

The Aggregate-Disaggregate-Aggregate (ADA) freight model system

Moshe Ben-Akiva, Massachusetts Institute of Technology
Gerard de Jong, Significance, ITS Leeds and CTS Stockholm

1. Introduction

Disaggregate models –defined here as models using observations at the level of the traveler, the traveling group, the business establishment or the shipment- have several advantages over aggregate models (which use groupings of those units as observations, e.g., groupings by geographic zone). Disaggregate models can be based on a foundation in behavioral theory, can include more detailed policy-relevant variables and do not suffer from the aggregation biases of aggregate models. Nevertheless, there are perfectly valid reasons why some of the components of a model system are modeled in an aggregate fashion. In this paper, we propose an **'aggregate-disaggregate-aggregate'** (ADA) model system for freight transport.

The context is that of a model system at the international, national or regional scale, designed by or for public authorities. Such international, national and regional freight transport models are used for different purposes, including:

- Forecasting transport demand (and through this, emissions from traffic, traffic safety, etc.) in the medium to long run under various scenarios;
- Testing transport policy measures, such as road user charging;
- Predicting the impacts on traffic (and traffic-related measures as mentioned above) of the provision of new infrastructure (roads, railway lines, canals, bridges, tunnels, ports, public freight terminals).

In the ADA model system, the production to consumption (PC) flows and the network model are specified at an aggregate level for reasons of data availability. Between these two aggregate components is a logistics model that explains the choice of shipment size and transport chain, including mode choice for each leg of the transport chain. This logistics model is a disaggregate model at the level of the firm, the decision making unit in freight transport.

Most (inter)national or regional freight transport model systems are lacking logistics elements, such as the determination of shipment size or the use of distribution centres. Exceptions are the SMILE and SMILE+ model in The Netherlands (Tavasszy et al., 1998; Bovenkerk, 2005), the SLAM model for Europe (SCENES Consortium, 2000, TNO, 2008), the EUNET model for the Pennine Region in the UK (Jin et al., 2005) and other regions in the UK (Bates et al., 2011), the model for Oregon (Hunt, 2003, Hunt et al., 2001, PbConsult, 2002), the work of Liedtke (2005), which includes an application to German long-distance markets and the FAME model for freight transport in the US (Samimi et al., 2010). Reviews of these developments are given in Tavasszy et al. (2012) and de Jong et al. (2012).

Section 2 of the paper explains the structure of the ADA model system. The various components of this model system are treated in section 3, focussing on the disaggregate ("middle") part of the ADA system (the logistics model) and on the disaggregation that comes immediately before the disaggregate part and the

aggregation that comes directly after it. The logistics model's data requirements are discussed in section 4. Estimation/calibration and validation issues for the logistics model is discussed, as well (section 5). These first five sections all present the model system in general terms, not in terms of an application to a specific study area. In section 6, various applications of the ADA model in the context of the freight transport model systems for Norway, Sweden Flanders and Denmark are presented. Finally, a summary and conclusions are provided in section 7.

2. The ADA model structure

2.1 The general concept

Figure 1 is a schematic representation of the structure of the freight model system. The boxes indicate model components. The top level of Figure 1 displays the aggregate models. Disaggregate models are at the bottom level. So in the ADA model system, we first have an aggregate model for the determination of PC flows, then a disaggregate "logistics" model, and finally another aggregate model for network assignment. In this section 2.1, we explain the general concept of the ADA model system. The relation between the first aggregate part (zone-to-zone PC flows) and the disaggregate part is further treated in section 2.2, and in section 2.3 the boundary lines between the disaggregate logistics model and the last aggregate part (assignment) are discussed.

The model system starts with the determination of flows of goods between production (P) zones and consumption (C) zones (being retail for final consumption; and further processing of goods for intermediate consumption). These models are commonly based on economic statistics (production and consumption statistics, input-output tables, trade statistics) that are only available at the aggregate level (with zones and zones pairs as the observational units). Indeed, to our knowledge, no models have been developed to date that explain the generation and distribution of PC flows at a truly disaggregate level.

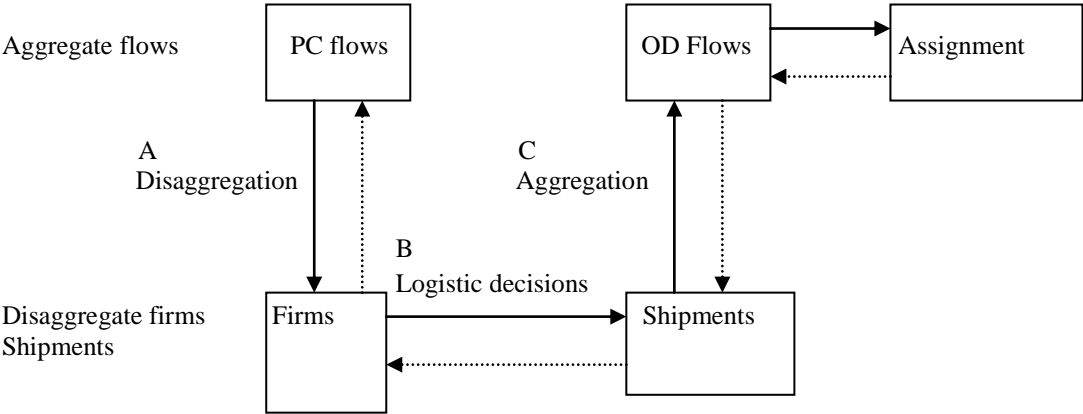


Figure 1. ADA structure of the (inter)national/regional freight transport model system

Most existing freight transport model systems include submodels for generating PC or origin-destination (OD) matrices (possibly by mode) and routines for assigning either one of these matrices to the networks (unimodal or multimodal). As explained

in section 2.2, assignment of PC flow to the networks would not be correct. A relatively new phenomenon of the ADA model is the inclusion of a logistics model that on the basis of PC flows produces OD flows for network assignment. The logistics model consists of three steps:

- A. Disaggregation to allocate the flows to individual firms at the P and C end;
- B. Models for the logistics decisions by the firms (e.g., shipment size, use of consolidation and distribution centres, modes, loading units, such as containers);
- C. Aggregation of the information per shipment to OD flows for network assignment.

This model structure allows for logistics choices to be modelled at the level of the actual decision-maker, along with the inclusion of decision-maker attributes.

The allocation of flows in tons between zones (step A) to individual firms are, to some degree, based on observed proportions of firms in local production and consumption data, and from a registry of business establishments. The logistics decisions in step B (to be modelled as a random cost choice model) are derived from minimization of the full logistics costs, including the transport costs.

The aggregation of OD flows between firms to OD flows between zones provides the input to a network assignment model, where the zone-to-zone OD flows are allocated to the networks for the various modes. Assignment can, in principle, be done at the level of individual vehicles (microscopic or mesoscopic models for simulating route choice, see Ben-Akiva et al., 2007; the Oregon model also has assignment at the level of individual vehicles). In such cases, the ADA model would be an ADD (aggregate-disaggregate-disaggregate) model (see Fig. 2).

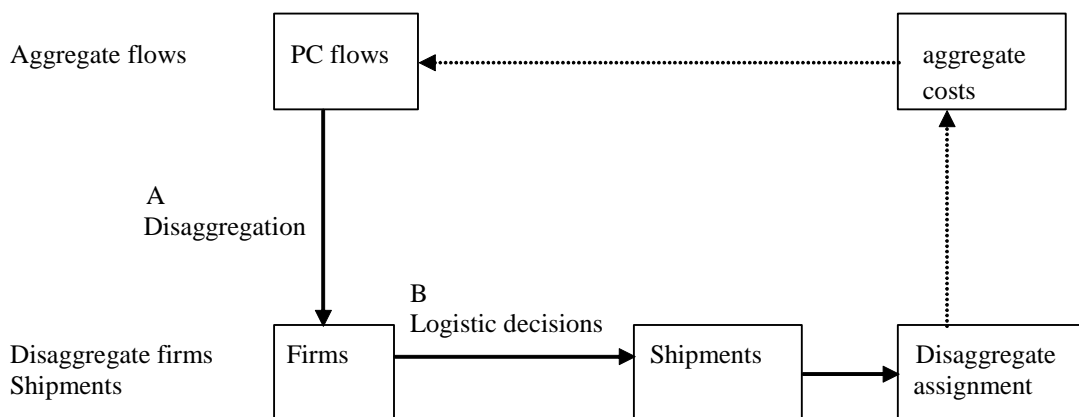


Figure 2. Proposed ADD structure of the (inter)national/regional freight transport model system

Most model systems perform an assignment of aggregate zone-to-zone flows (possibly with several user classes) to the networks in order to use available software, to keep the model tractable and to keep the run time manageable. This approach is also used because the network level is the level at which validation/calibration data are usually available (e.g., traffic counts at various

locations/screenlines). On the other hand, vehicle-level data that can be used for the estimation of micro-level network models are becoming more common.

There can also be backward linkages, as seen in Figure 1 (the dashed lines). The results of network assignment can be used to determine the transport costs that will be part of the logistics costs which are minimized in the disaggregate logistics model. The logistics costs for the various OD legs can be summed over the legs in the PC flow (and aggregated to the zone-to-zone level by an averaging over the flows). These aggregate costs can then be used in the model that predicts the PC flows (for instance, as part of the elastic trade coefficients in an input-output model). In case of assignment of individual vehicles in an ADD model, one might calculate aggregate transport costs at the PC level by adding the costs of the different vehicles that are involved in the same PC chain, and then averaging over PC flows (Fig. 2., dashed lines).

2.2 Relation between the PC flows and the logistics model

The PC flows between the production locations P and the consumption locations C are given in tons by commodity type. The consumption locations here refer to both producers processing raw materials and semi-finished goods and to retailers. The logistics model then serves to determine which flows are covered by direct transports and which transports will use ports, airports, consolidation centres (CCs), distribution centers (DCs) and/or railway terminals. It also gives the modes and vehicle types used in the transport chains. The logistics model, therefore, takes PC flows and produces OD flows. An advantage of separating out the PC and the OD flows is that the PC flows represent what matters in terms of economic relations - the transactions within and between different sectors of the economy. Changes in final demand, international and interregional trade patterns, and in the structure of the economy, have a direct impact on the PC patterns. Also, the data on economic linkages and transactions are in terms of PC flows, not in terms of flows between producers and trans-shipment points, or between trans-shipment points and consumers.

Changes in logistics processes (e.g., the number and location of depots) and in logistics costs have a direct impact on how PC flows are allocated to logistics chains, but only indirectly (through the feedback effect of logistics choices and network assignment) impact the economic (trade) patterns. Assigning PC patterns to the networks would not be correct. For instance, a transport chain road-sea-road would lead to road OD legs ending and starting at ports instead of a long-haul road transport that would not involve any ports. A similar argument holds for a purely road-based chain that uses a van first to a consolidation center, then is consolidated with other flows into a large truck, and finally uses a van again from a distribution center to the C destination. In this scenario, the three OD legs might be assigned to links differently than would be the case for a single PC flow. Therefore, adding a logistics module that converts the PC flows into OD flows allows for the trade-off between inventory, order and transport costs (endogenous shipment decisions) and for a more accurate assignment. The data available for transport flows (from traffic counts, roadside interviews and interviews with carriers) also are at the OD level or screenline level, not at the PC level.

2.3 Relation between the logistics model and the network assignment

In several existing freight transport model systems, a network model carries out both the modal split and the assignment to the networks (in a multimodal assignment). If the assignment in the ADA model system would be multimodal, the logistics model would not have to cover mode choice. This is labelled Option II in Figure 3. A better approach would be to include mode choice in the logistics model, and restrict the assignment to be unimodal. In the latter approach, the mode choice would be determined together with the other logistics choices (e.g., shipment size, number of legs in a transport chain, terminals used). This is Option I in Figure 3. The outputs of the logistics model will then be in terms of vehicle or vessel flows (not just tons) between OD pairs.

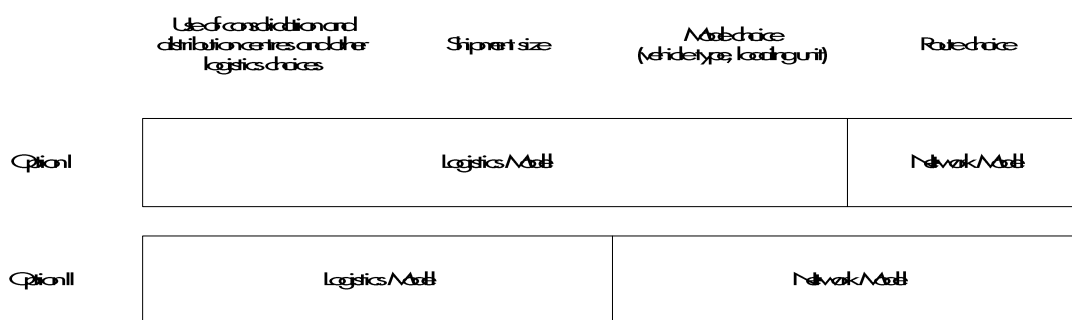


Figure 3. Two options for combining the logistics and the network model

3. Model specification

In this section we describe each component of the ADA model: generation of PC flows; logistics model in three steps; and assignment. Additional detail on the logistics model can be found in de Jong and Ben-Akiva (2007).

3.1 Generation of PC flows at the aggregate level

The type of model used here can be a multi-regional input/output (MRIO) model or a regionalized national input/output model (Cascetta, 2001; Marzano and Papola, 2004; Hunt and Abraham, 2005), or a spatial computable general equilibrium (SCGE) model (Bröcker, 1998; Tavasszy et al., 2002; Ivanova et al., 2007; Vold and Jean-Hansen, 2007; Hansen, 2010). Input data required for these models are input-output statistics (preferably multi-regional), production and consumption statistics by economic sector and international trade statistics. If the resulting PC flows would be in monetary values, conversion to tonnes would have to be done before going to the logistics model. Conversion factors can be based on data sets that include value and weight information for the same shipments (such as Commodity Flow Surveys) or trade/transport flows (e.g. customs data).

3.2 The logistics model

In this section 3.2 we discuss the three steps (A, B and C; see Figure 1) of the logistics model.

3.2.1 Disaggregation to firm-to-firm flows (step A)

Step A in the logistics model (conversion of zone-to-zone flows to firm-to-firm flows) is not a choice model, but rather a prerequisite so that logistics choices can be captured at the actor level. Instead of modelling trade between zones, this step makes modelling trade between firms possible. (These firms are manufacturers, wholesalers or retailers.)

The aggregate representation of the PC flows that are produced by the first aggregate model of the ADA system, and that are input to the disaggregation step A, is as follows:

Flows of goods in tons per year, by:

- r , zone of the sender (production zone)
- s , zone of the receiver (consumption zone)
- k , commodity type.

Step A disaggregates this to a disaggregate representation, characterised by:

Flows of goods in tons per year, by

- m , sending firm (located in zone r)
- n , receiving firm (located in zone s)
- k , commodity type.

Three general approaches to generate a disaggregate population or sample of firm-to-firm flows can be distinguished:

1. Re-weighting - use an existing sample or population and re-weight using marginal distributions (i.e., the row and column totals);
2. Synthetic - draw from a sequence of conditional distributions;
3. Hybrid - begin with re-weighting and enrich the set of characteristics using synthetic draws.

The re-weighting approach is the simplest, but a sample of actual firm-to-firm flows is only rarely available. The US and Sweden have a Commodity Flow Survey (CFS) sample, but these are samples of shipments for a limited time period (e.g, one to three weeks in Sweden). In this case, we are looking for all supplier/receiver pairs on an annual basis. Therefore, in practically all applications, a synthetic or a hybrid approach for step A is developed. In Annex 1 we give a practical example (from the implementation in Norway) of how the disaggregation step can be worked out

3.2.2 The logistics decisions at the disaggregate level (step B)

Step A produces disaggregate supplier-receiver relations (a business relation between two firms in which one is the sender of a good and the other the receiver). Each relation has an annual flow of goods in tons by commodity type. Even for a small area, there are millions of such relations. To reduce runtime, firm-to-firm relations are sampled, and expansion factors are used to obtain population estimates. For each relation, step B simulates the logistics decisions (micro-simulation), and adds the outcomes of these to the level of a firm-to-firm relation.

The different logistics decisions included in step B are:

- Frequency/shipment size (so inventory decisions are endogenous);
- Choice of loading unit (e.g., containerized or not);
- Use of distribution centres, freight terminals, ports and airports and the related consolidation and distribution of shipments. The locations of these trans-shipment points are taken as given, what is determined here is their use. This also gives the number of legs in the transport chain;
- Mode/vehicle type used for each leg of the transport chain. The choice set may contain: air transport, road transport, rail transport and maritime transport (possibly each with different vehicle/vessel types). Economies of scale in transport (larger vehicles have lower unit costs) are taken into account in the cost functions for the vehicle types

What step B of the logistics model does is to add dimensions to the disaggregate representation that was produced by step A. The full disaggregate representation after step B, consists of:

Shipments of goods in number of shipments, tons, ton-kilometers, vehicle-kilometers and vehicle/vessels per year, by

- m , sending firm (located in zone r)
- n , receiving firm (located in zone s)
- k , commodity type
- q , shipment size
- l , transport chain type (number of legs, mode and vehicle/vessel type used for each leg, terminals used, loading unit used).

The basic mechanism in the model for decision-making on all these choices is the minimization of total logistics costs. The total annual logistics costs G of commodity k transported between firm m in production zone r and firm n in consumption zone s of shipment size q with transport chain l (including number of legs, modes, vehicle types, loading units, trans-shipment locations) are:

$$G_{rskmnql} = O_{kq} + T_{rskql} + D_k + Y_{rskl} + I_{kq} + K_{kq} + Z_{rskq} \quad (1)$$

Where:

G: total annual logistics costs

O: order costs

T: transport, consolidation and distribution costs

D: cost of deterioration and damage during transit

Y: capital costs of goods during transit

I: inventory costs (storage costs)

K: capital costs of inventory

Z: stockout costs

In this minimization, it is assumed that the subscripts for the specific firms m and n (and also, for instance, firm size) do not matter. This assumption may be relaxed to accommodate economies of scale in for instance warehousing and ordering. Also, variation in the discount rate for the inventory capital costs and of other preferences between firms may be included.

The purchase costs of the goods from different suppliers are not part of the optimization, since the senders and receivers of the goods have already been determined in step A. However, the purchase costs do play a role through the capital costs of the goods that are included in the equation above.

The decision-making thus takes place at the level of the individual sender to receiver relation. Whether these decisions are taken by the sender or by the receiver will vary from sector to sector and even within sectors. One way to look at it is to regard the sender-receiver combination (and also the carrier if the transport is contracted out) as a single decision-making unit, carrying out the minimisation of the total logistics costs of this firm-to-firm flow. The idea of joint overall optimisation by shippers and carriers was supported by experimental economics (Holguín-Veras et al., 2011).

Equation (1) is expanded as follows: (see RAND Europe et al, 2004; RAND Europe and SITMA, 2005):

$$G_{rskmnq} = o_k \cdot (Q_k/q_k) + T_{rskq} + i \cdot j \cdot g \cdot v_k \cdot Q_k + (i \cdot t_{rsl} \cdot v_k \cdot Q_k)/365 + (w_k + (i \cdot v_k)) \cdot (q_k/2) + a \cdot ((LT \cdot \sigma_{Q_k}^2) + (Q_k^2 \cdot \sigma_{LT}^2))^{1/2} \quad (2)$$

Where:

o : the constant unit cost per order

Q : the annual demand (tons per year)

q : the average shipment size

i : the discount rate (per year)

j : the fraction of the shipment that is lost or damaged (might vary between modes)

g : the average period to collect a claim (in years)

v : the value of the goods that are transported (per ton)

t : the average transport time (in days)

w : the storage costs per unit per year

a : a constant to set the safety stock in such a way that there is a fixed probability of not running out of stock. For medium/high frequency products, a common assumption is that the demand (and lead-times) follows a Normal distribution. a will then be: $a = F^{-1}(CSL)$, where F^{-1} is the inverse Standard Normal Distribution and CSL is the cycle service level, which is the probability that the stock will not be empty during a replenishment cycle.

LT : expected lead-time for a replenishment (time between placing the order and replenishment)

σ_{LT} : standard deviation for the lead-time

σ_{Q_k} : the standard deviation for the yearly demand

The first term on the right-hand side (RHS) of eq. (2) gives the order cost, the second the transport cost, the third the cost of damage to the goods during transport, the fourth the capital costs on the goods in transit, the fifth the capital and storage costs of the (average) inventory and the last term represents the safety stock cost.

The optimal shipments sizes in the standard cases are not influenced by the safety stock, or vice versa. However, different transport alternatives with different transit times have an impact on the safety stock through the lead-time (and possibly through the standard deviation of the lead-time), and, thereby, also impact the inventory cost

(and the total cost). This may be the case for alternative modes. In principle, lead-time should be a function of the mode (h): $LT = LT(h)$.

3.2.3 Aggregation to zone-to zone flows (step C)

In step C of the logistics model, transport chain legs of individual firm-to-firm shipments for the same commodity type are aggregated by origin zone (on the basis of all sending firms and trans-shipment locations in the zone for the commodity) and destination zone (on the basis of all receiving firms and trans-shipment locations in the zone for the commodity). In this way we obtain OD flows in vehicles and tons.

The aggregate representation that is produced in step C is as follows:

Flows of vehicle/vessel units per year, by:

- r , origin zone (production zone or zone used for trans-shipment)
- s , destination zone (zone used for trans-shipment or consumption zone)
- k , commodity type (and distinguishing empty vehicles)
- v , mode (e.g., road transport) or vehicle/vessel type (e.g., heavy truck).

If the disaggregation was done properly, and if step B has not introduced errors, this is simply a matter of straightforward summation over shipments.

In reality not every shipment in a firm-to-firm flow is optimized and transported by itself, but multiple shipments from different firm-to-firm flows are often combined into a single vehicle (which we call 'consolidation'). This reduces the cost burden on the individual firm-to-firm flow, because now costs can be shared with other sender-receiver pairs. In all the ADA models developed so far, we allow for consolidation between trans-shipment locations: these terminals not only serve for modal transfer, but also for consolidating shipments (and deconsolidating these before final delivery). The degree of consolidation depends to a large extent on the presence of other cargo that could be moved between the same terminals, and this can be determined by an initial model run. The next model run then takes the predicted terminal to terminal flows from the first run and uses these to determine the share that an individual shipment has to pay of the transport cost of consolidated shipments, and so on, applying the logistics model in an iterative fashion.

Empty vehicle flows are calculated as follows: the loaded trips are first calculated as described above, and then vehicle balances between zones are used to let vehicles return from where they came, with specific shares for empty and loaded return trips. In this formulation, the probability that some of the empty capacity will also be used for transporting goods in the opposite direction is taken into account.

3.3 Assignment to networks

Standard aggregate network assignment software can be used to carry out an assignment of vehicles or tons (ADA model), or a (less standard) disaggregate assignment (ADD model). The latter is usually a simulation-based procedure, where individual vehicles are loaded one-by-one onto the network. This simulation can be microscopic, and includes all the movements of each vehicle on the network in detail (e.g., including lane changes) or mesoscopic, where more aggregate relationships

(such as speed-density curves) are used to model individual vehicle movements (see Ben Akiva, et al., 2007).

4. Data requirements

For the logistics model presented, the following data are needed for step A (disaggregation from flows to firms):

1. The number of firms (or the local units for firms with multiple establishments) by commodity type and zone.
2. The turnover of these local units and/or the number of employees of these firms.

This information is required both at the production and the consumption end. Another requirement is the consumption pattern (in terms of the commodity classification used) of the firms by commodity type produced. We assume that each firm (local unit) will produce goods in only one commodity class, but it may consume goods from several commodity groups.

Step B (logistic decisions) requires information on the following items:

1. Data on individual shipments: sector of sender and receiver, origin and destination, value of the goods, modes and vehicle/vessel type and size used, cargo unit, shipment size/frequency, use of freight terminals (including intermodal terminals and marshalling yards), consolidation and distribution centres, ports and airports. Preferably this is transport chain information: which shipments go directly from P to C, which use the above intermediate points?
2. Data on where the freight terminals, consolidation and distribution centres, ports and airports are located;
3. Data on logistics costs: transport costs per km, terminal costs, handling and storage costs for all available alternatives.

Most crucial are the data on the shipments of individual firms (item 1 for step B above). The spatial detail needs to be that of the zones used in the model. Step C requires no extra information.

5. Estimation, calibration and validation

We distinguish between model estimation (which takes place on disaggregate data using formal statistical methods), model calibration (which takes place on aggregate data and may or may not involve formal statistical methods) and validation (which takes place after having done the assignment and involves a comparison with traffic count data). Figure 4 represents an estimation, calibration and validation process for the ADA model system for freight transport. The estimation and calibration data are shown above the boxes, while the validation data are below the boxes. The matrices of PC flows are usually partly based on observations and partly synthetic (model-generated). The data used in this process comes from a CFS, regional input-output systems, economic statistics from national accounts and foreign trade data. The logistics model is estimated on a CFS or similar disaggregate data, coupled with information on terminals and time and costs data from the networks.

Estimation: on data up to the base-year

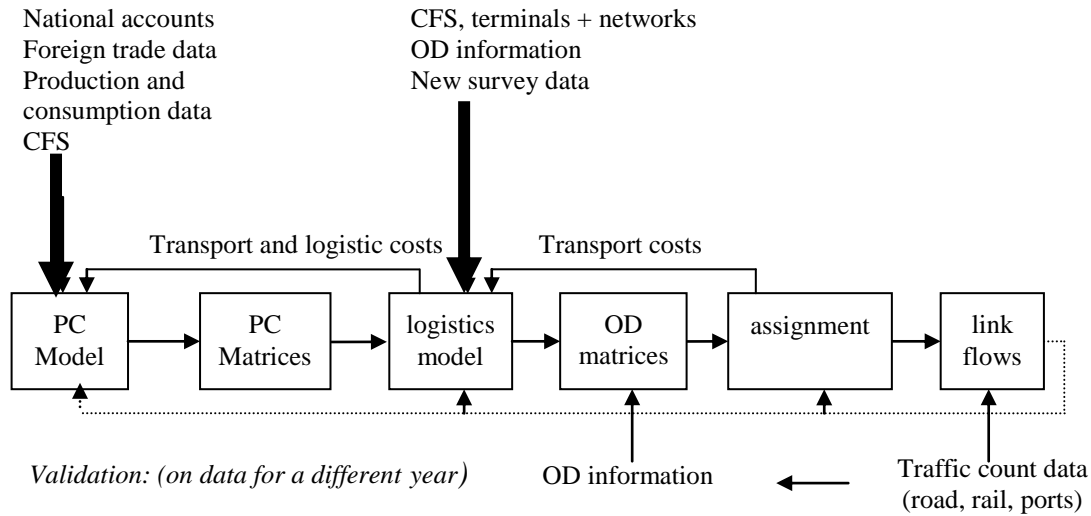


Figure 4. Estimation, calibration and validation of the model systems

The model application process is iterative (see Fig. 4, middle and upper part): after assignment, the new generalized costs are used to adjust the PC matrices, etc. This gives rise to an **inner loop**, which functions as follows:

1. The PC models (e.g., MRIO models) provide initial PC matrices;
2. The logistics model transforms these into OD matrices, using transport cost provided by the network model, and adds empty vehicles;
3. The network model assigns the OD matrices (including empty vehicles) to the networks;
4. The network model and the logistics model provide transport and logistics costs matrices to the PC model;
5. The PC model produces new base matrices on the basis of the new transport and logistics costs and provides these to the logistics model.

This loop continues until equilibrium is reached (in practice, until a pre-set maximum distance from equilibrium is reached). Estimation is not required within this inner loop. The inner loop addresses the adjustment of model variables (inputs and outputs), not model coefficients.

5.1 Estimation with disaggregate data

Data on logistics choices of individual shipments are used in model estimation for step B. The model is based on the total annual logistics costs, such as equation (2). A random cost discrete choice model can be obtained by using total annual logistics costs as the observed component and by adding random cost components ε that follow specific statistical distributions. These random components account for omitted variables, measurement errors and such.

$$C_{mnl} = G_{mnl} + \varepsilon_{mnl} \tag{3}$$

Where (dropping the subscripts r and s for the zones and k for the modes):

C_{mnql} : total logistic and transport cost

G_{mnql} : observed component of total transport and logistics costs

ε_{mnql} : random cost component.

Using equation (2) for G_{mnql} we get:

$$C_{mnql} = \beta_{0ql} + \beta_1 \cdot (Q/q) + X_{mnql} + J_{mnql} + \beta_2 \cdot j \cdot v \cdot Q + \beta_3 \cdot (t_{mnl} \cdot v \cdot Q) / 365 + (\beta_4 + \beta_5 \cdot v) \cdot (q/2) + a \cdot ((LT \cdot \sigma_Q^2) + (Q^2 \cdot \sigma_{LT}^2))^{1/2} + \varepsilon_{mnql} \quad (4)$$

Where:

β_{0ql} - alternative-specific constant

$\beta_1 = 0$

$\beta_2 = d \cdot g$

$\beta_3 = d$ (in transit)

$\beta_4 = w$

$\beta_5 = d$ (warehousing)

In eq. (4) we included a number of items, such as order costs, inventory costs and capital costs of goods during transit, in the coefficients to be estimated because the data on these items can be very difficult to obtain. As a result, the coefficients have specific logistical interpretations. We distinguish between the implied discount rate (d) of the inventory in transit (β_3) and of the inventory in the warehouse (β_5), because these need not be the same.

If we assume the extreme value distribution for ε , the model becomes logit. Nested logit is appropriate when some alternatives (e.g., road and rail transport) have a greater degree of substitution than other alternatives (e.g., road and sea transport). This is an empirical question, and statistical tests, particularly likelihood ratio tests, show whether or not the Nested Logit model is justified. A Logit Mixture Model provides additional flexibility and may be relevant for the logistics model, given the heterogeneity often found in freight transport. Heterogeneity can be captured in two ways:

- Two error components in ε :
 - one following the extreme value distribution;
 - and the other following, for instance, a multivariate normal distribution to allow for flexible correlation structures between alternatives.
- The coefficients in G (the β 's) follow a distribution. This is the random coefficients model or tastes variation model. It may capture heterogeneous preferences in freight transport decision-making.

5.2 Calibration to aggregate data

In the absence of disaggregate estimation data, it is possible to use a deterministic logistics model. This model can be more easily calibrated with aggregate data (OD information).

Through this ADA model, we have developed a procedure to calibrate parameters in the cost function to available aggregate data, which was tested in Norway. For the purpose of calibrating the model to aggregate data, a number of calibration parameters (e.g., for implied discount rates such as β_3 and β_5) in eq. (4), mode-specific constants, a constant for direct transport) were added to the cost function eq. (2). Observed OD data by mode and commodity type for aggregate zones (e.g. 10x10 zones for a country) are used as calibration targets. The calibration parameter values that minimize the difference between the model outputs and the calibration targets were then determined in an iterative process.

5.3 Validation to traffic count data

The validation process is depicted in Figure 4 (bottom part of the figure). After the assignment of the OD flows to the networks, the predicted link flows are compared to observed link flows from traffic counts (especially for road and possibly for rail and port throughput). Any large discrepancies that may arise require analysis. The parameters in all the models are then recalibrated, employing an iterative procedure. This process creates the **outer loop**. In the outer loop, or model calibration loop, model coefficients in all constituent submodels are adjusted to reach a good match with aggregate data.

6. Applications to Norway, Sweden, Flanders and Denmark

The ADA model was first specified in a project for Norway and Sweden (RAND Europe et al., 2004) to replace the existing multimodal network models for freight transport. Both countries have model components for deriving PC flows and for network assignment. Prototype versions (version 0) of the logistics model were developed and tested (RAND Europe and Sitma, 2005, 2006, RAND Europe 2006). There are two different models - one that is part of the Norwegian national freight model system, and one for the Swedish national freight transport forecasting system - but both have the same structure. The Norwegian and Swedish models differ in terms of the zoning system, the modes and vehicle types used, the commodity classification, the construction of the PC matrices, the cost functions and the consolidation rules.

In 2006/2007, version 1 models for Norway and Sweden were constructed. The logistics cost minimization in this improved deterministic model takes place in two steps. In the first step (transport chain generation), the optimal trans-shipment locations are determined for each type of transport chain (e.g. road-sea-road) between each origin and destination zone. The transport chain generation program determines the optimal transfer locations on the basis of the set of all possible multimodal trans-shipment locations (which is exogenous input). These terminals are coded as separate nodes and the program uses unimodal network information on times and distances between all the centroids and all the nodes for all available sub-modes to find the paths, which minimize the total logistics costs.

In the second step (transport chain choice), shipment size and transport chain (number of legs, selection of modes and vehicle types) are determined by enumerating all available options for a specific firm-to-firm flow and selecting the one with the lowest logistics costs.

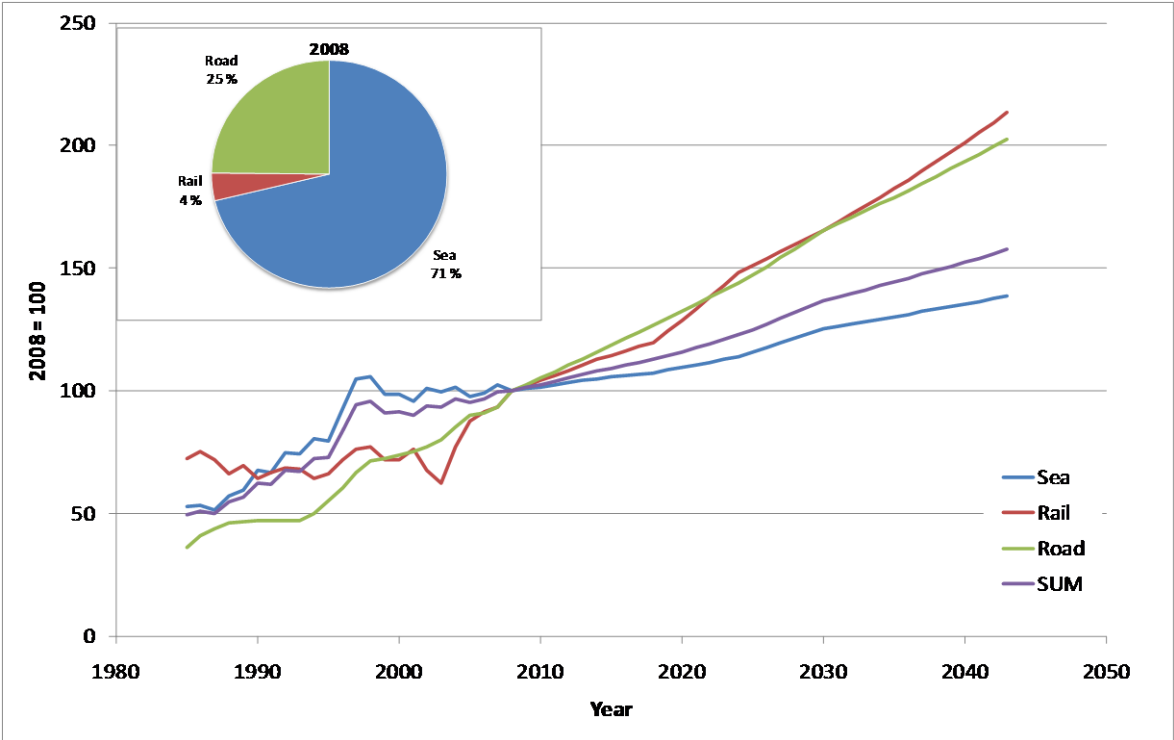
For Norway, this model uses all firm-to-firm flows based on data on firms by number of employees and municipality. No expansion is needed to determine the population of all goods flows in Norway. There is also a SCGE model (PINGO), that can be used to produce forecasts of PC matrices (Hansen, 2010).

In Sweden, the logistics model is combined with procedures for estimating PC matrices, which are based on the National Commodity Flow Survey and mainly use gravity type models (base matrices). A sample of firm-to-firm flows (for different size classes) is used for application of the disaggregate logistics choices, after which, an expansion procedure needs to be used to arrive at population totals.

For both countries, the logistics model produces OD matrices (e.g. in terms of vehicles and tons) that are assigned to the road, rail, sea and air networks. The Swedish model allows for consolidation of shipments within the same commodity type; in the Norwegian model, consolidation is possible within larger segments of the goods market (combinations of commodity types).

The Norwegian model has been used in many applications (including a national transport plan with long-run forecasts and various sensitivity analyses, corridor analysis, various cost-benefit analyses, the Norwegian Climate project; see Kleven, 2011). An example of long-run forecasts from this model is in the Figure below.

Figure 5. Realisations (1980-2008) and forecasts (2008-2050) by mode (2008=100) in the Norwegian national transport plan (Kleven, 2011)



The Swedish model has so far been used in a limited number of practical studies (changing the requirements for maritime fuels, impact of rail user charges, co-modality project, Sweden-Ruhr Area corridor study, see Vierth, 2011).

The freight transport model for the Mobility Masterplan of Flanders (the northern half of Belgium) also follows the ADA. The PC matrices come from an existing trade model, called Planet. Then comes a newly developed logistics model for the choice of shipment size and transport chain. This also is a deterministic model minimizing total logistics costs, that was calibrated using aggregate data on the mode shares for domestic transport, import and export of Flanders. Assignment for road transport is done jointly with the assignment of cars (de Jong et al., 2010). The ADA model for Flanders was used for several different scenarios for 2010-2040 and many sensitivity analyses. An interesting example of the outcomes of this model and the potential benefit of logistics modeling (with transport chain and shipment size optimization instead of just modal split) are the road transport costs elasticities. For the effect on tons by road, these elasticities are around 0, but for vehicle kilometers by road they are around -1 . This difference in elasticities is mainly due to:

- Shifts from direct road transport to transport chains like road-rail-road, which count twice for road tonnage, but which have much shorter road distances than direct road transport;
- Increases in shipment size.

A model system similar to that in Norway and Sweden is now being built for Denmark (Hansen, 2011b). A key change with regards to the Norwegian and Swedish models is that the Danish model will use a pivot-point procedure on the truck matrix (and similarly for sea and rail transport). This means that the logistics model will only deliver **changes** in the OD matrices between a base and a future year. These changes will then be applied to a base year truck OD matrix, which will be based to the maximum possible degree on observed data (e.g. traffic counts, surveys), making the overall model less synthetic and better empirically founded. The Danish model will make use of disaggregate data (partly from Stated Preference interviews in the Fehmarn Belt corridor).

The new Transtools 3 model (for the European Commission; see the chapter in this book by Nielsen et al.), also plans to use an approach for freight transport and logistics based on the experiences in Norway and Sweden (and Flanders, Denmark), in combination with base year OD matrices by mode, derived independently.

7. Conclusion

We presented an aggregate-disaggregate-aggregate (ADA) model system for international, national or regional freight transport, including: aggregate models for the determination of goods flows between production and consumption zones (such as input-output models); disaggregate models for logistics decisions, including shipment size, number of legs in the transport chain, use of consolidation and distribution centres and mode; vehicle type and loading unit for each leg; and assignment of the aggregate OD vehicle flows to the networks. The ADA model is specified at the disaggregate level (individual shipments) where disaggregate modelling is feasible and most attractive, but at the aggregate level where

disaggregate models are not possible or not attractive because of constraints on available data, software and runtime and for ease of interpretation.

When the middle part of a model system is disaggregate, and the parts that come before and after it are aggregate, additional disaggregation and aggregation steps become necessary. There are many ways to do a disaggregation and the problem has many feasible disaggregate outcomes. Sometimes data are available to determine which solution is most likely, but more often assumptions need to be made as the basis for disaggregation. The aggregation process however will have only one solution (though there can be different dimensions to represent the aggregate outcomes, e.g. in tons or in vehicles). It is important to do the disaggregation in such a way as to preserve the original PC zone-to-zone totals. Estimation of the disaggregate logistics model described in this paper requires data on individual shipments, which in most countries are not available to researchers. However, a calibration of such a micro-level model to data at a more aggregate level is also feasible.

In the longer run, an aggregate-disaggregate-disaggregate (ADD model), with assignment of each individual vehicle to the networks, may be possible, though it would likely incorporate a mesoscopic, not a microscopic, assignment. Modelling the international and interregional trade patterns at a disaggregate level (DDD model) may also be possible in the longer run (through choice of supplier by receivers, or the other way around). However, given the nature of the data that are available for this step (trade data from custom records, production and consumption by sector; at best, multi-regional input/output data), aggregate models for PC flows can be expected to remain mainstream.

ACKNOWLEDGEMENTS

This paper is, to a large extent, based on a series of projects carried out for the Work Group for Transport Analysis in the Norwegian national transport plan, the Samgods group in Sweden, the Government of the Flanders region in Belgium and the Technical University of Denmark (DTU). Organizations that were involved in these projects include Significance, RAND Europe, SITMA, IMOB University Hasselt, DTU, Solving International, Solving Bohlin & Strömberg and INRO Canada. The authors would like to thank Inge Vierth of VTI/CTS for comments on an earlier version.

REFERENCES

- Bates, J.J. (2011) "Freight modelling and policy analysis in The Unites Kingdom", Presentation at the *CTS-Seminar on European and National Freight Demand Models*, Stockholm.
- Ben-Akiva, M.E. (2004) "Choice analysis and market analysis for freight", presentation at the *colloquium in memory of Prof. Marvin L. Manheim*, Antwerp.
- Ben-Akiva, M.E., J. Bottom, S. Gao, H. Koutsopoulos and Y. Wen (2007) "Towards disaggregate dynamic travel forecasting models", *Tsinghua Science and Technology Journal*, 12-2, 115-130.
- Bierlaire, M., R. Mishalani and M.E. Ben-Akiva (2000) "General framework for dynamic demand simulation", Report RO-000223, Ecole Polytechnique Fédérale de Lausanne, Department of Mathematics.
- Bovenkerk, M. (2005) "SMILE+, the new and improved Dutch national freight model system", Paper presented at the *European Transport Conference 2005*, Strasbourg.
- Bröcker, J. (1998) "Operational spatial computable general equilibrium modeling", *Annals of Regional Science*, vol. 32, pp. 367-387.
- Cascetta, E. (2001) "Transportation systems engineering: theory and methods", Kluwer Academic Publishers, Dordrecht/Boston.
- Hansen, W. (2010) "Developing a new spatial computable general equilibrium for Norway", Paper presented at *European Transport Conference 2010*, Glasgow.
- Hansen, C.O. (2011) "Freight modelling and policy analysis in Denmark", Presentation at the *CTS-Seminar on European and National Freight Demand Models*, Stockholm.
- Holguín-Veras, J, N. Xu, G.C. de Jong and H. Maurer (2011) An experimental economics investigation of shipper-carrier interactions on the choice of mode and shipment size in freight transport , *Networks and Spatial Economics*, 11, 509-532.
- Hunt, J.D. and J.E. Abraham (2005) "Design and implementation of PECAS: A generalized system for the allocation of economic production, exchange and consumption quantities", Chapter 11, In: *Foundations of Integrated Land-Use and Transportation Models: Assumptions and New Conceptual Frameworks*, Elsevier, Oxford UK, 217-238.
- Hunt, J.D. (2003) "Agent behaviour issues arising with urban system micro-simulation", *European Journal of Transport Infrastructure and Research* 2(3/4), 233-254.
- Hunt J.D., R. Donnelly, J.E. Abraham, C. Batten, J. Freedman, J. Hicks, P.J. Costinett and W.J. Upton (2001) "Design of a statewide land use transport interaction

model for Oregon”, *Proceedings of the 9th World Conference for Transport Research*, Seoul, South Korea.

Ivanova, O., A. Vold and V. Jean-Hansen (2002) “PINGO: a model for prediction of regional and interregional freight transport”, Report vol. 578/2002, Institute of Transport Economics, Oslo.

Ivanova, O., C. Heyndrickx, K. Spitaels, L. Tavasszy, W. Manshanden, M. Snelder and O. Koops (2007) “RAEM version 3.0”, Transport Mobility Leuven, Leuven

Jin, Y, I. Williams and M. Shahkarami (2005) “Integrated regional economic and freight logistics modeling: results from a model for the Trans-Pennine Corridor, UK”, Paper presented at the *European Transport Conference 2005*, Strasbourg.

Jong, G.C. de and M.E. Ben-Akiva (2007) “A micro-simulation model of shipment size and transport chain choice in Norway and Sweden”, Paper submitted for publication in *Transportation Research B*.

Jong, G.C. de, J. Baak, K. Ruijs, M. Pieters, T. Bellemans, D. Janssens and G. Wets (2010) “Passenger and freight transport in Flanders 2010-2040 under three scenario’s”, Paper presented at the *European Transport Conference 2010*, Glasgow.

Jong, G.C. de, I. Vierth, L. Tavasszy and M.E. Ben-Akiva (2012) “Recent developments in national and international freight transport models within Europe”, Paper accepted for publication in *Transportation*.

Kleven, O. (2011) “Freight modelling and policy analyses in Norway”, Presentation at the *CTS-Seminar on European and National Freight Demand Models*, Stockholm.

Liedtke, G. (2005) “An actor-based approach to commodity transport modeling”, PhD thesis, University of Karlsruhe.

Marzano, V. and A. Papola (2004) “Modeling freight demand at a national level: theoretical developments and application to Italian demand”, Paper presented at the *European Transport Conference 2004*, Strasbourg.

PbConsult (2002) “Commercial transport and transport supply components of the TLUMIP second generation model”, *Third Oregon Symposium on Integrated Land Use and Transport Models*, 23-25 July, 2002.

RAND Europe (2006) “Technical report on the further development of a logistics module in the Norwegian and Swedish national freight model systems”, Report for the NTP and Samgods group, PM-2216-SIKA, RAND Europe, Leiden.

RAND Europe, Solving and INRO (2004) “The specification of logistics in the Norwegian and Swedish national freight model systems, Model scope, structure and implementation plan”, Report for the NTP and Samgods group, TR-225-SIKA, RAND Europe, Leiden.

RAND Europe and SITMA (2005) "The development of a logistics module in the Norwegian and Swedish national freight model systems, Deliverable 1b: further model specification", Report for the NTP and Samgods group, PM-1968-SIKA RAND Europe, Leiden.

RAND Europe and SITMA (2006) "Documentation and clarification of deliverable 4 and the associated program delivery for a logistics module in the Norwegian and Swedish national model systems", deliverable 4a, PM-2055-SIKA, RAND Europe, Leiden.

Samimi, A., A. Mohammadian and K. Kawamura (2010) "Freight Demand Microsimulation in the U.S.", *World Conference on Transport Research (WCTR)*, Lisbon.

SCENES Consortium (2000) "SCENES European transport forecasting model and appended module", Technical description, SCENES deliverable 4 for the European Commission DG-TREN, ME&P, Cambridge.

Tavasszy, L.A., M. van de Vlist, C. Ruijgrok and J. van de Rest (1998) "Scenario-wise analysis of transport and logistics with a SMILE", Paper presented at the 8th WCTR Conference, Antwerp.

Tavasszy, L.A., M.J.P.M. Thissen, A.C. Muskens and J. Oosterhaven (2002) "Pitfalls and solutions in the application of spatial computable general equilibrium models for transport appraisal", Paper presented at the 42nd Congress of the Regional Science Association, Dortmund.

Tavasszy, L.A., K. Ruijgrok and I. Davydenko (2012) "Incorporating logistics in freight transportation models: state of the art and research opportunities", *Transport Reviews: a transnational transdisciplinary journal*, DOI: 10.1080/01441647.2011.644640.

TNO (2008) "TRANS-TOOLS, Deliverable 6: TRANS-TOOLS Final Report", WP 6, Delft, TNO Inro (The Netherlands).

Vierth, I. (2011) "Freight modelling and policy analyses in Sweden" Presentation at the *CTS-Seminar on European and National Freight Demand Models*, Stockholm.

Vold. A and V. Jean-Hansen (2007) "PINGO: A model for the prediction of regional and interregional freight transport in Norway", TOI, Oslo

Annex 1. Practical method to disaggregate zone-to-zone (z2z) flows to firm-to-firm (f2f) flows (as implemented in Norway)

Below, the procedure is explained for a hypothetical z2z relation. This example is for a commodity type k , but we drop the subscript k for convenience.

There are 400,000 tons going from zone r to s according to the PC matrices. We want to allocate this number to firm-to-firm relations within the zone pair rs .

We know (input data for the base year) that there are 10 firms sending k from r (possibly to all s). We also know (again from the input data) that there are 20 firms receiving k in s (possibly from all r).

So within rs there could be at most $10 \times 20 = 200$ firm-to-firm relations.

From logistics experts we got estimates for the number of receivers per sender, by commodity (a national average number for each commodity type). Let there be 30 receivers per sender for k . Suppose that for Norway as a whole there are 1,000 receivers for k (all zones s) and that we also know the number of senders of k (here: 500) from all zones r . So in total for k there should actually be $30 \times 500 = 15,000$ relations. As a by-product this gives the implied number of senders per receiver: $15,000 / 1,000 = 15$.

The maximum overall number of relations for k is $1,000 \times 500 = 500,000$. So $15,000 / 500,000 = 3\%$ of the maximum number of relations materializes.

In equation form:

$$\begin{aligned} \text{Fraction} &= (\text{ReceiversPerSender} * \text{TotalSenders}) / (\text{TotalReceivers} * \text{TotalSenders}) \\ &= \text{ReceiversPerSender} / \text{TotalReceivers} \end{aligned}$$

In which:

ReceiversPerSender: average number of receivers per sender (from logistics experts).

TotalReceivers: number of receivers in all the zones in Norway (from business register data).

TotalSenders = number of senders in all the zones in Norway (from business register data).

Now for the zonal pair rs we assume that this 3% fraction is applicable. With 200 potential f2f relations there should be 6 actual f2f relations:

$$\text{Actual number of f2f relations from zone } r \text{ to zone } s = \text{fraction} * (\text{Senders} * \text{Receivers})$$

In which:

Senders: number of senders in a zone r .

Receivers: number of receivers in zone s .

We now select these 6 mn relations at random from the 200 available by using proportionality to the product of the production volume of firm m and the consumption volume of firm n for the commodity in question. Then we can divide the 400,000 tons over the 6 relations proportionally to the share of a mn relation's product in the sum of the products over all 6 mn relations. The sum of the allocated flows over the 6 relations will equal 400,000 tons (preservation of PC flow).

Sometimes, the data we use on firms in the domestic zones (production and consumption files) do not include any producing or consumption firm in a zone for a commodity type for which there is z2z information in the PWC base matrices. In these cases we have generated a single artificial firm (sender or receiver) for that commodity in that zone.

Furthermore we use some extra rules to prevent getting too many small f2f flows. For export and import flows we have no information on firms at the foreign end. We assume that there will only be one sender per zone for import and only one receiver per zone for export.