediates (I) are more complex. They are calculated by taking the minimum of on the one side, the sum of each sector's demand of intermediate deliveries of a commodity on the county level (using input coefficients), and on the other side of the share of production delivered to intermediates on the county level. This implies that the regional share of intermediates will never exceed the county share of intermediates.

These five points illustrate how demand (absorption) is estimated in PANDA and which data that is used. For each sector, this leaves us with

RPCr = Pr/RAr = P * nr / (Ir + PCr + SCr + KCr + Jr + Er)

If these calculations are made for each municipality in a county, and then added together, we will observe consistency problems since there is no control on the sum of RAr, and because the single region model does not take repercussions from other regions in the same county into consideration. If the RPCs are the same in all municipalities, there will be no inter-regional trade within the county. This matter can however be discussed further.

The statistical information applied for constructing and using tables for the RPC's are collected from the Regional National Accounts (the RNA), or rather the National Accounts by County. This is the same source that is used to construct the Social Account Matrices (SAMs) in PINGO. There are a couple of properties to these data, and to the use of them in PANDA, that deserves to be mentioned:

1. PANDA is a single region model, not concerned about development outside the region that is under discussion (except growth rates for final demand of 'exports to rest of country'. Regional Purchasing Coefficients (RPCs) show the region's share of purchases of each good.

2. The RNA information is used for constructing RPCs.

Since RPCs are based on RNA, we need additional information to construct RPCs for non-county regional levels.

PANDA is a single region model, developed for analysing regional development for regions consisting of municipalities located within the same county. During the past couple of years, PANDA has become a model that can be operated in regions across county borders. We can now, in principle, choose any combination of municipalities, irrespective of location, and aggregate these into a single region. Since economic data only exist on the county level, we have to construct RPCs using different indicators, depending on which regional level we operate the model. The construction of RPCs is built into the model, in the sense that the user does not influence the estimation procedures. New coefficients are estimated each time a new project is generated in the system, depending on the region chosen. There are three types of RPCs that can be used in PANDA, each of them with its own estimating procedure:

1. *If the region consists of exactly a county*: The construction and use of tables are not altered and the figures are used directly to estimate RPCs.
2. *If the region consists of some municipalities in one county*: In this case, non-survey methods are applied to the county figures, which are broken down to the sub-county level (the analysing region).
3. *If the region consists of municipalities across county borders*: Procedure number 1 and/or 2 is used, together with other available information (some surveys on interregional trade have been made, but these are not precise), and RPCs are estimated according to the procedures described by Stokka and Vik (2000). Information on transport flows is not used.

We have pointed out that the empirical information on flows of commodities in PANDA is restricted to the make and use of tables from the RNA. This implies that *there is no information on flows of commodities between zones at the municipality level*. There is no flow of commodities between the region and 'outside', and equally important, no back-flow of commodities into the region.

In addition, the commodities have been transformed, so the input-output structure of PANDA is sector-sector (not commodity-sector).

Since no information on flows of commodities exists below the county level, and we have to use non-survey methods to construct RPCs for lower regional levels, we need to have some kind of information in the form of indicators. In PANDA, we know employment by production sector on the municipality level. This information is collected mainly from the AA-register. By applying this information, and assuming that productivity is similar in all municipalities, we can construct production figures for lower regional levels. The RPCs can be calculated in several manners. It is important to take into account the size of the new region, as well as the industrial structure in the new region (compared to the county), which is done in PANDA.

There is no information on the weight (tonnes) of commodity production and flows in PANDA. The only measurement is in NOK (value).

4.4.2 Application of RPC's for developing of NEMO and PINGO

It is recommended to take a more detailed exploitation of the opportunities that is in data availability from the PANDA model to be used in the NEMO and PINGO models. The exploitation should involve experts on both data sources (NEMO and PANDA). It is clear however that the RPC's can be applied for the purpose of subtracting regional purchase (i.e., municipality internal transports) from the marginals before the gravity model is run when establishing delivery pattern in the base matrices in NEMO.

The input-output tables and the RPC's in PANDA that are obtained by applying indicators to spread the economic activity to a higher zonal level can probably be used to refine the SAM matrix applied in PINGO. If regional models are established there would be a need for also refine the PINGO-level, because with a regional view in the freight models forecasts at county level is too coarse. PANDA can most probably be an actual source for giving the necessary data that would be needed for developing a forecast model at regional level.

As for NEMO, it is necessary, however, to obtain additional data that represents the geographical commodity flows within the new zonal level[18].

18 The new SIP on regional development that has been proposed by TØI will discuss this further (on theories, methods, existing data and data requirements, but will not start work on programming a model)

5 Changes in existing data sources

A plan for collection of new data needs to take into account to what extent new data can complement existing data and to what extent the quality of new and old data is good enough to make it possible to establish models where output results are within acceptable errors.

Some statistics that have been carried out for a number of years have gone through some changes that lead to lower quality of the data. In the following sections we describe the main changes in the statistical sources.

5.1 Lorry surveys

The Lorry surveys have been carried out continuous since 1993 by Statistics Norway. Before 1993 the survey were carried out every fifth year. The survey consists of a sample of 2 000 vehicles each quarter out of a population of 41 000 vehicles, with minimum payload of 3,5 tonnes and maximum total weight of 35 tonnes. A new sample is drawn each quarter, based on revised data from the Road Administration's vehicle register. No cars are represented in the survey twice a year. The survey is drawn in such a way that the results should be reliable at county level over a tree year time period.

Until year 2000 freight vehicles with a maximum payload between 1 and 3,5 tons also were included in the survey. The exclusion of these vehicles is a result of EU's ordinance with respect to transport statistics. These smaller vehicles constitute of nearly 50 per cent of the traffic work. It is therefore a weakness that they are removed from the survey, especially for analysis on regional level.

There are plans for chancing the commodity grouping in the lorry surveys to a system called NST 2000, that would replace NST/R. This is decided in Eurostat and the new commodity grouping is closer to industry activity than the NST/R system. The NST 2000 is based on the following assumptions:

1. The criterion for classification of goods are based on the economic activity from which the goods originate. This is the same approach as used in the CPA[19] (1996), where the structure (divisions) of CPA is the same as NACE Rev 1
2. NST-2000 is based on the CPA categories
3. The NST-2000 is not based on the physical form of the goods

The advantages of the NST 2000 system is the close link between commodity groups and economic activities, and thereby closer link between the lorry survey and the economic register data. The disadvantages is in the discontinuity in time series from the lorry survey. There would be possible to made conversion keys between the NST/R and the NST 2000, but conversions keys are never 100 %.

There is a general (international) problem that lorry surveys are underestimating the activity in the lorry surveys. This is caused out of the time-consuming job for the drivers related to fill out the driving diary. Therefore questionnaires are returned empty with the

[19] CPA: Classification of products by activity

argument: The lorry is out of order. SSB are now adjusting the annual mileage by information available from the EU checks were level of mileage are registered from the lorries' mileage recorder.

Electronic data subtraction from the four main transport operators in Norway with a national covering network, can be a start in the right way for correcting for under reporting and sample squewnesses in the Lorry survey.

5.2 Coastal surveys

Until 1993 a regularly survey were carried out among coastal transport. The survey has not been carried out since 1993. Therefore there are no updated data sources for the transport pattern for costal shipping in Norway.

From Statistics Norway we have got information that they in cooperation with the Norwegian Coastal authorities are seeking for financial sources for a new coastal survey in 2005.

5.3 External trade surveys

Norwegian Customs and Excise are reorganized from January 1st 2004. During 2004 a number of customs offices will be closed or merged with other units. Table 4.1 shows the number of custom offices before and after the reorganizing.

Table 4.1. Number of custom offices per county before and after the reorganizing in 2004.

		Number of custom offices	
		Before 2004	After 2004
1	Østfold	6	2
2	Akershus	2	1
3	Oslo	1	1
4	Hedmark	3	4
5	Oppland	2	0
6	Buskerud	1	1
7	Vestfold	4	1
8	Telemark	2	1
9	Aust-Agder	1	0
10	Vest-Agder	1	1
11	Rogaland	4	3
12	Hordaland	3	1
14	Sogn og Fjordane	1	1
15	Møre og Romsdal	6	1
16	Sør-Trøndelag	2	2
17	Nord-Trøndelag	3	1
18	Nordland	11	5
19	Troms	3	2
20	Finnmark	8	6
	Sum	64	34

TØI-report 731/2004

The table shows that the number of custom offices will be almost halved after the reorganizing, and Aust-Agder and Oppland would not have any offices at all after 2004. The effects would be a more coarse External trade statistics. For import we are using the office of toll clearing when making OD-matrices. Closing of toll offices can give as result that when toll clearing at the nearest office, this may be in another county.

6 Recommendations for collection of new data

This chapter gives a summary of the main data needs and recommendations for collections of new data.

6.1 Available, but unused data sources

In chapter 2, 3 and 4 we described data needs and existing data sources, while chapter 5 described changes in existing data sources. It is possible to obtain a lot of information about the commodity inflow and outflow of zones at a more disaggregated level than already used in freight modelling:

- Trade and manufacturing statistics are available at postal zone level at a quite disaggregated commodity level
- The Panda model can give necessary information about economic activity at a suitable level of aggregation (municipality level) to establish forecasts model(s) connected to network models at regional level
- External trade can in principle be transferred domestic from county to postal zone level
- Skogdata have detailed information about timber transports from the forest to the timber industry and timber terminals (information available at coordinate level)

6.2 Missing data

Missing data related to further development of the freight models at both national and regional level are, based on the discussion in chapter 2 to 5, shortly summarized:

1. Data for delivery pattern (both geographical and between business links)
2. Information about consignment size
3. Information about freight volumes through enterprise internal warehouses (inventory)
4. Service related transports
5. Transport costs, especially seaborne transport cost are hardly available

A survey where all these variables are charted, seems to be a good choice both with regard to improvement of the national model and with regard to establishment of regional models. A CFS is the only kind of survey that would satisfies all these demands, but is quite expensive to carry out. Required sample size would increase when geographic level of aggregation decrease.

6.3 Main data needs

Earlier chapters have explored the need for methodological improvement, development of new regional model and requirements for new data. This section put focus on the insufficiencies and concretized methods for data collection.

A relevant question to be asked is whether it should be the transport statistics that should be the important nominator or for instance the statistics for the trade sector. An argument for making use of the latter is that the commodity information is connected to the transport user and not the transport operators. This argument is supported by chapter 2, where there are arguments to use data for economic activity which are to a great extent linked to the transport users.

6.3.1 National model

The national model can be subdivided in (1) base matrices, (2) a logistic module and (3) a multimodal network model with cost functions (STAN). There are data requirements to improve base matrices and the network model and to establish the logistic module. The main data needs were identified in chapter 2 and 3 to primarily be information about the spatial delivery pattern between producer and consumer for each of the commodity groups in NEMO. The missing information about delivery pattern is today substituted by an OD-matrix computed from the lorry and coastal counts supplied with information from the National rail freight operator CargoNet about transported volumes between railway terminals. This gives OD-matrices with information about spatial transport pattern, but because mode specific counts don't have information about delivery pattern from producer to consumer, only transports between reloading addresses (and therefore short trips are overrepresented), this is not a satisfactory foundation for calibrating the delivery pattern in the base matrices. Information about transport costs for seaborne transport is also missing, and today's cost functions for international shipping transport are based on rather weak data. If a CFS is carried out, there is the possibility to replace the mode specific counts by data for the delivery pattern between the producer and consumer.

Moreover, as mentioned in section 2.2.2, a CFS would also make it possible to collect better information about the volume weight relationship of the transported goods and information about the share of the freight that is transported between the different business links (i.e., the share between producers and producers, between producers and wholesalers etc.). The current NEMO version suffer from the lack of such data and could probably be significantly improved by applying such data according to the methodology outlined in section 2.2.2. New data in terms of a commodity flow survey would also become useful, in that this would make it possible to apply the CFS data in the matrix balancing instead of the mode specific counts and then save the mode specific counts for evaluation purpose.

The data requirements for the establishment of the logistic module is still unclear since a final decision has still not been made about the model structure. The plan of Jong et al. (2004) was commented on in section 2.3. It seems quite obvious that a CFS needs to be carried out to obtain the data required to establish the most disaggregated logistic flows (i.e., by commodity type per shipment size and firm). Although a CFS survey would be beneficial also in the case that a more aggregate model (e.g., by commodity type and shipment size), it is less critical that such a survey is carried out.

6.3.2 Regional models

Service related transports are more or less totally missing in today's statistics since from 2000 the lorry survey only covers vehicles with a maximum payload over 3,5 tons. A starting point for information about these transports are the survey of Rideng and Strand (2004), but the survey don't have any geographical level.

Developing regional transport models would in addition to the data needs for the national model, also raise needs for more detailed surveys than the today's sample in the lorry survey, were precision level are at county level based on tree years surveys. Therefore if regional models should be established, sample sizes has to be increased for the study area of the model. A better alternative is maybe to use data extraction from the transport operators computer systems, were information about delivery patter between zones at a more coarse level is available (p ex postal codes). The main challenge is however to get a covering picture of the freight flows caused out of the very scattered transport marked in Norway, with a lot of one man enterprises and a own account share that are considerable but decreasing. Information about commodity groups are totally missing in this data source.

6.4 Different kinds of surveys

6.4.1 Carrier based surveys

For example: Truck and coastal counts in Norway.

Advantages: Relatively cheep to carry out. It is able to capture imports as soon as they enter the country (if the truck is Norwegian registered). Load per trip and commodity are registered in addition to municipality of origin or destination for the transport.

Disadvantages: Shipment characteristics (commodities, value) are often not known, especially if more than one shipment are loaded together in a lorry or container. This is quite common in Norway because of a scattered industry and settlements. If goods are reloaded during a trip, origin and destination will differ between transport and delivery pattern. It is not possible to identify multi-mode shipments from the mode specific surveys.

Lorry counts

Freight vehicles with maximum payload up to 3,5 tons have been excluded from the Lorry survey since 2000. A simplified survey among smaller freight vehicles were carried out by TØI in 2004 (Rideng and Strand), but the survey does not include information about delivery structure/OD-pattern.

If regional models should be developed, then service transports have to be included since their share of the total traffic work by freight vehicles play a very important role in urban areas.

Coastal count

Coastal transports still play an important role for distribution to smaller towns and coastal areas in Norway. Oil and other bulk are of course important commodities shipped by vessels, but also general cargo are often distributed by vessels to smaller municipalities in the north and west of Norway. A regularly survey among coastal shipping in Norway were carried out every fifth year until 1993. The 1993 survey did not cover inland liner shipping. Since 1993 this survey have been the only source for seaborne transport pattern in Norway. It is an important weakness in the transport statistics that the inland seaborne transport pattern is missing.

In Norway we have two different ship registers; one for ships in international lines and one for domestic vessels that mostly visit Norwegian ports (and only occasionally visit other North Sea ports located outside Norway[20]).

Important information to collect from coastal surveys for our purpose will be:

- type of ship with a measure of loading capacity, age and type of trade
- tons of cargo carried for every trip and each leg of the trip
- type of commodity carried
- routing of the ship and typically the hub for loading
- ship costs (or data as number of crew, bunkers, port costs such as infrastructure costs and handling costs of cargo, length of stay in a port, number of days in operation over the year)

Most of the Norwegian coastal vessels are private owned and old (often more than 30 years). Because Norway traditionally has had an international merchant fleet there are a lot of trained persons to run small coastal vessels. A fear is that when such persons retire, the coastal vessel fleet capability will vanish. Probably such transports will then be substituted by road transport. Better infrastructure of roads and high labour costs make this more profitable and flexible for the user, but it is not a good solution for the environment.

Rail transports

For rail transport there are hardly public available statistics showing yearly freight amounts transported by rail, and not at all statistics showing spatial transport pattern. During the development of the national freight model and experiences from the terminal survey shows that information about rail transport, given by the National freight operator on rail in Norway (CargoNet), seems to be strictly confidential.

6.4.2 Roadside surveys

For example: A road survey carried out in Vestfold in 2002 by SINTEF and another survey carried out in the Northern part of Norway in 2003 by the National Road Administration.

In roadside surveys freight transport vehicles in traffic are stopped along a road, and an interviewer are asking the drivers about load, commodities, origin and destination of the transport, etc.

The most important information in roadside surveys is information on the choice of logistics of the commodities. Where do the lorries depart and where is the hub of transport for non-direct transports? It is also interesting to compare the results of the STAN model (road routes) to empirical data. The Swedish model work experienced a lot of discrepancies between the two.

Another important source of information is the purpose of travel in cities where transports of goods are done in combination with the service industry. These types of transports will not be captured in freight surveys and due to the large amount of such transports it is important to collect this information.

---

[20] NOR (Norwegian Ordinary shipRegistry) differently from the international NIS (Norwegian International Shipregistry)

Advantages: These surveys are often worked out in connection with more local studies and can give quickly indicators of the traffic composition and size for a smaller area.

Disadvantages: It is difficult to make a representative sample in respect to roads, were and when to stop the vehicles, time on day, week, month, etc. Also delivery pattern can be hard to establish out of these surveys.

6.4.3 Commodity flows survey (CFS)

USA was probably the first country to work out a commodity flow survey with a spatial dimension. The first survey was carried out as early as 1963, and therefore a lot of experiences are gained. Sweden has fulfilled tree surveys; two pilot surveys in 1996 and 1998 and one full-scale survey in 2001. The Swedish surveys seem to be more or less a blue print of the survey in USA. What is probably less known is that in 1982 there were carried out a CFS for Norway, but without the spatial dimension.

Experiences from Norway

The Production Accounts in the Norwegian National Accounts (NA) are based upon a commodity flow matrix in value terms. It is based upon detailed commodity data from imports and production delivered to input, consumption, capital formation or exports. Each production account (each commodity) is calculated in basic values. (The basic value is a value calculated from the production side where the VAT and other indirect taxes are not included from both sides.) The trade sector is however included in the production accounts as a separate addition for each commodity. All the statistical information is collected yearly by Statistics Norway and is calibrated for the purpose to be fitted in the NA matrix.

There are however from a transportation point of view three weaknesses with “the Norwegian NA CFS matrix”:

1. There are no geographical location specified (except for the regional NA)
2. The trade sector is not explicitly specified
3. Due to lack of resources input are not so frequently covered

The NA by county gives a location on the production side, but give no flows between counties. The data in the NA by county are heavily aggregated compared to the NA data for the production accounts.

Due to very rigorous NA calculations and a rapid increasing service industry, the need for a NA revision became obvious. In order to make such a revision successful several reports were produced by Statistics Norway. One of these gives a CFS for 12 specific commodities (goods) for the national economy at producer costs. (Strøm K F, 1983).

Experiences from USA

In USA a commodity flow survey has been carried out approximately every fourth/fifth year since 1963 (1963, -67, -72, -77, -83, -88, -93, -97 and finally 2002). Until 1972 the survey covered establishments in selected manufacturing industries in the 48 contiguous States, but since 1977 the survey have covered all manufacturing establishments in the US. From 1983 the survey had an expanded coverage to include selected mining industries and grain wholesalers. In 1993 the coverage of the survey was further expanded, by including mining, manufacturing, wholesale, selected retail and service industries and the sample size consists of nearly 200.000 establishments. In 1997 the main work was related to make a more efficient sample design: The sample size and respondents burden were reduced. It still surveyed approximately 100 000 freight shippers within the United States – some five million shipments both within and across the US border in 1997. The 1997 survey required shippers to report the origin, destination, commodity type, tonnage, value and mode sequence used to move a sample of their shipments drawn at four different times during the cource of the calendar year. The survey does not ascertain the routing and therefore the mileages involved in getting freight from one place to another. Southworth and Peterson (2000) has developed a intermodal network model that assesses the routing, mileage and missing sequency of reported modes and other transport indicators missing in the CFS.

The surveyed establishments was reduced to 50 000 for the 2002 survey.

The US CFSs are based on tree stage sampling stratified after 1) Establishments, 2) Weeks of the year and 3) Outbound shipments. The CFS Industry coverage excludes:

- Crude petroleum and natural gas
- Farms
- Service industries
- Governments Establishments
- Imports (until shipment reaches 1st domestic shipper)
- Trans boarder shipments

Sample of weeks of the year are drawn to cover all 52 weeks of the year:

- 4 one-week reporting periods assigned to each establishment
- 1 reporting period in each calendar quarter

Sample of outbound shipments:

- for each reporting period respondents report for a sample of outbound shipments
- sample of shipments is selected by respondents using provided instructions
- average number of shipments reported = 25
- maximum number reported is set at 40

For each sampled shipment, the respondent report:

- origin, destination
- mode of transportation
- commodity (5 digit SCTG)
- value of shipment
- weight of shipment

For export information is given about:

- export mode
- foreign destination

Hazardious material information

- four digit United Nation/North American Code (UN/NA)

CFS Modes of transportation

- Parcel, delivery, courier or Postal services
- Private truck
- For-Hire truck
- Railroad
- Shallow draft vessel
- Deep draft vessel
- Great lakes
- Pipeline
- Air
- Other

A new survey will be conducted in 2007.

Applications of the US 1993 CFS shows that local transportation of freight is important for the economy of Wisconsin since in 1993 roughly 35% of the value and 70% of the weight of the total shipment from the state were shipped to destinations within the state

(Sorrantini, 2000). The surveyed showed that 84% of the value and 88% of the weight of commodities originating in Wisconsin were moved specifically by truck.

To establish a trip tables for Wisconsin, Huang and Smith (1999) used the US 1993 CFS to establish production and attraction of freight transport from and to zones in Wisconsin and applied gravity models to form the transport pattern that were subsequently evaluated with regard to distance distribution obtained from the CFS. The zonal production and attractions were then adjusted such that network assignment of matrices obtained by the gravity model gave correspondence between selected link volumes in the network model and actual corresponding volumes obtained by counts. Somewhat surprisingly Huang and Smith (1999) doesn't use the information about the OD pattern that is inherent in the CFS. In the Norwegian context, where register data can be used to generate production and attraction data, it would be attractive to include the trip patterns from a CFS directly as constraints into the gravity model. This could be accompanied by adjustment of either the production and attraction figures or parameters in the gravity model to ensure coherence between network model link flows and actual counts.

Chin, Hwang and Greene (2001) looked at both the 1993 and 1997 US CFS's to determine freight tonnkm of commodities shipped by truck within, to, from, and through each state and thereby provide a measure of the extent to which state economies are linked. They found that 73% and 76% of the truck tonnkm in 1993 and 1997 were between states. Through state shipments account for more than 50% of the truck ton-miles in 19 states. They conclude that the CFSs are critical important as the principle source of freight data for the United States to understanding the role of the nation's highways in supporting freight transport. But they also say that it is still necessary to supplement the CFS data with transborder trade flows.

Experiences from Sweden

A CFS has been carried out tree times in Sweden, first as a pilot survey in 1996 for the northern part of Sweden, then a simplified national survey in 1998, and finally a full-scale (but still a pilot) survey for whole Sweden for 2001.

The 2001 survey covers the manufacturing, mining and wholesale sectors, were sample is stratified after type of production and STAN99 commodity group (deduced from economic sector), number of employees and geographic site (3100 enterprises per quarter of a total population of 38000 enterprises). This is more or less quite identical to the sampling method in the US-survey. Each firm registered a sample of consignments during a specified reporting period. The reporting period is one week for big enterprises and two to tree weeks for small enterprises (varying between business sectors). Consignments variables were, among other factors, value (excluding VAT and transport costs), weight (excluding tara), postal code (for deliveries within Sweden), modes used during the transport chain and own versus hired transport.

The survey from 2001 gives data on the movement of goods from consignors in Sweden and shipments received and consigned from and to foreign recipients/consignors regardless of transport mode. The target population is commodity shipments originating from local units or destined from abroad to local units within companies.

- Consignment invoice value, excluding freight-costs and value-added tax
- Consignment weight, excluding packaging
- Cargo type
- Hazardous goods
- Transport modes within Swedish borders in order of occurrence (location or reload terminals are not asked for)
- Transport modes outside Swedish borders in order of occurrence

- Geographic destinations/origins by country code and nearest larger city
- Post code for consignments with Swedish recipients
- Geographic point of entry/departure to/from Sweden for foreign consignments

Data from the 2001 survey will be used for calibration in order to establishing new base (PC) matrices in the SAMGOS model (Inregia et al, 2004). Base data for establish the PC matrices are economic statistics, were the steps are to establish production and attraction (sum out of and sum into each zone) for each commodity group and then to use restricted entropy models to establish delivery pattern. Restrictions in the entropy model will be based on delivery pattern from the CFS.

A new survey started the 28. June and will run for one year. Only small changes have been done since the survey in 2001. The main conclusions of the survey in 2001 were that it worked well as a survey. Some cut-off-borders were done to the retail sector and some more business sectors are added (slaughterhouses, corn dealers and car dealers). Information about receiver and delivery business sectors has been complimented with import to get the information about producer and consumer localisation. The survey is also supplemented by a question about sale value last month to intercept mistakes in the answers.

Advantages: Whole transport chains are covered from production to consumption, even if freight is reloaded between modes through the transport chain.

Disadvantages: Quite expensive to carry out. The next Swedish national survey has an estimated cost frame of 8 million SEK. Difficult for the establishments to specify correctly which business the receiver is part of (e.g., whether the firm is classified as a service- or commodity trade company).

6.4.4 Shipper based surveys

Shipper based surveys have frequently been carried out in the Netherlands and France since the second parts of the eighties. An European survey was carried out in the Mystic project. These surveys seems to be quite similar to the commodity flow surveys, were the main difference are the sampling unit shipper instead of industry in the CFS surveys.

The purpose of the Dutch surveys was ambitious. The idea was to collect data from shippers about transport in such a way that it would be a representative sample of domestic transport (transport within the Netherlands) imports and exports. The work started with a large-scale shipper survey by mail, followed by a set of small-scale surveys: In depth interviews among shippers and surveys for forwarders and carriers. On this sample a new set of models (e.g. mode choice models) could be based. For a number of reasons this objective was never met.

A simplified survey was held by TNO Inro in 1998 to feed the SMILE database. The survey consists of telephone interview with 300 logistics and product managers, asking them about key characteristics of their goods and the organization of their distribution chains. One of the general, but rather important findings was that for almost half of the products studied (in terms of shipments), intermediate inventories were in use. This confirmed the idea that modelling of distribution chains is important to accurately represent freight flows in and around the country.

The French shippers' survey was carried out in France during the first six months of 1988. This survey had two main objectives:

- To obtain better knowledge of transport demand and identify the practices and constraints of the shippers depending on their economical environment.

- To describe the ways of transport organisation implemented to satisfy this demand and to reconstitute transport chains with their whole complexity both physical and organisational.

Advantages: Private truck activity is captured. Detailed information about shipment size is uncharted (e.g. commodity, weight, distance). Information about the whole transport chain is captured. Only a small difference to CFS, were shipper is sample unit in shipper-based surveys, while industry and wholesalers are sample units in CFS.

Disadvantages: Probably more difficult to obtain right estimates for freight volumes in total based on sample by shipper than by industry (as in the CFS). The surveys doesn't capture import movements from the border to initial domestic location and excludes transit transports.

6.5 Recommendations for future data sampling in Norway

A commodity flow survey (CFS) would contribute with information about delivery pattern, consignment size, the whole transport chain from production to consumption, transport costs including loading, reloading and unloading costs and inventory costs. This is valuable data and missing data that can be used in the calibration process of establishing delivering pattern in the base matrices. Data outcomes about transport costs could also be used in the logistic module. It would also be valuable for checking out what kind of commodities that are usually being transported together with other consignments during the long distance in a transport chain. Finally it would be valuable to get information about load factor for different kinds of containers.

Ideally, we need data for tonnes transported within and between pairs of municipalities for every business link. If we had this information for all pairs of municipalities, the data would themselves create the necessary OD pattern. It is self-evident that it is impossible to collect data at that level of detail. Representative data at a higher level of aggregation is more realistic, and can be used as constraints in the calibration process. Such data are most easily obtained through a commodity flow survey. A commodity flow survey is also well suited to obtain data for vehicle type and cost of transport, reload and inventory costs. These data can be used in the model system (e.g., in the logistic module). This is further discussed in chapter 4.

6.5.1 Data collection in different time horizons

First of all we remind about the available information on the register data that is still unused, with respect to level of aggregation both in the geographical dimension, but also with respect to commodity aggregation. What is scarce information about is first of all delivery pattern, but also information needed in the logistic module as consignment size and internal inventories are scarce. Therefore we recommend a commodity flow survey were these sort of information can be chartered. It is however not recommendable to start with a full-scale CFS. It would be better to start with a pilot survey for a restricted area or some selected business sectors. Choice of area should be where the need for a regional model are greatest. This could be an area close to one of the bigger cities in Norway. Oslo is a natural choice, but have a quite difficult transport pattern.

Short time horizon - 1 to 3 years

While we are waiting for a national CFS to be worked out, we recommend to complete the mode specific statistic and also supplement the sample based surveys with actual transport data . More specified this is to:

- Carry out a new coastal count, and
- Start the work of planning a pilot CFS

Statistics Norway would probably start with an electronic data exchange from transport operators data systems connected to the principle of reducing the drivers burden related to the Lorry counts.

Medium time horizon - 2 to 5 years
In the medium time horizon the work of improve base transport data must be further improved. We recommend to:

- Carry out a pilot CFS
- Carry out a pilot survey among service transports
- Provide that the time horizon between two coastal surveys not became more than five years, eventually carry out a study were the relation between AIS and the port statistic are combined to establish delivery pattern for seaborne transport.

Long time horizon – 5 years and longer

- Based on the experiences of the pilot CFS, carry out a full scale CFS, including service transports.

6.5.2 Other applications of outcomes from a CFS

There are at least additional applications of a CFS, than just the purpose of transport modelling. These are:

1) For use in development of National Accounts statistics by Counties (Statistics, Norway)
2) For use in regional economic analyses (Ministry of Local Government and Regional Development)
3) To quantify spillover effects from measures (Ministry of Trade and Industry)

The National Account by counties needs information about delivery patterns of both input and output, but they hardly need the geographical pattern but more the structure of delivery pattern between economic activities.

In regional economic analyses one missing link is information about trade between regions. A CFS can give useful information to such analyses. We also think that the Panda model could benefit from a CFS survey.

In analyses of economic activities between regions, an important factor is spill over effects, or more detailed knowledge between activities in a chain of deliveries in the regions.

In addition transport operators with a national or regional covering network or someone who want to establish a transporting activity would have benefits from these kind of survey.

6.5.3 Cost estimates for further data use and data collections

To give cost estimates for data uses (the manufacturing and trade statistics and the external trade statistics are quite difficult. We have tried to get estimates from Statistics Norway but have not gotten any answers. Therefore the estimates are based on earlier bought statistics at a more aggregated level. The table below give a summarising of expected costs.

Table 6.1. Rough cost estimates for more detailed data for establishing a more detailed level of aggregation in the freight network models.

	Source	Zone level	Commodities	Cost estimate
Manufacturing industry	SSB	Postal zones	Detailed	30.000 NOK
Trade industry	SSB	Postal zones	Detailed	30.000 NOK
External trade industry	SSB	Postal zones	Detailed	40.000 NOK
Timber	SkogData	Coordinate level	Timber, chips and pulp	30.000 NOK

TØI-report 731/2004

Working hour related to data preparation and establishing revised base matrices come in addition.

Cost estimates for a commodity flow survey are even more complicated to give estimates for. Such kind of survey seems to be most rationale that is carried out by per example Statistics Norway that have the main competence in collecting statistics. A rough estimate is however around 2 millions for a regional survey covering the Oslo/Akershus area. For a full-scale CFS, the Swedish estimate is around 8 millions Swedish kroner. We would think that the cost for carry out a similar survey for Norway would be approximately at the same level of costs.

6.6 Main conclusion

The main conclusion from the project is that much information can be got from use of more detailed statistic than already used when developing the freight model system for Norway. Register data from SSB are available at postal code level, that is tree times more detailed than the recommended zonal level in this report, equal to the level used in NTM-5 (NaTional passenger Model), where number of zones are 1428.

The main missing data is information about delivery pattern (both with respect to geography but also with respect to witch business link the delivery goes through). Information about delivery pattern between detailed zonal sizes can be got for truck transports for use in regional models from data extraction from transport operators data systems. The main problem is however how to get a covering picture, because the transport marked consists of many small enterprises (some with only one employer) were the small transport operators are not expected to have any electronic system for their transports at all. Another problem is how to catch transports on own account.

With respect to the logistic module information is also missing about consignment size and volumes through internal warehouses (inventory). These information in addition to information about delivery pattern and transport costs can be gotten from a commodity flow survey (CFS). There are experiences from carry out such a survey from both USA and Sweden. The main drawback with a CFS is however the high level of costs related to such a survey, but a survey can be limited to either selected regions or some selected business sectors.

References

Chin S-M, H-L Hwang and D L Greene (2001): In, out, within and through: Geography of truck freight in the lower 48. *Transportation Research Record, 1768:pp.18-25.*

De Jong G, M Ben-Akiva, S Bexelius, A Rahman, M Van de Voort, M Florian, M Baker, P Gibbs, U Hester, L Lingbrant and L Lundqvist (2004): T*he specification of logistics in the Norwegian and Swedish National Freight Model Systems – Note on the scope of the model and preliminary ideas on the model structure.*

Civitas (2004): *Økt kunnskap om varestrømmer i Osloregionen. Forprosjekt*. Civitas-rapport, juni 2004.

Hovi I B and A Vold (2003): *An overview over the national freight model for Norway*. Paper presented at Conference on National and International Freight Transport Models. Helsingør 2003.

Huang W-J and R L Smith Jr (1999): Using commodity flow survey data to develop a truck travel demand model for Wisconsin. *Transportation Research Record, 1685:pp 1-6.*

Ivanova O, A Vold and V Jean-Hansen (2003): *A SCGE Approach to Forecasting Interurban Transport Flows in Norway*. Paper presented at the Conference on National and International Freight Transport Models. Helsingør 2003.

Inregia, TNO and SCB (2003). *DISTRA – pre-study on modelling local/regional distribution and collection traffic*. 2003-03-22.

Inregia, VTI and SINTEF (2004). *National specification of base matrices for the Swedish National freight model system*. Sub deliverable 2 prepared for the concertation meeting in Stockholm 2nd of September 2004.

Ivanova I, A Vold and V Jean-Hansen (2002): *PINGO A model for prediction of regional and interregional freight transport, Version 1*. TØI report 578/2002. ISBN 82-480-0266-7. Oslo: Institute of Transport Economics.

Jean-Hansen V and J Andersen . *City logistics*. Paper presented at the Aalborg University, Trafikkdage på Aalborg Universitet. August 2004.

Johansen L (1960): *A Multi-Sectoral Study of Economic Growth. Contribution to economic analysis*. North-Holland Publishing Company. Amsterdam, New York, Oxford.

Lervåg, L-E, S Meland og R Wahl 2001: *Utvikling av NEMO/REGO. Parameterverdier*. SINTEF Bygg og Miljø, Trondheim.

NTP Transportanalyser (2004). *Handlingsprogram for utvikling av tverretatlige godsmodeller i perioden 2004 – 2008.*

Rideng, A and S Strand (2004): *Transport performance for vans and small lorries.* TØI report 729/2004. Oslo: Institute of Transport Economics.

Silborn, H, A Kroneberg, P Pedersen and O Eidhammer (2004): *Freight terminal structure in Norway. General cargo terminals. Proposal*. Working paper of March 5th 2004 TØ/1614/2004. Oslo: Institute of Transport Economics.

Sorrantini J A (2000): Estimating statewide truck trips using commodity flows and input-output coefficients. *Journal of Transportation and Statistics*, April 2000.

Southworth F and Peterson B E (2000): Intermodal and international freight network modelling. *Transportation Research Part C, 8:pp.147-166.*

Strøm K F (1983): *Varestrømmer i engros- og detaljhandel*. SSB Rapporter 83/31. 1983. Oslo. Statistics Norway.

Statistics Norway (SN):

Agriculture statistics: See "Statistikkbanken" at SN's internet page. Data per municipality must be purchased from SN.

External Trade Statistics - NOS. (ETS) Published continuously on the internet (See SN's internet page "Statistikkbanken"). In order to have detailed estimates on geographical locations (e.g. Custom Offices) data has to be bought from the data producer (SN).

Fishery statistics: Published by both SN and the Directory for Fisheries (DF). From DF one may obtain more geographical detailed estimates. In fact there are defined zones in the North Sea (more than 100 zones) where the fish are catched and the movements of fishing vessels to and from the field of each type of fishery (cod fish, pelagic fish and other).

Forestry statistics: See "Statistikkbanken" at SN's internet page.

Manufacturing Statistics (MS). The same description can be given for MS as for the TS above. It is possible to get ton estimates for single enterprise-establishments by location from MS. These data have to be purchased by the SN.

National Accounts statistics (NA): See SN's internet page on this topic. Yearly estimates are given for the production account – gross output and value added for NA publishing sectors. More detailed estimates (diaggregated both on the input and on the output side of the sector. E.g. imports to and exports from each sector (300 sectors are specified) and a lot more information) can be obtained from the SN.

National Accounts statistics on Counties (NAC): Published by SN on its page on the internet. Other specifications can be obtained from the SN e.g. private consumption specified by where the consumption actually has been purchased and where it is consumed. Wage income for the counties where the wages have been earned and where it has been taxed (by location where the wage earners live – important distinction in PINGO). NAC are produced for 1973, -76, -80, -83, -86. –89, -92, 94, -97, -99, -2000, and -01. But the input side and consumption is not published since 1994. For the years 1973 to 1986 full input output tables are published (the input side for every NAC sector is broken down on commodities).

Retailer and Wholesalers' statistics – NOS (RWS): As for the ETS. More detailed data than given in the "Statistikkbanken" has to be bought from the data producer (SN).

Vold A et al (2002): *NEMO. Nettverksmodell for godstransport innen Norge og mellom Norge og utlandet. Versjon 2*. TØI-rapport 581/2002. Oslo: Transportøkonomisk institutt.

Vold A and I B Hovi (2004): *Construction of basis OD matrices for 1999 freight flows between municipalities in Norway*. TØI report 699/2004. Oslo: Institute of Transport Economics.

Appendices

Appendix 1

Table A.1. Employment sectors in PANDA.

1 Agriculture	29 Manufacture of electrical equipment and supplies
2 Forestry and Logging	30 Manufacture of instruments and technical equipment
3 Fishing	31 Building of ships, boats and oil platforms
4 Breeding of Fish	32 Manufacture of other transport equipment
5 Crude petroleum and natural gas productions, pipeline transport	33 Manufacture of furniture and fixtures
6 Coal mining	34 Manufacture of industrial products
7 Metal ore mining	35 Electricity, gas and water supply
8 Other mining	36 Construction
9 Manufacture of fishery products	37 Repair of motor vehicles and household equipment
10 Manufacture of food and foodstuffs	38 Wholesale and retail trade
11 Manufacture of beverages and tobacco	39 Hotels and restaurants
12 Manufacture of textiles	40 Ocean transport
13 Manufacture of clothing	41 Coastal and inland water transport
14 Manufacture of leather, leather products and footwear	42 Land- and air transport
15 Manufacture of wood and wood products	43 Post- and telecommunications
16 Manufacture of paper and paper products	44 Financial services
17 Printing and publishing	45 Real estate
18 Manufacture of industrial chemicals	46 Business services
19 Manufacture of other chemical products	47 Insurance and development services
20 Manufacture of products of petroleum and coal	48 Private services
21 Manufacture of rubber and plastic products	49 Production of local government services
22 Manufacture of ceramic products	50 Production of general government services
23 Manufacture of glass and glass products	60 Educational services not included in governmental services
24 Manufacture of other mineral products	61 Public administration not included in governmental services
25 Manufacture of iron, steel and ferroalloys	62 Health services not included in governmental services
26 Manufacture of non-ferrous metals	63 Cultural services not included in governmental services
27 Manufacture of metal products except machinery and Equipment	64 Sports, other leisure services not included in governmental services
28 Manufacture of machinery and equipment	99 Production of other services not included elsewhere

TØI-report 731/2004

Appendix 2

Register statistics in Norway

Production and consumption statistics available in Norway

There seem to be some misunderstandings of what type of economic and volume data available in Norway in chapter 4.2.2. We have therefore made a table that gives an overview of available economic and volume statistics.

Table A2. Available data for production and trade in Norway (in value and volume).

Sector (coinciding by typical product)	Value	Volume	Geographical M= Municipal C= County N=National	Yearly or less frequent
Agriculture	Yes	Yes	M	Year
Fisheries	Yes	Yes	M	Year
Forestry	Yes	Yes	M	Year
Mining (establishments)	Yes	No	M	Year
Oil and gas	Yes	Yes	M	Monthly
Manufacturing at different levels:	**Data available at different levels – see below**			
Enterprise	Yes	Yes	Location of head office	Year
Establishments	Yes	No	M	Year
Construction	Yes	No	(M)	Year
Trade	**See for each trade sector (65 sectors (coinciding with commodity) specified in Norway):**			
Retail	Yes	No	M	Year
Wholesale	Yes	No	M	Year
Foreign Trade	**For both sides for every country and by transport mode**			
Imports	Yes	Yes	Location of custom office	Monthly
Exports	Yes	Yes	M	Monthly