significance
quantitative research

Towards a full-fledged four-stage transport model first and an ADA model next

A thinkpiece on WP2 FAME Forecasting Methodology

Gerard de Jong | 25 February 2026

Towards a full-fledged four-stage transport model first and an ADA model next

A thinkpiece on WP2 FAME Forecasting Methodology

Gerard de Jong | 25 February 2026

Authors:
Gerard de Jong | Significance

Project number:
[26003]

Version control:

Version	Date	Description
V1.0	13-2-26	Draft for discussion
V2.0	25-2-26	Final report

Table of contents

1. Introduction

DfT has requested bids for thinkpieces on four different topics in relation to the new Freight Analysis and Modelling Environment (FAME)[1]. After having evaluated the bids, DfT awarded a contract to Significance for a thinkpiece on one of the four topics, the forecasting methodology (WP2). This document contains Significance's thinkpiece for this topic.

We agree with Phil Goodwin (Goodwin, 2024) that a model that is used for the appraisal of schemes that change transport times and costs should include, whenever possible, the various impacts of these factors on both passenger and freight transport (i.e. variable demand modelling). This includes impacts on the modal split in freight transport and on the spatial distribution of freight flows. Insight in these impacts is needed for a good analysis of the costs and benefits of these schemes.

DfT's brief on this topic requires that the recommendations in each thinkpiece should provide further detail or alternative suggestions than is already provided within the FAME Feasibility Report issued by MDST in 2024 (MDST, 2024). In this thinkpiece, we work out a model structure and specification that goes beyond the Feasibility Report. The proposed new model development consists of two phases, both of which extend variable demand modelling. In the first phase, an aggregate four-stage model would be developed in which the production/attraction and distribution stages of the forecasting model are worked out into an economy model that predicts how PC flows (see section 2.1 for an explanation) change as economic indicators and transport times and costs change. This is currently lacking in the recommendations in the Feasibility Report.

The second phase builds on the first one by replacing its aggregate modal split model by a disaggregate model for transport chain choice, so that the model as a whole becomes an aggregate-disaggregate-aggregate (ADA) model. Both phases are further worked out below, also including a discussion on data requirements and availability, and of validation of model outcomes against the data and against elasticities from the literature for each phase.

Section 2 of this document discusses the aggregate four-stage model that includes an economy model for forecasting long-run changes (phase I). The extension to an ADA model (phase II) is discussed in section 3. In section 4, a summary of the recommendations in this thinkpiece is given. Annex A briefly discusses some ideas for new data collection.

[1] Actually, there is also a freight transport model from the US called FAME (Freight Activity Microsimulation Estimator), by Samimi et al. (2010).

2. Phase I: a full-fledged aggregate four-stage transport model starting from PC flows

2.1 Introduction to phase I

Just as the peer reviewer Ian Williams (Williams, 2024), we support most of the recommendations of the Feasibility Report. With regards to zones, modes, vehicle types, commodities and domestic/international the model that we specify in this section for development in a first phase can be fully in line with the recommendations from the Feasibility Study.

However, we propose alternative solutions for the PC flows and the link to OD patterns (again similarly to Ian Williams). The Feasibility Report takes a step in the direction of variable demand modelling (VDM) by proposing an aggregate logit modal split model that would include transport time and cost. On the total freight transport volume (all modes), its recommendations are not worked out in detail but clearly exclude an economy model that predicts future PC flows and that DfT could run itself as part of the new freight model. Our interpretation of the Feasibility Report is that it proposes a gravity model (with explanatory variables like floorspace, employment and generalised transport costs), but only to determine the base year OD matrices, not for forecasting the future and not for PC-based modelling.

2.1.1 Variable demand modelling (VDM)

TAG Unit 2.1 on variable demand modelling (Department for Transport, 2025) says under 1.1.5: “Any response in the demand for transport of freight is not considered here, and it is generally assumed that total freight traffic is fixed, but susceptible to re-routeing.” It rules out variable demand modelling (VDM) for (road) freight, only leaving route choice/network assignment to react to changes in transport times and costs. What can be included within VDM for passenger transport can be found under 2.3.3: “Typically, the demand model should address the responses of frequency, mode choice and destination choice, as well as, in some instances, time of day choice, and should be applied on a Production/Attraction (P/A) basis.” For freight transport however, VDM can include similar choices (though destination choice might be production location (supplier) choice given the destination) and a number of other choices as well (see the overview of possible reactions to price changes in freight transport in de Jong et al. (2010), that was also quoted in Goodwin (2024). For some of these choices, there is national and international literature that yields transport price elasticity values. These have been reviewed several times (e.g. in de Jong et al., 2010; Tavasszy and de Jong, 2014, chapter 9; de Jong et al., 2023). The picture that emerges is that the price elasticities in freight transport are unlikely to be zero, the average values often being of similar magnitude as those for passenger transport (price elasticity values will be presented in section 3.6 of this thinkpiece). On most roads, freight vehicles constitute a minority of all vehicles (say 15%). However, given that the values of travel time for LGVs and HGVs are quite substantial, showing a substantial willingness of decision-makers in freight transport to pay for time savings, the impact of traffic schemes on freight traffic is likely to be non-negligible. We therefore think that models used in appraisal should include at least the most important reactions to time and price changes in freight transport. This may not be needed for all kinds of road investment schemes (e.g. not for small local projects) but should certainly not be ignored for motorway projects.

2.1.2 Production-Consumption (PC) versus origin-destination (OD) flows

At the macro-level, a distinction can be made between PC (production-consumption) flows and OD (origin-destination) flows. This is the macro-level equivalent of the distinction at the micro-level between a transport chain, and a leg of a transport chain (see Figure 1). The transport chain can potentially be a multi-modal transport structure whereas the leg is a uni-modal structure that is part of a transport chain. By adding the volumes of the transport chains to and from the same zones, one obtains PC matrices. By adding the volumes of the legs to and from the same zones, the OD matrices are obtained. PC matrices contain commodity flows all the way from the production zone to the consumption zone. These flows may consist of several OD flows. This is because a transport chain may be used with multiple modes and/or vehicle types as well as one or more transhipments along the route.

PC flows represent economic relations and transactions within and between different sectors of the economy. Changes in final demand, international and interregional trade patterns, as well as in the production structure of the economy, have a direct impact on the PC flow patterns. Also, the available data on economic linkages and transactions are in terms of PC flows. They are not in terms of flows between producers and transhipment points, or between transhipment points and consumers. Changes in logistics processes such as the number and location of terminals, and in logistics costs, have a direct impact on the composition of the transport chains. This would lead to different OD flows that would only indirectly impact the economic and trade patterns (and hence the PC flows).

Assigning PC flows to the transport networks require that they be first transformed into their corresponding OD flows. For instance, the transport chain road-sea-road would lead to road OD legs ending and starting at different ports. This would take the place of a long-haul road transport that would not involve any ports. A similar argument holds for a purely road-based chain that first uses vans to ship goods to a consolidation centre, then consolidates the goods with other flows of goods into a large truck for the main haul and finally uses a van again to deliver goods from a distribution centre to the C destination. In this scenario, the link assignment depends on the three OD legs. Therefore, converting the PC flows into OD flows, which can be done by a 'logistics model', (see de Jong and Ben-Akiva, 2007) is required for a meaningful network assignment. The data available for transport flows from traffic counts, roadside surveys and interviews with carriers are also at the OD level, not at the PC level.

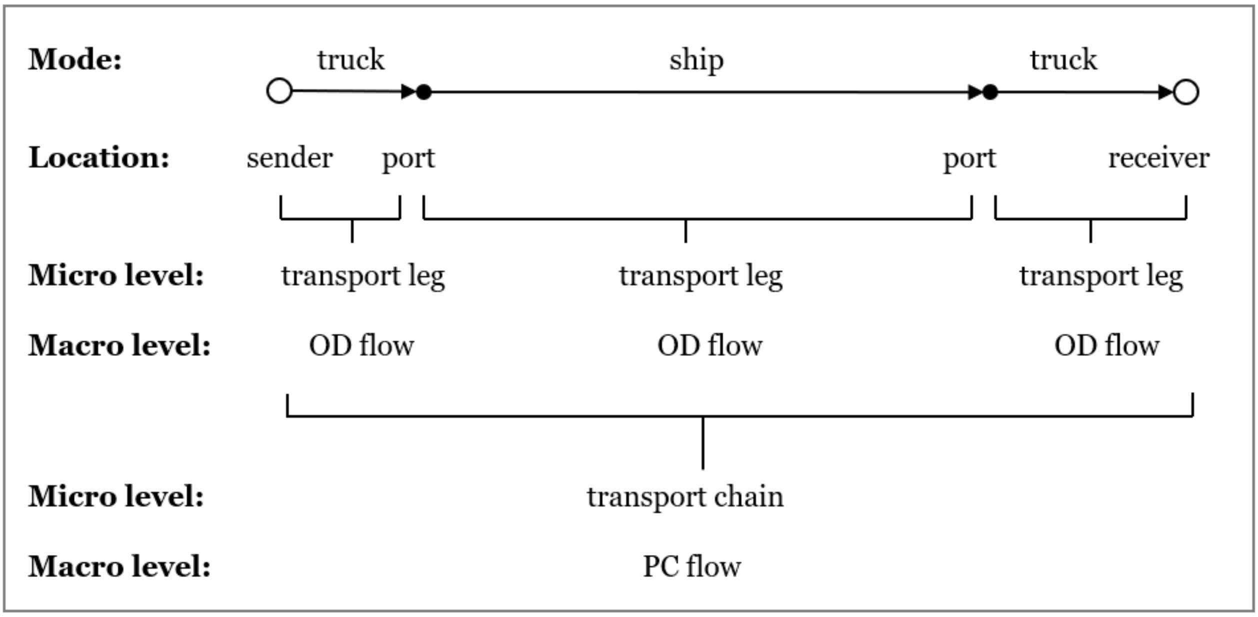

Figure1: Transport chain, transport legs, PC flows and OD flows

In this section of the thinkpiece, we specify such a PC model. An advantage of using the PC level in the first and second stage of the four-stage model is that it relates the flows to economic developments at the locations where goods are produced and consumed, instead of using intermediate terminals/distribution centres that act as origins and destinations in the transport statistics.

2.1.3 Overview of different model types for PC flows

The economy module for PC flows could for instance be a gravity model (as in the national freight transport models of The Netherlands (de Jong et al., 2011), Denmark (Willigers et al., 2025) and Sweden (Abate et al., 2019) or an input-output model (as in the national freight transport models of Italy (Cascetta, 1997) and Austria (Grebe at al., 2020)). The transport time and cost from the subsequent transport chain choice model can be used (for instance in the form of logsums) as explanatory variable in the gravity model (as in the Dutch and Danish national models). This adds another choice dimension (besides modal split) to the variable demand modelling: when time or cost increases, different zone pairs can be selected in such a way as to reduce the trip lengths.

2.2 The four-stage model and an alternative more comprehensive three-stage model

Many modelling concepts applied in freight transport forecasting have originally been developed for passenger transport. Many authors seem to agree that the four-step transport modelling structure from passenger transport can fruitfully be applied to freight transport as well. However, within each of the four steps the freight models can be very different from those in passenger transport. Important differences between the freight and passenger transport markets are the diversity of decision-makers in freight (shippers, carriers, intermediaries, drivers, operators), the diversity of the items being transported (from parcel deliveries with many stops to single bulk shipments of hundred thousands of tonnes) and the limited availability of data (especially disaggregate data, partly due to confidentiality reasons).

The four steps in the context of a freight transport model system are (also see Figure 2):

- Production and attraction. In this step, the quantities of goods to be transported from the various production zones and the quantities to be transported to the various consumption zones is determined (the marginals of the production-consumption (PC) matrix, but several models use origin-destination (OD) matrices here). The output dimension is tonnes of goods. In intermediate stages of the production and attraction models, the dimension could be monetary units (trade flows).
- Distribution. In this step, the flows in goods transport between origins and destinations (cells of the PC or the OD matrix) are determined. The dimension is tonnes.
- Modal split. In this step, the allocation of the commodity flows to modes (e.g. road, rail, inland waterways) is determined.
- Assignment. After converting the flows from tonnes to vehicle-units, they can be assigned to networks (in some models this is about assigning truck flows together with passenger cars to road networks).

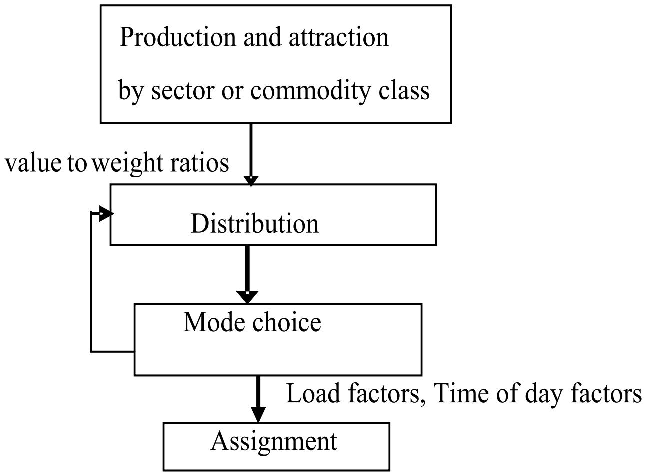

Figure 2: The conventional four-step model in freight transport

However, additional steps are often needed to transform trade flows in money units to physical goods flows in tonnes and further into vehicle flows with specific vehicle utilisation factors. These additional processes can be modelled as fixed rates, but also by explicit representation of logistics choices. Also, other logistics aspects that are related to the trade-off between transport and inventory costs are usually not included in freight transport models, even though logistics solutions of firms influence the mode split.

Besides these four steps, a number of transformation modules are usually required within a comprehensive freight transport model system. Such transformations could involve converting trade flows in money units into physical flows in tonnes to determine production and attraction. This can be done by using value-to-weight ratios for different commodity groups. The ratios used here may have a large impact on the final predictions and therefore it is important to assemble good data on the conversions and, if possible, to make it an endogenous, policy sensitive choice within the model system. Another transformation module is that for going from flows in tonnes to vehicle units, such as heavy goods vehicles (HGVs), as might happen between mode choice and assignment. Actually, this is influenced by a great number of decisions on shipment frequency, shipment size, return loads and vehicle utilisation rates. These decisions could be modelled explicitly in additional logistic modules, but often fixed conversion rates (load factors) are used here as well. Another type of transformation module is a regionalisation module to go from a coarse to a fine zoning system. Time of day in freight transport models (e.g. AM peak, PM peak, off-peak) is usually determined on the basis of fixed factors as well. One of the few freight transport models that explicitly models time of day choice is the model for Flanders (see de Jong et al., 2016). The MASS-GT model for the Rotterdam area and the Dubai Freight Model have a micro-level submodel for tour formation and scheduling of delivery trips (see De Bok et al., 2018, 2020)

A better representation of the freight transport model system might be the following.

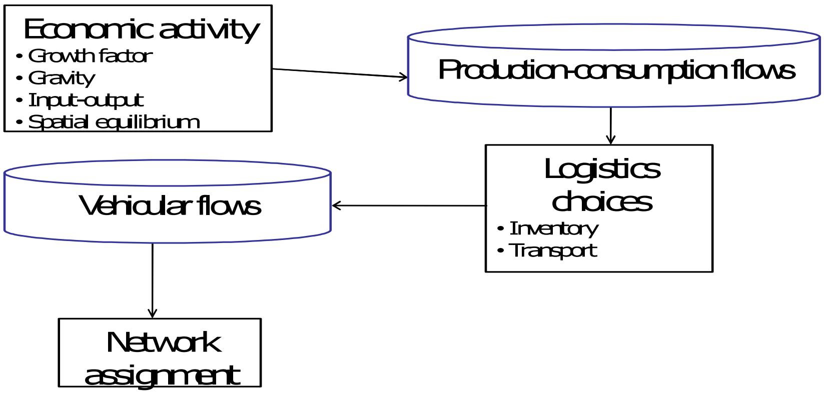

Figure 3: A revised representation of the structure of the freight model system

This structure has fewer steps (only three models: economic activity choices, logistics choices and route choice), but each of the first two steps includes several related choices. This model representation is consistent with the distinction between PC flows and OD flows that was discussed earlier in this section.

An important question then is whether each of the submodels (the three groups of choices) in Figure 3 should be modelled at the aggregate or the disaggregate level. Each of the choices in freight transport can be analysed on either *aggregate* data, the data at the macro-level of geographic zones, or *disaggregate* data, the data at the level of the decision-maker such as shippers and carriers. There are partly disaggregate national freight transport models in for instance Italy and Scandinavia. But most freight transport models are aggregate models that are sometimes highly segmented by commodity type. We think that, given the current data situation and computer technology, much can be said for modelling the logistics choices at the disaggregate level and the other choices at the aggregate level. This leads to the aggregate-disaggregate-aggregate (ADA) model (for national/international freight transport), which will be discussed in Phase II of the thinkpiece.

In phase I, we will focus on the models for production, attraction and distribution (together: the trade model or the economy model or the PC model). This discussion follows the reviews of freight transport models in de Jong et al. (2004, 2013, 2021) but also extends and updates these and applies the lessons learned to the model aims and the data situation in the UK. Modal split, the logistics choices and the assignment step will be discussed in more detail in section 3 of this thinkpiece.[2]

2.3 Different types of PC models

Within this first step we can distinguish the following types of models that have been applied in practice:

- Trend and time series models;
- System dynamics models;
- Zonal trip rate models;
- Regression models for production and/or attraction and gravity-based distribution models;

[2] The Feasibility Report also recommended some submodels that are not not commonly included in national or regional freight transport models, such as an HGV and LGV fleet model, an HGV driver labour market model and maybe also a port competition model (for UK ports), which we all support. The Dutch national freight transport model includes prototypes for a freight vehicle fleet model and for a port competition model for the Hamburg-Le Havre range (following Zondag et al., 2010 where the port choice depends on sea transport characteristics, port characteristics and hinterland characteristics, together constituting a full maritime transport chain).

- Input-output models;
- Input-output models for production and/or attraction and gravity-based models for distribution;
- Spatial Computable General Equilibrium (SCGE) models.

All these models are based on aggregate data. In the international literature we have found very few examples of disaggregate production, attraction or distribution models in freight transport (Samimi et al., (2010) model the sourcing decision of receiving firms and Gliebe et al. (2014) use game theory for matching suppliers and receivers; both in the US).

In trend models, historical trends are extrapolated into the future. Time series data have been used to develop models of various degrees of sophistication, ranging from simple growth factor models to complex auto-regressive moving average models (e.g. Garrido, 2000). The latter model uses information only on truck flows and is meant for short-term forecasting. Time series models with explanatory variables (not just explaining a variable from its own past time series), such as gross domestic product GDP, have been developed as well.

ASTRA (Assessment of Transport Strategies) was developed in a project for the European Commission and is an example of a system dynamics model (ASTRA, 2000). It is still used by the European Commission in conjunction with the network assignment model TRUST (TRT, 2018), where ASTRA produces growth factors over time to upscale the matrix. In ASTRA, the changes in the transported quantities over time and feedbacks to/from the economy, land use and the environment are modelled explicitly. In the macro-economic module of ASTRA, growth of GDP is predicted. The results are fed into the regional economics module, which yields freight demand in terms of flows in tonnes by OD pair. In the transport submodel this is assigned to modes and virtual links. Changes in transport demand in turn may affect GDP, through the transport cost. The parameters of a system dynamics model system are usually not obtained from statistical estimation, but from existing literature and by trying initial values and checking the resulting dynamic behaviour of the system (trial and error). A system dynamics model might include both the economy and modal split steps. System dynamics models, however, usually do not contain sufficient spatial and network detail to yield zone-to-zone flows and link loadings.

Zonal trip rates for production and attraction are usually derived from classifying cross-sectional data on transport volumes to/from each zone in the area under investigation (or another similar area) into a number of homogeneous zone types. Examples of such rates from the US can be found in the Guidebook on Statewide Travel Forecasting (FHWA, 1999) or chapter 3 by José Holguín-Veras, Miguel Jaller, Ivan Sanchez-Diaz, Shama Campbell and Catherine T. Lawson in Tavasszy and de Jong (2014).

An approach used in several models is to employ regression models for production and attraction, use the outcomes of these as the row and column totals of the forecasted matrices and then rely on gravity modelling to determine the cell values of the matrices (double-constrained gravity model). The (cross-sectional) regression model can use GDP, employment or population of production and/or location zones and can be specified as a linear, loglinear or double-logarithmic (constant elasticity of substitution) model. In the double constrained gravity model for distribution the trade flows (in tonnes) between production zones and consumption zones are determined based on measures of production and attraction (the outcomes of the regression models) and a measure of transport resistance. The latter is expressed as transport cost or generalised transport cost. More specifically, in such models the flow between zone i and zone j is a function of the product of production and attraction measures of zone i and zone j respectively divided by some measure of the (generalised) transport cost. The regression and gravity model approach fits with PC data (because these give the locations for which economic growth is relevant), but is sometimes used for OD data, even though this is not suitable.

Double-constrained gravity models for distribution in freight are included in:

- The PC base matrix for Sweden (estimated largely, but not exclusively, on the Swedish CFS, Edwards, 2008);
- The Danish national freight transport model LTM (Vejdirektoratet, 2021).

In these double-constrained models, there is still an influence of transport time and cost on the matrices: higher cost or time will lead to shorter trips, but the degree of response is limited by the fact that the row and column totals cannot change when cost or time change.

In the freight model for Sund & Bælt (the agency that administers the large existing and potential fixed links in and around Denmark), a single constrained gravity model is used: here the gravity model includes the split over production zones, given the consumption zone (Tapia et al., 2025). Behaviourally this mimics the sourcing decision: the choice of supplier by a receiving firm. This is the reverse of destination choice modelling in passenger transport modelling where the origin is kept fixed. In freight transport we think that production zone choice is more relevant/realistic than consumption zone choice. The single constrained gravity model is combined here with a regression-based attraction model. This combination includes an extra sensitivity to changes in cost and time, the choice of production zone, that is fixed in the double-constrained gravity model.

Even more flexibility can be found in the Transtools3 model for Europe (de Jong et al., 2017). This model uses an unconstrained gravity model, so both production and consumption can react to changes in time and cost, in this case through a logsum variable that was calculated from a transport chain choice model.

In modern gravity models (e.g. Anderson and van Wincoop, 2003) trade between two countries (and the same applies to zones within countries) is not simply determined by the absolute trade costs between the two countries, but by the relative trade cost (the trade cost of country i from importer j relative to its overall trade cost for all the countries from which it imports). In an empirical gravity model, this can be taken into account by adding multilateral resistance terms. However, a simpler method is to use importer or exporter fixed effects (de Jong et al., 2017).

Input–output (IO) models are basically macro-economic models that start from input-output tables. These are tables that describe, in money units, what each sector of the economy (e.g. textile manufacturing) delivers to the other sectors, also including the final demand by consumers, import and export. National input-output tables have been developed for many countries, usually by a central statistical office. A special form of input-output table, which for many countries does not exist, is a multiregional input-output table (MRIO), also known as a spatial input-output table. A multiregional input-output table not only includes deliveries between sectors, but also between regions (trade flows). Please note, as for the gravity model, that the regions here should be PC regions, not OD. Most multi-regional input-output tables distinguish only a few, large regions within a country. The input-output model assumes that for forecasting, the multi-regional input-output table can be scaled up on the basis of predicted sectoral growth. The new input-output table can then give the future trade flows between regions, using either:

- Fixed technical and trade coefficients: the present production and trade patterns are extrapolated into the future;
- Elastic technical and trade coefficients: functions are estimated (e.g. multinomial logit) in which the fraction that is consumed in region i of the production of sector s in region j depends on the total production of region j in sector s and the (generalised) transport cost, in relation to other regions. This makes production/attraction/distribution sensitive to changes in transport cost and time.

A multi-regional input-output model can thus supply both production/attraction and distribution. Examples of multi-regional input-output models in freight transport are:

- The Italian national model system for passengers and freight (Cascetta, 1997), which uses 17 sectors and 20 regions and also has elastic coefficients. In this model, the freight OD flows follow from a multi-regional input-output analysis with elastic coefficients (after transformation from money units into tonnes and after regionalisation);
- The SCENES European model system for passengers and freight and its predecessor STREAMS, Strategic Transport Research for European Member States (Leitham et al, 1999), with 33 sectors and more than 200 zones in Europe and elastic coefficients (SCENES Consortium, 2001).

The Dutch national freight transport model BasGoed (de Jong et al., 2011; Rijkswaterstaat, 2022) uses a double-constrained gravity model, where the row and column totals come from an input-output model, not from regression models.

Another option for production, attraction and distribution is the computable general equilibrium (CGE) model, that establishes equilibrium in several related markets (not only transport, but also goods markets, labour markets, land markets). CGE models in economics (not focussing on transport) often include economic issues that are not handled in transport models, such as type of competition and economies of scale. Spatial CGE (SCGE) models (e.g. Bröcker, 1998; Tavasszy et al., 2002; Ivanova et al., 2006; Chapter 2 by Olga Ivanova in Tavasszy and de Jong, 2014) have been developed, that might become operational models for practical (inter)national and regional forecasting, but there is still a long way to go before these models could work as detailed transport models or be combined with detailed transport models. The main advantage of SCGE models is also their main disadvantage: these models are very broad (transport, economy, land use) and can give interactions between markets that other models are not capable of, but the way in which transport is depicted in these models is often highly simplified. On the other hand, the European model TRIMODE (Martino et al., 2018) developed for the European Commission, also contains an SCGE Economy model to determine the PC flows. In principle a SCGE model could also be used for PC flows in/to/from the UK, but we expect that this will be too complicated for this purpose and possibly unstable.

The advantages and disadvantages of the different model types for the PC model are listed in Table 1. This table makes clear that trend and time series model and zonal trip rates by themselves are unlikely to include impacts of transport time and cost and therefore can be ruled out for variable demand modelling. System dynamics models can be omitted because they are only evidence-based to a limited to moderate degree (no econometric/statistical estimation). SCGE models are not sufficiently focused on solving the lack of transport demand sensitivity: they include all kinds of other markets than transport but have difficulty in providing enough detail on transport itself to be useful as a transport model or a part of it. They cover much more than just trade relations and take a very long time to develop, implement and test.

What remains as options for a new PC model for the UK are therefore the combination of cross-sectional regression models with gravity models, elastic input-output models by themselves and the combination of input-output and gravity models. The choice between these three options depends to a large extent on the data requirements and data availability in the UK. This is discussed in section 2.5 below.

2.4 The subsequent steps: the aggregate chain choice model and network assignment

After the PC model follows in phase I the aggregate split model for transport chains (e.g. OGV2-rail-OGV2 and OGV2-inland waterways-OGV2, but examples also contain a direct OGV2 chain and a direct block train) between the PC locations. Apart from transhipments between modes, transport chains can also distinguish several vehicle types, such as the chains OGV1-OGV2-OGV1. The shifts to a different mode or vehicle type take place at consolidation centres, distribution centres and intermodal terminals (including ports). By using transport chains instead of modes, the transport chain choice model becomes the link between the PC flows and the OD flows. It takes the PC flows from the PC model and allocates these to transport chains, the legs of which are the OD flows. The OD flows can be collected from the transport chain outcomes (as an example the first chain mentioned above results in one rail OD trip and two OGV2 trips). The modes covered in this transport chain choice model would be road, rail, inland waterways and short-sea shipping (with air transport and pipelines included only as generators). Within road transport, the vehicle types LGV, OGV1 and OGV2 would be distinguished, as recommended in the Feasibility Report.

Table 1. Summary of freight transport production, attraction and distribution models

Type of model	Advantages	Disadvantages
Trend and time series models	Limited data requirements (but for many years)	Little insight into causality Limited scope for policy effects
System dynamics	Limited data requirements Can give land use interactions External and policy effects variables can be included	No statistical tests on parameter values
Zonal trip rates	Limited data requirements (zonal data)	Little insight into causality Limited scope for policy effects
Regressions for P and C, gravity for cell values	Limited data requirements Some link to economy and economic theory Policy effects through transport cost function in gravity model	Limited scope for including explanatory factors Limited number of calibration parameters (which rules out a good fit to the data)
Input-output	Link to the economy Link to economic theory Can give land use interactions Policy effects if elastic coef-ficients	Need input-output table (data), preferably multi-regional Restrictive assumptions if fixed coefficients Need conversion from values to tonnes if data only in money values
Input-output for P and C, gravity for cell values	Link to the economy Link to economic theory Policy effects through transport cost function in gravity model	Need input-output table (data), preferably multi-regional Need conversion from values to tonnes if data only in money values
SCGE	Link to the economy Based on economic theory Can give land use (and other) interactions Can give policy effects	Needs at least as much data as input-output Includes many assumptions Needs conversion from values to tonnes Mostly still in development phase

The aggregate transport chain choice model can be estimated on the OD data for the different modes (rail statistics, waterway statistics, (air)port statistics, CSRGT) as inputs for the dependent variable. The difference with an aggregate modal split model is that for zones that do not have railway platforms or inland waterways quays, block trains and direct waterway transport are not available, but road access-egress to container and wagonload trains and vessels is possible through transshipment at a terminal. There are several ways for handling these road access/egress parts of the chain. The simplest way is to use an (empirical) average distance to/from the rail or inland waterways terminal in the transport cost calculation. This does not give terminals as origin and destination. To get it, one needs a spatial list of terminals with road access and egress links as part of a road network. Then one can use the least cost solution for each chain type (e.g. OGV2-rail-OGV2) as the choice alternatives in the transport chain choice model (as used in the Swedish national freight transport model), or the shortest access/egress distance to/from a terminal of a certain type (used in the Dutch national freight transport model BasGoed.

The OD flows from the transport chain model in tonnes can be converted to flows in numbers of vehicles by using appropriate load factors and empty vehicle rates. The predicted relative shifts over time in the OD flows in vehicles can then be pivoted on OD base year vehicle matrices, that are based on the data at the OD level (transport statistics, OD surveys and count data) and assigned to networks. It will probably be best to use existing passenger transport models for this by adding the freight vehicles, because both types of road vehicles cater for the same road capacity.

2.5 Data requirements for PC models and data availability in the UK

PC flows start at the location of producers, but if wholesalers are distinguished in the data used, they can also start from wholesalers. A wholesaler differs from a terminal, consolidation and distribution centre. It marks a change of ownership of the goods, as can be recorded in the data on economic transactions. Using terminals, consolidation and distribution centres for transshipment and storage does not entail a change of ownership of the goods. The PC flows end at the location of consumers of the goods, which (also depending on the data used) can include manufacturers that do further processing, wholesalers, retailers and households. It is usually not possible in gravity or IO models (or any of the production, attraction and distribution models discussed above) to distinguish between physical shopping (these trips are part of the passenger transport model) and online shopping with home delivery. It seems best to abstract from these very last mile considerations in a national PC model and exclude the households as C locations. This makes sure that the shopping trips will not be double counted. If one would be very interested in modelling delivery trips to households (currently more a topic for urban models than for national ones), one could add a sub-model for the tour formation decisions of the delivery firms as a final transport demand modelling step (as in de Bok et al., 2018, 2020).

An overview of which type of data source can be used for which type of freight transport model can be found in chapter 10 by Gerard de Jong on data availability and model form of Tavasszy and de Jong (2014). The key points are:

- A gravity model, together with regression functions for production and/or attraction requires data from trade statistics for the base year. A base year PC matrix from FAME WP1 (the project on base year matrices) could also be used for the estimation of the regression and gravity model. One might have to be careful here if this PC matrix itself is also based on modelling methods (and one would be modelling on model data), but given that this PC matrix is likely to be the best approximation available of the true PC flows, it would be best to use this as input in the forecasting model development and be consistent with it;
- An input-output model requires input-output tables for the flows between sectors or between products from the national accounts for the base year;

- Transport statistics (CSRGT and other road freight statistics, port statistics, airport statistics, rail statistics, domestic waterborne statistics) are usually by mode and can only be used for an OD-based model, which is less appropriate for production/attraction/distribution than a model at the PC level. They can be used to validate the predicted OD flows from the transport chain model that follows the PC model (i.e. is conditional on its outcomes). Vehicle flow data (from traffic counts and similar) can be used for OD matrices in terms of numbers of vehicles and for validation at this level.

An important issue in this regard is the zoning level of the trade data and the input-output data. The estimation of the gravity model requires international trade data for flows to/from the UK, but also regional trade data, for trade flows within the UK and for the domestic production and consumption locations of the international trade flows with the UK. Some data gaps for regional detail are allowed, since a gravity model can be estimated on a system with a more limited number of zones and then applied to a finer zoning system to produce forecasts at the finer geography.

For the input-output data, it is important that these are not just national tables, but regional ones, including information on flows between zones/regions (multi-regional input-output MRIO data). A multi-regional input-output table with larger domestic and international zones can be regionalized (converted to smaller zones) by using a gravity model approach, as was done for the input-output table in the Austrian national freight transport model (Grebe et al., 2020). Input-output data also use different classification systems of sectors and commodities than a transport model (that will use commodities from the transport statistics) and the conversion from one to the other may not be straightforward.

For the UK, input-output data are available from the national accounts at the ONS website (https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/datasets/ukinputoutputanalyticaltablesdetailed for input-output data on the basis of the CPA product classification and https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/datasets/ukinputoutputanalyticaltablesindustrybyindustry/current for input-output data on the basis of the SIC sector classification). Both tables are available at the moment up to 2023, are only at the national level (so an IO-table, but no MRIO-table) and are both denoted in money units not tonnes. In order to make these tables useful for developing a MRIO model, one would have to:

- Convert the product codes (CPA) or sector codes (SIC) to the commodity types used in the transport model (presumably based on the NST codes);
- Convert the IO table in money units to an IO table in tonnes, using commodity-specific value-to-weight ratios;[3]
- Convert the national IO table to a multiregional one, using a gravity approach.

So, for the creation of the database for input-output modelling, considerable efforts would be required. Furthermore, all these steps (especially the second and third) diminish the quality of the final result, since representative value-to-weight ratios are hard to get and the regionalisation of a national table would be a crude process. A further data problem with IO data is that the national accounts follow the money flows and not (necessarily) the physical flows – there could for instance be large flows of

[3] A disadvantage of using IO models could be that the IO data and models are in money units. A freight transport model should produce flows in terms of tonnes, tonne-kilometres, vehicles and vehicle-kilometres. In transport model systems with IO models, a conversion takes place from money units to tonnes using fixed value-to-weight ratios. These ratios are usually based on mean values from the trade statistics and mean weights from the transport statistics for commodity classifications that are assumed to be uniform. This is often a weak point of model systems using the IO approach, or indeed any other economic model. We recommend strengthening this weak link by carrying out specific surveys among shippers asking about the value and the weight for the same shipment and about its frequency (see Annex A).

agricultural goods for further processing into the central parts of London if that happen to be the locations of the head offices of the firms involved.

The trade statistics from HM Revenue & Customs https://www.uktradeinfo.com/trade-data/) include international as well as regional tables. They use the SITC trade classification or HMRC's commodity classification (which is the same as the EU's). The international tables provide the countries at the foreign end. At the domestic end, large regions of the UK are used (e.g. NE England, NW England) as well as UK ports. The trade tables are all in money units; tables in tonnes were not found for recent years. For the estimation of gravity models per NST commodity group one would have to do the following data processing first:

- Convert the commodity codes in the trade data/SITC codes to the commodity types used in the transport model (presumably based on the NST codes);
- Convert the trade tables in money units to tonnes, using commodity-specific value-to-weight ratios;
- Convert the large regions and ports to more detailed zones.

The first step is relatively standard, the second is more problematic (as for IO models). The third step is not as large as for the IO data (where the starting point is purely national) and commonly done in gravity modelling by estimating the gravity model at a more condensed level and then applying it at a finer zoning level. This does require data on GDP, employment, population or land use at the finer zoning level, but these are available.

2.6 Which model type is recommended for the FAME PC model?

In a number of best-practice transport models, such as the Italian and Austrian national model, input-output models are used. However, some other models use a gravity approach (Sweden, Denmark, BasGoed model in The Netherlands). The major advantage of IO models over gravity models is that they use more data on the economic structure (such as linkages between sectors and regions). This gives the IO model a better foundation in economics (theory and statistics) than a gravity model. If up-to-date multiregional IO data would be available for the UK, we would recommend using an IO approach for the FAME PC model. To include the whole range of factors such as economic development, world market prices, production, consumption and trade in a freight transport model, the only possible way seems to be to use an economic model (such as IO) that uses money units. However, this advantage disappears when no multiregional IO data is available, or when the data do not give a reliable picture of freight flows (after conversion to tonnes). As discussed in section 2.5 on the data, recent multi-regional IO tables from the national accounts are not made available for the UK. In the past, agencies have carried out regionalisations, but these suffered problems of flows being recorded to/from main offices instead of local units. Doing a new regionalisation together with building a new IO table is substantially more work than the calibration of a gravity model.

Given the above data problems and the relatively limited time and budget that is needed for developing a gravity model, we recommend using the combination of regressions for production and attraction and a gravity approach for distribution. More specifically we recommend developing a single constrained gravity model together with a regression-based attraction model. This gravity model then takes the attractions as given and models the choice of production zone (sourcing). This combination will then include an extra sensitivity to changes in transport cost and time (as part of variable demand modelling), going beyond the double-constrained gravity model.

We recommend that the freight transport demand model (including the PC model and the transport chain choice model) be developed at what is called 'tier 2' in the Feasibility Report: 374 zones in the UK (ITL3/NUTS3) plus NUTS3 in the Republic of Ireland, NUTS1 and port zones in continental Europe and world regions outside Europe. A zoning system with hundreds of zones instead of thousands is quite common in a national freight transport modelling (e.g. Norway, Sweden, Austria). The Dutch national freight transport model (for a geographical smaller country) has only 45 domestic zones. In freight transport, zones can be larger than in passenger transport, because most trips in freight transport are much longer than in passenger transport (so having small zones to reduce the share of internal traffic and to obtain accurate level-of-service to/from the centroids is not so important). This tier 2 freight transport demand model can then also be much faster in operation than a model with thousands of zones. Before network assignment, the tier 2 matrix predictions are converted to tier 1 (8,985 zones in the UK plus the international zones), using tier 1 data on employment, population, GDP or land use, for joint assignment with cars, This kind of conversion before assignment is also carried out in the national freight models of Denmark, The Netherlands and Austria.

2.7 Validation of model outcomes against elasticities from the literature for trade models

Validation of the model outcomes against OD-data (in tonnes and in vehicles) has been discussed in section 2.5 above. Another type of model validation compares the sensitivities to changes in GDP, transport time and cost implied by the model against elasticities from the literature. A recent overview of the international literature on elasticities in passenger and freight transport was carried out by de Jong et al. (2023) to provide benchmarks for assessing the sensitivities of the Danish national model. This also includes the impact of changes in GDP on tonne-kilometres over all modes together from trade models. The 80% interval from the literature was established to range between 0.5 and 0.8, with a most likely value for the GDP elasticity of freight tonne-kilometres of 0.65. The review includes a number of studies from the UK: Piecyk and McKinnon (2013), Dunkerley et al., (2015) and Rohr et al. (2018).

Only limited literature was found on the impact of transport cost through production/attraction/distribution on freight tonne-kilometres. A few studies from Denmark, The Netherlands and Canada show a distribution effect that is just a few % of the mode choice effect. But there are also a few studies (Denmark again) where the distribution effect is substantial and even exceeds the mode choice effect.

3. Phase II: further development into an ADA model

3.1 Introduction to phase II

The text below on phase II is partly based on Ben-Akiva and de Jong (2013), updated with recent experience in several countries and applied to the UK situation in terms of model aims and data.

In phase II, the aggregate economy model and the network assignment remain the same as in the phase I model, but the aggregate transport chain model is replaced by a disaggregate transport chain model (at the level of annual firm-to-firm flows). The model as a whole then becomes an ADA model, as has been developed for the official national freight transport models of Norway, Sweden, Denmark and Austria and also for Flanders, Dubai and Indonesia (Hidayati et al., 2024). These models were originally developed in the first decade of the 21st century (see de Jong and Ben-Akiva, 2007) to include more aspects from private sector logistics decision-making in public sector forecasting models. The Feasibility Study largely missed these developments, but the peer review by Ian Williams includes ADA (Williams, 2024).

The aggregate-disaggregate-aggregate (ADA) model is a framework for freight transport models at the international, national or regional level that distinguishes three model steps:

- An aggregate trade model (this can a gravity-based model as recommended in section 2), that explains trade between the location of production and the location of consumption (the PC flows).
- A disaggregate (firm-to-firm flow level) logistics model for the choice of transport chain and shipment size; in this way the logistics model translates the production-consumption (PC) flows into mode-specific origin-destination (OD) flows; this model also looks at the use of terminals for consolidation and distribution (which could include ports and airports), and at empty or loaded return trips for the vehicles (vessels) used.
- An aggregate network assignment, by mode (this can be combined with the assignment of passenger transport vehicles, e.g. for road transport).

The ADA philosophy is that it is most important that the logistics choices (the 'middle part') are modelled at the level of the individual decision-makers (disaggregate level) because decision-makers on all goods flows between two zones per commodity type do not exist; this would assume a greater level of coordination and optimality than what happens in reality. Modelling the PC flows and the network assignment at the aggregate (zone-to-zone) level on the other hand is acceptable for reasons of data availability and computation time. In order to have a disaggregate logistics model, the outcomes of the PC model have to be allocated to individual firms at both ends (firm-to-firm flows), and before network assignment flows have to be aggregated into OD matrices.

3.2 The structure of the ADA model

The boxes in Figure 4 indicate model components. The top level of Figure 4 displays the aggregate models (at the zone-to-zone level). 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 (including transport chain choice), and finally another aggregate model for network assignment.

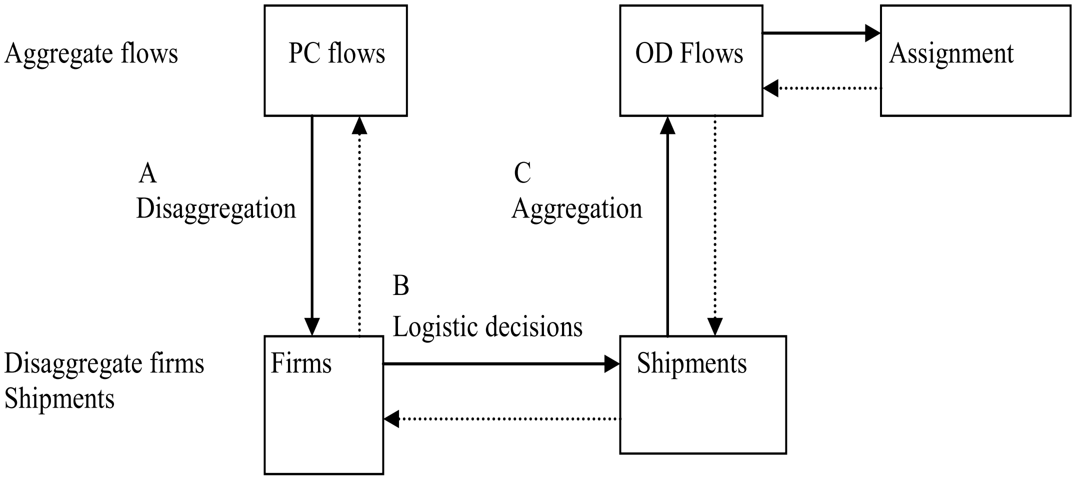

Figure 4: Structure of the ADA (inter)national/regional freight transport model system

The model system starts with the determination of flows of goods between production (P) zones and consumption (C) zones (being retail (R) 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). It is possible to also include flows to and from wholesale (W) here to include a representation of supply chains. This is done in the Norwegian and Swedish national freight transport models; their PC flows are actually PWRC flows.

The aggregate PC model produced flows at the zone-to-zone level by commodity. These flows are then disaggregated to one or more firm-to-firm flows between sending firms in the production zone and receiving firms in the consumption zone (step A Disaggregation in Figure 4). This disaggregation to firm-to-firm flows is necessary as a prerequisite to the disaggregate chain choice model which works at this micro-level. Step A uses available data on the number of firms in each zone, the size of each firm, and the types of commodities produced and consumed.

In the Swedish ADA model, the disaggregation of the PC matrices could partly be based on the Commodity Flow Survey (CFS), with information on millions of individual shipments. Such a data source was not available in Denmark, as was also the case for the ADA-application to Norway (however, Norway later also launched a downsized CFS). For these countries, we allocated PC flows to firm-to-firm relations on the basis of sampling of firms, using available more aggregate statistics on the location, sector and size of firms (in terms of the number of employees) by zone by commodity type. The data used for this are register data. Therefore, the firms used in the model are existing firms, but the flows between firms in the model are of a synthetic nature (in the absence of disaggregate data on shipments).

The information on the number of firms by size by zone is used both at the production and the consumption end. We assume that each firm (local unit) will produce goods in only one commodity class, but it may consume goods from several commodity groups.

For domestic flows and for the domestic side of export and import, the firm-to firm flows model uses the assumption of proportionality with the product of the firm sizes of the sender and the receiver: the higher this product, the higher the probability that this sender-receiver combination will be selected as firm-to-firm flow and the bigger the flow that will be allocated to this pair of firms (if selected). The international flows are split by assuming a number of firms at the foreign end that is proportional to the numbers of senders and receivers for a given zone size at the domestic end.

The firm-to-firm flows admittedly have a partly synthetic character, but the advantages of this setup are greater consistency with private sector logistics choice logic and not having to assume a decisionmaker that optimises the full zone-to-zone level flows of a commodity type. Consequently, the ADA model constitutes a step towards an agent-based model, but is not a full one, since it does not model how shippers and carriers interact with each other.

Step A thus creates disaggregate supplier-receiver relations (a business relation between two firms in which one is the sender of a good and the other the receiver), or firm-to-firm flows. Each relation has an annual flow of goods in tonnes by commodity type. Even for a relatively small country like Denmark, Sweden or Norway, there will be millions of such relations. For the UK there will be many more than in a Scandinavian country. To reduce runtime, firm-to-firm relations can be sampled, and expansion factors can be used to obtain population estimates (as is already used for Sweden and Indonesia).

The core of the ADA model system is a logistic model that, for each of these firm-to-firm flows, determines:

- The choice of annual shipment frequency (and thus shipment size);
- The choice of transport chain type (encompassing mode choice);
- The choice of vehicle type on each leg of the chosen transport chain.

These choices are usually modelled in the ADA models as a deterministic or stochastic minimisation of the logistic costs (step B Logistic decisions in Figure 1). The logistic costs consist mainly of the transport costs, the order costs, and the storage costs. Low shipment frequencies, implying large shipment sizes, will result in low order costs and, due to economy of scale, low transport costs. On the other hand, greater inventories will increase the storage costs. The optimal solution, that is the solution with the lowest total logistic costs, is determined by an evaluation of the logistic costs for a series of shipment frequencies, transport chain types and vehicle types, keeping track of the cheapest solution. The highest shipment frequency that will be evaluated corresponds to the economic order quantity (EOQ).

When the optimal shipment frequencies, chain types and vehicle types for all firm-to-firm relations have been determined by the logistics model, the transport flows are aggregated to origin-destination (OD) matrices by transport mode (step C Aggregation in Figure 4). For transport chains consisting of a single leg, for example a direct OGV2 connection, the origin and destination zones will coincide with the production and consumption zones. This will be different for more complex chain types. For example, a OGV2-rail-OGV2 chain will contribute to the aggregated OGV2 matrix by the first chain leg between the production zone and the first rail terminal, as well as by the last chain leg between the second rail terminal and the consumption zone. The second chain leg in this example will contribute to the aggregated rail matrix.

The ADA model also specifically handles the possibility of empty return trips for vehicles: one of the principles of the ADA model is that the vehicles and vessels should return to where they came from.[4] This leads to two type of vehicle flows in the reverse direction as the original commodity flow: loaded and empty vehicle trips. The relative share of loaded and empty vehicle trips is determined dependent on the mode, vehicle type and distance class.

[4] In reality some trips may not be returning to where they came from and it might be possible to relax the assumption of round tours for some fraction of the transports or vehicles.

Once the aggregated matrices by transport mode have been determined, they can be used for further processing. For each mode they can be assigned to a network (the Assignment in Figure 4). Pivoting and assignment (in vehicles) follows the same procedures as in phase I.

The ADA model has been implemented in practice in a number of countries, where gravity models (Sweden, Denmark), input-output models (Austria) or SCGE models (Norway) are used to obtain the PC matrices (the first A). Deterministic and/or stochastic (random utility model) minimisation of the full logistics costs (including transport, inventory and capital cost) are applied to yield the logistics outcomes at the micro-level (the D) and existing commercial packages are used for network assignment (the second A). The calibration methods for the disaggregate logistics model are discussed in more detail in section 3.3.

3.3 Different methods for calibration/estimation of the disaggregate logistics model (including transport chain choice)

There are three options for modelling transport chain choice in an ADA model:

1. Base this choice, together with determination of terminals used and of shipment size, on deterministic minimisation of the total logistics costs (e.g. Norway);
2. Estimate a logit-type transport chain and shipment size model on data on individual shipments (like the Swedish Commodity Flow Survey); the use of terminals is still determined deterministically;
3. Estimate a disaggregate logit model for transport chain choice on aggregate and partly synthetic data (e.g. Denmark); the terminals used and the shipment size still come from deterministic minimisation of the full logistics costs.

This first approach does not require econometric estimation, since it just chooses the lowest logistics cost solution for every firm-to-firm flow, in a similar way as the shortest path is usually chosen in network assignment. Calibration takes place at a more aggregate level, by superimposing constants on the cost functions so that the model outcomes optimally reflect the observed OD-based transport statistics data on mode and/or vehicle type shares for import, export, domestic transport, key intercity flows etc.

The total logistics costs contain more components than just transport costs (especially inventory costs are included as well). More specifically, the total logistics cost function consists of:

- Order cost;
- Distance-based transport cost (this includes fuel cost);
- Time-based transport cost (this includes driver cost and vehicle depreciation costs)
- Loading at sender; unloading at receiver;
- Transfer at terminal (different for containerised transport or not);
- Deterioration and damage;
- Capital costs in transit (time);
- Inventory and capital cost not in transit, e.g. at the receiver.

The fully deterministic approach 1 can lead to problems. The lack of a smooth response of the model to generalised transport cost changes can result in either sticky or flip-flop behaviour. Secondly, the lack of parameters in the model does not allow a tractable estimation of the model and limits the possibilities for calibration of the model to optimise the fit of the model output to observed flows.

The second alternative entails the estimation of discrete choice models of transport chain and shipment size choice on disaggregate (shipment level) data, as has been done in a number of research papers (e.g. Abate et al. (2019)). The transport chain choice sub-model, in the European model Transtools3 (Jensen et al., 2019) is estimated on two disaggregate data sources: the Swedish Commodity Flow Survey (CFS) and the French ECHO survey. The main disadvantage of this approach is that for the econometric estimation of the logistics model for transport chain and shipment size choice, data is required about goods flows at the shipment level (such as the Commodity Flow Surveys (CFS) in the USA or Sweden, or the ECHO survey in France), which for most countries will not be available. Estimation of the logistics model (as a random utility logit model) within ADA is preferred (should sufficient disaggregate data be available) over a deterministic model, because the latter could lead to less plausible sensitivities for transport time and costs changes (Abate et al., 2019).

The third way of calibrating the coefficients of the logistics model within ADA is estimating the disaggregate model on available aggregate OD-level information. This approach was developed in recent years for the new version of the national freight transport model of Denmark (Willigers et al., 2025).

The general approach in option 3 is the introduction of a smooth choice function of which coefficients are estimated instead of the all-or-nothing approach of option 1. A plausible choice model to use is the standard multinomial logit (MNL) model, which is a smooth function and can be estimated relatively easily. In the new cost function the shipment size is implicit, including alternative and commodity specific constants as well as coefficients for cost and time.

The systematic utility (V, unit utils, at firm level) for moving an annual firm-to-firm flow Q of goods of type g by transport chain m between sending firm i and receiving firm j is as follows:

$$V_{ijgm} = \beta_g^C C_{ijm} + \beta_g^T T_{ijm} + \frac{\beta_{mg}^S}{\sqrt{Q_{ijg}}} + ASC_{mg} \qquad (1)$$

The first two terms are related to transport costs C per tonne and transport time T (in minutes per transport), with corresponding coefficients indicating their importance. The third term shows the shipment size-effect on logistics costs with a mode and commodity specific coefficient, dependent on the total flow between the producing firm and the consuming firm in tonnes per year Q (inspired by Combes & Tavasszy (2016)). The fourth term is a system wide mode-specific constant.

In terms of the logistics trade-off between storage and transport costs via shipment size as decision variable, the model differs from the one in option 1 in that shipment size considerations are now implicit - the various cost factors (ordering, storage, value density) and shipment sizes do not need to be known. Their joint effect is now represented by the coefficient β^S. It is therefore reasonable to allow the shipment size-related coefficients to depend on the type of goods moved, as with the time (cost) coefficient. This construct has the main advantage that no assumptions need to be made that cannot be backed up by observations while it does follow the logistics logic of the current model; the main drawback is that these cost factors are not available anymore as input for to the choice model.

Evaluation of the choice model can follow the MNL probability function; the shares of the modes of transport follow by multiplication with Q at the company level. Addition of flows by mode for all firms with the same (i,j) provides the OD flows in tonnes by mode between regions.

The equation above works at the level of firms (making it a disaggregate model), but no data on mode and vehicle type choice at the firm level are used to estimate it. On the contrary, the estimation of the coefficients of the disaggregate utility function takes place on aggregate data, as described below, except for the partly synthetic disaggregate firm-to-firm PC flow Q. The estimation involves the following calculation loop until convergence to find the optimal values for the betas and ASCs:

1. Evaluation of the above MNL transport chain choice model to derive all OD relations that are implied by each P to C transport chain, at firm level;
2. Aggregation of flows across firm-to-firm flows by mode;
3. Comparison with observed OD flows (by zone pair) per mode to see how close they are to each other;
4. Adjustment of the coefficients and repeat 1) until the difference at step 3) is sufficiently small.

In the project in Denmark, Significance developed an estimation program to carry out this disaggregate logit estimation on observed disaggregate PC and OD data, level of service data, unit transport costs and partly synthetic firm-to-firm flows and used this to find the coefficients, that will be used in the next version of the national model.

3.4 Data requirements and data availability in the UK for the disaggregate logistics model, (including transport chain choice)

The data that are required for building and for estimating/calibrating the disaggregate logistics model (including its transport chain choice) differ between the options 1-3 mentioned at the beginning of section 3.3.

The first option has relatively small data requirements. Level-of-service (LOS) data (transport time and distance) at the zone-to-zone level are needed for each mode or vehicle/vessel type in the model. Then, one needs transport cost functions and mode/vehicle/vessel capacity and average load factor and share of empty trips that read in the LOS-data and yield transport cost of each alternative. Furthermore, one also needs a spatial terminal list (linked to the networks) for the transhipments, and unit costs for different kinds of transhipments. Also, unit costs are needed for inventory cost (inventory in transit and storage in a warehouse) and order cost. The total logistics cost that is optimised can be calculated using this collection of inputs. The model is applied on the firm-to-firm flows and its main outcomes (after aggregation) are calibrated on market shares of the modes in the OD data, by commodity type. All these data are available for the UK.

To carry out the disaggregation step to firm-to-firm volumes (step A in Figure 4), there is a further data requirement in options 1-3:

- The number of firms in each zone;
- The size of each firm (employment in classes);
- The types of commodities produced and consumed by each sector.

For the UK, these data seem to be available (as in other western countries) in the land use data/business data.

The second option additionally requires data at the shipment level (transport chain used, weight, value, commodity, production and consumption location, terminals used) for all modes. This is not available in the UK, since (unlike Sweden) it does not have a commodity flow survey. The Bill of Laden data is in principle disaggregate, but at best this can be obtained for container transports through UK ports, not for non-container transport or domestic container transport and even here unrestricted use may not be possible because the data is commercially sensitive.

For the third option, the commodity flow survey is not needed. The data for option 1 are needed here as well. The model estimation uses predictions of firm-to-firm annual goods flows, which are based on disaggregating zone-to-zones flows using the same data as used for the disaggregation in options 1 and

2 (number of firms in each zone; size of each firm and types of commodities produced and consumed by each sector).

3.5 Which model type is recommended for the FAME transport chain choice model?

The disaggregate transport chain choice model can be a deterministic model, a stochastic model estimated on disaggregate data or a disaggregate model estimated on aggregate PC and OD data and disaggregate (partly synthetic) annual shipment volume data. For the UK, the second option needs to be ruled out, because a commodity flow survey covering all relevant modes is missing (a possible new data collection on multimodal disaggregate transport chains can be found in Annex A).

The first option could be used in the short run, because the data are available and it is relatively simple to build. It does however potentially lead to sticky (overly stable) behaviour when transport times and costs change until a point is reached where the behaviour becomes flip-flop (too jumpy). This is due to the all-or-nothing character of the solution used. The third option provides a smoother behaviour to transport time and cost changes and uses the same data as option 1. This makes it preferable to option 1. It has been applied for the new national freight transport model of Denmark, but this is the only application of this estimation method so far (Willigers et al,. 2025).

3.6 Validation of model outcomes against elasticities from the literature for mode choice

The international elasticity literature review of de Jong et al. (2023) contains 80% intervals for then impact of transport cost (price) changes by mode on tonnes and tonne-kilometres by mode, see Table 2.

Table 2. Own price elasticities for validating a freight transport model (mode effect only)

Exogenous change	Impact on	Lower bound	Most likely value	Upper bound
Road price (per tkm)	Road tkm	-0.14	-0.4	-1.3
Road price (per tkm)	Road tonnes	-0.04	-0.3	-0.5
Road price (per tkm)	Road vkm	-0.14	-0.4	-1.3
Rail price (per tkm)	Rail tkm	-0.4	-0.8	-1.4
Rail price (per tkm)	Rail tonnes	-0.3	-0.6	-1.1
Rail price (per tkm)	Rail vkm	-0.4	-0.8	-1.4
Sea price (per tkm)	Sea tkm	-0.2	-0.6	-1.0
Sea price (per tkm)	Sea tonnes	-0.2	-0.6	-1.0
Sea price (per tkm)	Sea vkm	-0.2	-0.6	-1.0

For instance, the range for the impact of the road transport price on road tkm is between -0.14 and -1.3 with -0.4 as most likely value. This is the mode choice elasticity. There is much more literature on mode choice than on transport chain choice, that is why we look at the mode choice elasticity here.

In joint mode and shipment size selection models, the resulting elasticities are often smaller (in absolute values) than in pure mode of transport selection models, possibly because shippers can also divert to other shipment sizes. This was also found in a recent model in this class, Abate et al. (2019), estimated on microdata for Sweden, although this study only determined elasticities for two specific types of goods.

The above price elasticities refer to changing the full transport cost of some mode. A specific transport cost component (e.g. fuel cost, driver cost, vehicle depreciation cost) will have its impact on mode choice/transport chain choice through the full transport cost. The size of this effect on tonne-kilometres of the mode depends on the share of this cost component in the total transport cost. For establishing these shares, one can use sources on transport cost in the UK, such as the annual cost tables published by Motor Transport, a road freight industry publication. The impact of transport time on the tonne-kilometres of some mode work mainly through the time-dependent transport cost (including driver cost, vehicle depreciation cost), which forms the majority of the full transport cost. Besides the impact of time through cost, time also can have a direct effect on the tonne-kilometres of a mode. This is the cargo-dependent effect of transport time, that occurs when highly valuable or perishable cargo is away in transit. For most goods, this additional effect of time is likely to be fairly small.

4. Summary

The FAME Feasibility Report (MDST, 2024) recommended a zoning system, specific modes, vehicle types and commodities and a distinction between domestic versus international trips that we all support for a new national forecasting system for freight transport for the UK. We have different recommendations on the model structure and specific model components.

First, we propose developing a combination of a regression model for attraction and a single-constrained gravity model for sourcing choice/distribution (together forming a trade model), estimated on production-consumption data. This would give the DfT a submodel that it can run itself to forecast future PC matrices and to calculate the impact on these choices of changes in GDP, population or employment and of changes in transport times and costs. It would add a new dimension of variable demand modelling (sourcing choice/distribution) in freight transport forecasting at the national level in the UK. The trade model would be followed by an aggregate logit model that splits the PC matrix into transport chains. By explicitly including the use of multimodal terminals and consolidation and distribution centres in road transport, the transport chain choice model forms the link between PC flows (that it takes as inputs) and origin-destination (OD) flows (that it produces). The transport chain model is also dependent on transport time and cost and therefore also a form of variable demand. The resulting OD-flows can be assigned to the networks. This set of three submodels covers all four steps of the classic four-stage model. We propose that initially (phase I) this model is developed.

Secondly, in a subsequent phase II, the trade model and the network assignment remain the same, but the aggregate transport chain model is replaced by a transport chain choice model (possibly also including the choice of shipment size) at the level of individual firm-to-firm flows. This has the advantage of trying to reproduce decisions at the level where that actual decisions are made by individual firms, instead of the overall zone-to-zone level. The disaggregate chain choice model starts with a disaggregation of zone-to-zone flows to firm-to firm flows using register data on firms, employment of these firms and the type of goods they produce and consume. This partly synthetic firm-to-firm flow is then used together with transport time and cost data, (aggregate) OD and PC data and terminal location data to estimate the disaggregate transport chain choice model. The OD flows can be taken from this model, after aggregation. The phase II model will have the same capabilities in terms of variable demand modelling as the phase I model, but with a better behavioural foundation and as such it forms a step in the direction of agent-based modelling.

References

Abate, M., I. Vierth, R. Karlsson, G.C. de Jong and Jaap Baak (2019) A disaggregate stochastic freight transport model for Sweden, Transportation, 46, 671–696.

Anderson J.E. and E. van Wincoop (2003) Gravity with gravitas: a solution to the border puzzle, Am Econ Rev 93-3, 171–192.

ASTRA consortium (2000) Final report: assessment of transport strategies; University of Karlsruhe, Germany.

Bates, J. and I. Williams (2026) FAME Base Year Matrices Thinkpiece, Draft final version, Thinkpiece for DfT.

Ben-Akiva, M., G.C. de Jong (2013) The aggregate-disaggregate-aggregate (ADA) freight model system, in Freight Transport Modelling (Eds.: M.E. Ben-Akiva, H. Meersman and E. van de Voorde), Emerald.

Bok, M. de and L. Tavasszy (2018) An empirical agent-based simulation system for urban goods transport (MASS-GT), Procedia Computer Science 130: 8.

Bok, M. de, L Tavasszy and S. Thoen (2020) Application of an empirical multi-agent model for urban goods transport to analyze impacts of zero emission zones in The Netherlands, Transport Policy https://doi.org/10.1016/j.tranpol.2020.07.010.

Bröcker, J. (1998) Operational spatial computable general equilibrium models, The Annals of Regional Science, Vol. 32, 367-387.

Cascetta, E. (1997) National modelling in Italy; Simulation and evaluation models for the Italian DSS, Paper presented at Seminar on National Transport Models: The State of the Art, Noordwijk.

Combes, F., L.A. Tavasszy (2016) Inventory theory, mode choice and network structure in freight transport. EJTIR 16(1), 38-52.

Department for Transport (2025) TAG UNIT M2.1, Variable Demand Modelling, DfT, London.

Dunkerley, F., C. Rohr and A. Daly (2015) Road traffic demand elasticities: A rapid evidence assessment, Report for the UK Department for Transport, RAND RR888, RAND Europe, Cambridge.

Edwards, H. with assistance from J. Bates and H. Swahn (2008) Swedish base matrices report – estimates for 2004, methodology, data and procedures, SIKA PM 2008.

Federal Highway Administration (1999) Guidebook on Statewide Travel Forecasting, prepared for the FHWA by the Center for Urban Transportation Studies, University of Wisconsin-Milwaukee and the Wisconsin Department of Transportation.

Garrido, R.A. (2000) Spatial interaction between trucks flows through the Mexico-Texas border, Transportation Research A, 33(1), 23-33.

Gliebe, J. Doyle, K. Wies, C. Smith, K. Shabani, M. Outwater (2014) An Agent-Based Computational Economics Approach to Forecasting Freight Flows for the Chicago Region, TRB Innovations in Travel Modeling – Baltimore.

Goodwin, P. (2024) The deep flaws in how we model freight, TAPAS.network, 8 May 2024, Commentary.

Grebe, S., G.C. de Jong, C. Wampera, C. Obermayer, R. Pompl (2020) VMÖ – A new strategic transport model for Austria. Special ETC 2019 issue of Transportation Research PROcedia (TR-PRO), 49, 95-106.

Hidayati, L., G.C. de Jong, A. Whiteing (2024) A stochastic logistics model for Indonesia's national freight transport model: Transport chain choice from the shipper perspective, Asian Transport Studies, Vol. 10, 2024.

Ivanova, O., A. Vold and V. Jean-Hansen (2006) PINGO – A model for prediction of regional and interregional freight transport, TØI report 578/2002 (revised), Institute of Transport Economics, Oslo.

Jensen, A.F., M. Thorhauge, G.C. de Jong, J. Rich, T. Dekker, D. Johnson, M. Ojeda Cabral, J. Bates and O. Anker Nielsen (2019) A disaggregate freight transport chain choice model for Europe, Transportation Research E, 121 , 43-62, 2.

Jong, G.C. de, H. Gunn and W. Walker (2004) National and international freight transport models: an overview and ideas for future development, Transport Reviews, 24 (1), 103 - 124.

Jong, G.C. de and M.E. Ben-Akiva (2007) A micro-simulation model of shipment size and transport chain choice, Special issue on freight transport of Transportation Research B, 41, pp. 950-965, 2007.

Jong, G.C. de, A. Schroten, H. van Essen, M. Otten and P. Bucci (2010) The price sensitivity of road freight transport – a review of elasticities, in E. van de Voorde and Th. Vanelslander (eds.), Applied Transport Economics, A Management and Policy Perspective, De Boeck, Antwerp.

Jong, G.C. de, A. Burgess, L. Tavasszy, R. Versteegh, M. de Bok and N. Schmorak (2011) Distribution and modal split models for freight transport in The Netherlands, paper presented at European Transport Conference 2011, Glasgow

Jong, G.C. de, I. Vierth, L. Tavasszy and M. Ben-Akiva (2013) Recent developments in national and international freight transport models within Europe, Transportation, Vol. 40 (2), 347-371.

Jong, G.C. de, M. Kouwenhoven, K. Ruijs, P. van Houwe and D. Borremans (2016) A time-period choice model for road freight transport in Flanders based on stated preference data, Transportation Research E, February 2016, 20-31.

Jong, G.C. de, R. Tanner, J. Rich, M. Thorhauge, O.A. Nielsen and J. Bates (2017) Modelling production-consumption flows of goods in Europe: the trade model within Transtools, Journal of Shipping and Trade, 2:5, 1-23, DOI 10.1186/s41072-017-0023-9.

Jong, G.C. de, M. de Bok and S. Thoen (2021) Seven fat years or seven lean years for freight transport modelling? Developments since 2013, Journal of Transport Economics and Policy, Vol. 55 (2), 1-17.

Jong, G.C. de, L. Eggers, A.J. Daly (2023) Update of knowledge on elasticities and time values for GMM, report for the Danish Road Directorate, Significance, The Hague,

Leitham, S., J. Downing, A. Martino and D. Fiorelli (1999) European transport forecasts for 2020: the Streams model results, Paper presented at 1999 European Transport Conference, Cambridge.

Lindgren, S., I. Vierth, M. Johansson and G.C. de Jong (2020a) Using the Swedish Commodity Flow Survey in freight transport research, Paper presented at IEEE Forum ISTS 2020, online.

Lindgren, H. Petterson, I. Vierth and G.C. de Jong (2020b) Scope for automation in the Swedish Commodity Flow Survey, Paper presented at IEEE Forum ISTS 2020, online.

Martino, A, I. Williams, D. Fiorello, K. Noekel, P. Capros, P. Siskos, G. Zazias, I. Charalampidis, P. Karkatsoulis and W. Schade (2018), TRIMODE – integrated transport model for Europe, Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria.

MDST (2024) DfT FAME: Feasibility Report, Final report for the Department for Transport, MDS Transmodal.

Piecyk, M. and A. McKinnon (2013) SRF Roadmap – Part 1: Road Freight Transport in the UK, Heriot Watt University and the University of Cambridge.

Rijkswaterstaat (2022) Actualisatie basisjaar 2018 BasGoed, Analyserapportage, Rijkswaterstaat, Ministerie van Infrastructuur en Waterstaat, Rijswijk.

Rizet, C. and M. Guilbault (2004) Tracking along the transport chain with the shipper survey, Proceedings of the ITSC Conference, Costa Rica, Elsevier.

Rohr, C., B. Whittaker, J. Fox, J. Glenesk and J.J. Collins (2018) Latest Evidence on Factors Impacting Road Traffic Growth: An Evidence Review, report for the UK Department for Transport, WSP and RAND Europe.

SCENES consortium (2001) SCENES transport forecasting model: calibration and forecast scenario results, deliverable 7 to EU DGTREN; SCENES Consortium, Cambridge.

Samimi, A., A. Mohammadian and K, Kawamura (2010) Freight demand microsimulation in the U.S., World Conference on Transport Research (WCTR), Lisbon.

Significance, NEA and Demis (2010) Schatting BasGoed. Rapport for DVS, Significance, The Hague.

Tapia, R., D. Mulders and G.C. de Jong (2025) Freight Model Stage 2 Estimation report, Report for Sund & Bælt, Significance, The Hague.

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 European Congress of the Regional Science Association, Dortmund.

Tavasszy, L.A. and G.C. de Jong (2014) Modelling Freight Transport, Elsevier Insights Series, Elsevier.

TRT (2018) Description of the TRUST model, TRT Trasporti e Territorio, Milan.

Vejdirektoratet (2021) The Danish National Freight Transport Model, Dok.nr. 12084-001 rev. 2, Vejdirektoratet, Copenhagen.

Williams, I. (2024) Peer Review of Freight Analysis and Modelling Environment (FAME) – Stage 1, Report for DfT, Ian Williams.

Willigers, J., D. Mulders and G.C. de Jong (2025) Report on the second and third round of estimating the logistics model within the GMM, Report for the Danish Road Directorate (VD), Significance. The Hague.

Zondag, B, P. Bucci, P Gützkow en G.C. de Jong (2010) Port competition modeling including maritime, port and hinterland characteristics, Maritime Policy & Management, 37 (3), 179-194.

Annex A. Possibilities for new data collection in the UK

The Phase I and II forecasting models recommended in this thinkpiece do not depend on new data collection: these models can be based on existing data sources without having to carry out new surveys. Nevertheless, in this Annex A we would like to discuss some potential data collection efforts that would lead to data that would be very useful for future freight transport modelling exercises in the UK.

First of all, there would be benefits from doing some kind of commodity flow survey. The key point of such data is that it collects transport-related information at the level of individual shipments not just for one mode, but for all the modes used in the transport chain. The Swedish Varuflödesundersökningen (VFU, CFS), carried out by Statistics Sweden for the transport authorities, asks a sample of Swedish production and wholesale companies (i.e. shippers) to record their shipments in a 1-3 week period. This includes outgoing shipments (domestic and international) and incoming (international). This data collection is done about once every five years. It contains data on 3.5 mln outgoing shipments, including data on production and consumption location (municipality level), industry, weight, value, commodity type and mode chain (e.g. truck-train-truck). The data have been used for the estimation of models of transport chain choice and of joint transport chain and shipment size choice, but also as one of the sources of PC data (Lindgren et al., 2020a).

Collecting shipment data at a large scale (many thousands of shipments, if not millions) would be very welcome for freight modelling in the UK too, but to obtain a sufficiently high response rate from the shipper firms, their participation should be obligatory. This may be at odds with attempts to reduce to administrative burden on firms. This is also an issue in Sweden and attempts to automate the data collection of the CFS, by using data from the logistics systems of the firms, have been made (Lindgren et al., 2020b). For the estimation of a model of transport chain (and shipment size) by commodity type, it is not necessary to have millions of records (shipments). A survey like ECHO (see below) would probably be big enough for this. But if the survey is to be used as key input for a spatially detailed PC matrix, a large commodity flow survey would be required.

The US has a similar CFS in terms of sample size and data items collected. The French Envois-Chargeurs-Opérateurs (ECHO) survey (Rizet and Guilbault, 2004) has a much more limited sample size (10,000 shipments in total). ECHO was carried out by IFSTTAR and ISL in 2004 using as starting point a sample of French shippers and their most recent shipments in the last (up to the three) months. ECHO is richer than the Swedish or US CFS (e.g. it also has transhipment locations), because it not only collects transport chain information from 3,000 shippers, but also information on each leg from the 27,000 carriers, logistics service providers and receivers involved. In the European Transtools3 project, transport chain choice models were estimated jointly on the Swedish CFS and ECHO. Data like a CFS or ECHO could also be used to obtain good information for both domestic and international flows on the value-to-weight ratios per commodity, since these can both be asked per shipment.

Another example of using data from the logistics systems of firms themselves (without having to do a survey) is the collection of XML data by Statistics Netherlands on individual tours in the deliveries of a set of firms active in transport operations in The Netherlands. These data were used to estimate a tour formation (multi-drop delivery tours) and scheduling model in the MASS-GT project (de Bok et al., 2018, 2020) and in a similar tour-based sub-model within the Dutch national freight transport model BasGoed (Rijkswaterstaat, 2022).

significance
quantitative research

Grote Marktstraat 47
2511 BH Den Haag
Nederland

info@significance.nl
+31 70 312 1530