

Freight Demand Forecasting Models

The different approaches

Research Master's Thesis submitted by
Simon Boland

With the view of getting the degree in
Master 120 in Business Engineering, professional focus

Supervisor
Bart Jourquin

Academic Year 2017-2018

Acknowledgements

First of all, I would like to thank my supervisor Bart Jourquin for advising and supporting me throughout the making of this master thesis. I sincerely appreciated his reactivity, his encouragement and his expertise.

I would also like to thank Catriona Hudson for proofreading this paper.

And finally, many thanks to my friends and family for their overall support throughout the entirety of the project.

Table of contents

1	Introduction	1
2	Literature review	3
2.1	Freight demand characteristics	3
2.2	Factors affecting freight movement.....	4
2.3	Freight/Non-Freight Intensive Sectors	5
2.4	Four layers of decisions and activities.....	5
2.5	The Classic Four-Stage Transport Model.....	6
2.5.1	<i>First Stage: Trip Generation</i>	<i>8</i>
2.5.2	<i>Second Stage: Trip Distribution</i>	<i>11</i>
2.5.3	<i>Third Stage: Modal choice.....</i>	<i>16</i>
2.5.4	<i>Limits of the Four-Stage Model.....</i>	<i>18</i>
2.6	Activity-Based Models and Multi-Agent Systems	19
3	Classifying Freight Demand Models.....	22
3.1	Aggregate vs. Disaggregate Data.....	22
3.2	Attributes Influencing Freight Trip Generation	23
3.2.1	<i>Example of attribute: linking the forecasts to employment.....</i>	<i>25</i>
3.3	History of model classifications.....	28
3.4	The most recent classification system: Freight Demand Forecasting in Class.....	29
4	Real-world models.....	32
4.1	PLANET Model.....	32
4.1.1	<i>Description of the model.....</i>	<i>32</i>
4.1.2	<i>Methodology.....</i>	<i>35</i>
4.2	TRANS-TOOLS model.....	38
4.2.1	<i>Level of Service calculation</i>	<i>41</i>
4.2.2	<i>Demand response</i>	<i>42</i>
4.2.3	<i>Route choice and network loading.....</i>	<i>43</i>
4.2.4	<i>ETISplus Database.....</i>	<i>44</i>
4.3	FAME.....	48
4.3.1	<i>Firm-type Generation module.....</i>	<i>49</i>
4.3.2	<i>Supply Chain Replication module.....</i>	<i>50</i>
4.3.3	<i>Shipment Forecasting module.....</i>	<i>50</i>

4.3.4	<i>Logistic Decision module</i>	51
4.3.5	<i>Network analysis module</i>	52
4.4	The Integrated transport system in the Arab Mashreq (ITSAM).....	53
4.4.1	<i>Description of the model</i>	53
4.4.2	<i>Development of the model</i>	53
4.4.3	<i>Assumptions</i>	54
4.5	Comparison of the real-world models	56
5	Conclusions and Managerial Implications	60
6	Bibliography	62
7	Appendices	67

List of acronyms

ADA	Aggregate-Disaggregate-Aggregate
ALOs	Administrative and Logistical Operations
DGSIE	Directorate General Statistics and Economic Information
EDI	Electronic Data Interchange
EOQ	Economy Order Quantity
ESCWA	United Nations Economic and Social Commission for Western Asia
ETIS	European Transport policy Information System
FA	Freight Attraction
FAME	Freight Activity Micro-simulation Estimator
FG	Freight Generation
FIS	Freight Intensive Sectors
FP	Freight Production
FSA	Freight Service Activity
FTA	Freight Trip Attraction
FTG	Freight Trip Generation
FTP	Freight Trip Production
HB	Home-Based
IFSTEM	International Freight Simultaneous Transportation Equilibrium Model
ITSAM	Integrated Transport System in the Arab Mashreq
LoS	Level of Service
LP	Linear Programming
MNL	Multinomial Logit
NAICS	North American Industry Classification
NHB	Non-Home-Based
non-FIS	Non-Freight Intensive Sectors
NUTS	Nomenclature of Territorial Units for Statistics
O-D	Origin-Destination
PC	Production-Consumption
SIC	Standard Industrial Code
STEM	Simultaneous Transportation Equilibrium Model
TRANS-TOOLS	(TOOLS for TRansport Forecasting AND Scenario testing

List of Tables and Figures

TABLE 1: A GENERAL FORM OF A TWO-DIMENSIONAL TRIP MATRIX.....	12
TABLE 2: VARIABLES EXPLAINING FREIGHT MODE CHOICE	17
TABLE 3: SUMMARY OF FINDINGS CONCERNING ATTRIBUTES INFLUENCING FREIGHT TRIP GENERATION	24
TABLE 4: REVIEW FACTORS FOR FREIGHT MODELING TECHNIQUES	25
TABLE 5: OUTLINE OF ESTIMATION PROCESS FOR TYPICAL APPLICATION SCENARIOS.....	28
TABLE 6: CLASS ADVANTAGES AND DISADVANTAGES.....	30
TABLE 7: FREIGHT MODEL CHARACTERISTICS BY INPUT/OUTPUT	31
TABLE 8: INPUT VARIABLES FOR THE EXPLANATORY AND PARSIMONIOUS MODELS.....	52
TABLE 9: ITSAM MODEL VARIABLES	55
TABLE 10: COMPARISON OF THE REAL-WORLD MODELS	56
FIGURE 1: THE CLASSIC FOUR-STAGE TRANSPORT MODEL	7
FIGURE 2: TRIP PRODUCTIONS AND ATTRACTIONS	9
FIGURE 3: FUNCTIONS TO MEASURE FSA	27
FIGURE 4: MODULES OF THE PLANET MODEL.....	33
FIGURE 5: LINKS BETWEEN TRANSPORT MODULES IN PLANET	34
FIGURE 6: MAIN MODEL-GROUPS IN THE FREIGHT AND LOGISTICS MODEL FLOW	40
FIGURE 7: FRAMEWORK OF THE FREIGHT ACTIVITY MICROSIMULATION ESTIMATOR	49

1 Introduction

Freight demand forecasting is of paramount importance to many companies. More than \$13 trillion worth of commodities are sent yearly over the United States' freight transportation system. In order to meet the demand's requirements, appropriate forecasting must be held. It, however, remains rather complex (Doustmohammadi et al., 2016).

Over the past few decades, the demand for freight transportation services has been steadily increasing due to the abundance of international trade. Bottlenecks have highlighted the inefficiency and overall weak points in the system which need improvement to face ever-increasing demand. It concerns public and private instances over all means of freight transport (Holguín-Veras et al., 2016).

Freight demand literature has been less abundant than passenger demand although freight participates heavily in traffic congestion, ambient noise and pollution (Ortúzar & Willumsen, 2011). The modelling of freight transport is more intricate than the modelling of passenger transport. It gathers many more actors and includes specificities depending on the type of commodities being transported.

Emerging countries whose economy relies mainly on exports cannot neglect freight demand optimization. The rail and road compromise requires reflection further than the cost perspective: delay or early arrivals influence storage, manufacturing and downstream sales (Ortúzar & Willumsen, 2011).

Transport problems are the same as in the past (congestion, accidents, pollution, costs, etc.). We have learned a lot from times of weak transportation planning. The 21st century has however brought new problems such as fast urbanization, high demand for public transport, scarcity of resources (capital, data and skilled personnel) ... The number of cars and commercial vehicles in use worldwide has increased by more than 38% from 2006 to 2015 (Statista, 2016). Hence the urge to design and plan operations to counter these problems has increased also (Holguín-Veras et al., 2016).

In order to acknowledge, analyze and improve this situation, different tools were developed to aid decision-making in the transport literature. As in most research topics related to economic sciences, the transport models confront the supply to the demand.

The transport demand is commonly defined as the quantities to transport from each origin to each destination. Therefore, the transport demand is frequently represented in origin-destination matrices. These matrices assign quantities to each pair of origin-destination.

Additional frameworks are required to obtain more information on the whereabouts of these quantities. Hence allocation models assign one, or several paths, to the quantities transported for each pair of origin-destination. This allows to observe the flows on a given network.

The transport supply is defined as the combination of the services transporters offer, the infrastructure available, the relevant laws and policies, the modes available and miscellaneous transporter characteristics.

The goal of this master thesis is to provide an overview of the freight demand forecast models. The first part of this paper will be a literature review of the existing theoretical models to predict freight demand. After a brief explanation of the basic concepts of transport planning, we will examine the classic four-stage transport model as a point of reference and address its limitations. Next, we will provide an introduction to multi-agent systems and review the methods to model freight generation and distribution.

The final part of the paper will focus on the review, analysis and comparison of practical models, namely the PLANET model, the TRANS-TOOLS model (which used the ETISplus project database), the FAME model and the ITSAM.

2 Literature review

2.1 Freight demand characteristics

Freight transportation refers to the five basic modes of transportation: highway, rail, water, air and pipelines (Doustmohammadi et al., 2016).

The key influencers on the movement of goods are:

- The shipper: requires the movement of goods to be delivered to a consignee. He will also choose the forwarder based on reliability, price and quality factors.
- The freight forwarder: organizes the shipping process (which mode, at what time, etc.)
- The carrier: responsible for the transportation (provides vans, lorries, trains)
- The driver: guides the vehicle

There are, of course, more actors, such as insurance companies, the state and officials (Ortúzar & Willumsen, 2011).

An important aspect of the transport demand is that it is derived. The demand for transport is not an end goal, the transport is a way to satisfy one or several needs. We first need to understand when, where and how the demand occurs in order to obtain plausible forecasts (Ortúzar & Willumsen, 2011).

Time and space are capital elements in demand forecasting. A misinterpretation of where the demand stands greatly impacts the supply and demand equilibrium. Plus, deliveries can be highly inefficient at certain hours of the day. The time element of the demand is dynamic; peak hours, flexible working hours and work shifts must be considered. The planning gets even more intricate when trips are gathered in **Tour or Trip Chains** (set of linked trips).

Often times, infrastructure and the vehicles are not owned by the same company. The infrastructure takes a lot of combined effort and planning to be changed. Major transport tracks can take from 5 to 15 years to be modified. Therefore, some of them are intricate to change: one cannot add a third of a train station or half a runway to respond to a slight demand raise (Ortúzar & Willumsen, 2011).

Transportation agencies know very little on the whereabouts of trucks and cargoes; they mostly account for the deliveries. They only have an idea of the network truck traffic and the land use patterns, but they know very little of freight flows. Thus, no agent involved in the freight has enough information to describe the whole network, which causes inefficiency.

On the one hand, theoretically sound models guarantee stable results, consistency, give confidence in the forecasts, help to understand the context and the relevant variables.

On the other hand, practitioners don't want to use theoretically sound models just to satisfy their academic colleagues. Plus, they are often quite complex, they require scarce data, they cannot always offer an appropriate level of detail and they don't often offer a lot of flexibility.

The choice between the two depends on the context, and the level of detail required. Balance should be found between theory and expedience in every decision taken (Ortúzar & Willumsen, 2011).

2.2 Factors affecting freight movement

The factors affecting freight movement are:

- Location factors: the location of all inputs and the place where the commodity is sold or consumed will greatly influence the amount of Freight movement.
- The variety of products needed to assemble the freight: the more specific pieces required, the more intermediate transportation.
- Physical factors: the way commodities must be transported (in cold/warm cargoes, in bulk, packaged, with high security...)
- Operational factors: the size of the firm, its geographical dispersion.
- Geographical factors: where the customers are, how dispersed they are.
- Dynamic factors: seasonalities,
- Pricing factors: as opposed to passenger demand, prices are generally not published because of bargaining power and negotiation (Ortúzar & Willumsen, 2011).

2.3 Freight/Non-Freight Intensive Sectors

There is an overall tendency for humans to pay attention to events of a greater size, whereas repeated events of a smaller size pass unnoticed. For example, distribution centers with large trucks that deliver large quantities are easily detected, but few people would suspect the food and retail sector in a big city generates several times more traffic than most ports and distribution centers (Holguín-Veras et al., 2016).

It is therefore useful to classify the industry sectors into two groups:

1. **Freight Intensive Sectors (FIS)**, for which “consumption and production of cargoes are central to the activity performed by the establishment” (Holguín-Veras et al., 2016, p. 14).
2. **Non-Freight Intensive Sectors (non-FIS)**, for which “production and consumption of cargoes are of secondary importance” (Holguín-Veras et al., 2016, p. 14).

Metropolitan areas are of paramount importance. It’s also important to note that deliveries to urban areas using delivery vans and small trucks typically represent 80% to 90% of total freight traffic. These are short trips with a high frequency to counter the fact that urban areas have little storage room. This mode of freight transportation is also more adequate to increasingly important “reverse logistics” of returns and exchanges (Holguín-Veras et al., 2016).

2.4 Four layers of decisions and activities

Freight movements can be seen as the result of four layers of decisions and activities.

- 1 – Decisions on production: destination, type of product, volume and trade relationships (who produces what, in what quantities and for what intermediate or final consumer)
- 2 – Logistics: decisions on the use and location of inventories and supply chain management
- 3 – Choice of transport modes (including multimodal)
- 4 – Transport route followed, more complex for multimodal

The attention allocated to each layer depends on the scope and geography of the study.

To sum up, the research has been less significant in freight modeling than passenger modeling for the following reasons:

- Many aspects of freight demand make it harder to model than passenger movements.
- Urban congestion is mainly caused by people movement.
- The movement of freight involves more actors: shippers, receivers, transporters, storage facilitators, manufacturers...

The most classical approach to the problem is to adopt the conventional four-stage demand model to the case of commodities (Ortúzar & Willumsen, 2011).

2.5 The Classic Four-Stage Transport Model

This model is built around the concepts of zones and networks. The geographical area or establishment group studied is seen as a zone. The model is represented in Figure 1. It is divided in four stages (or four steps):

- **Trip Generation:** Firstly, data is assessed to determine the number of trips entering and leaving the studied zone.
- **Distribution:** Secondly, the destination of these trips is recorded in an Origin-Destination matrix.
- **Modal Split:** Next, the trips are characterized depending on their transportation mode.
- **Assignment:** Finally, the last step assigns the trips to their corresponding networks: private or public transport.

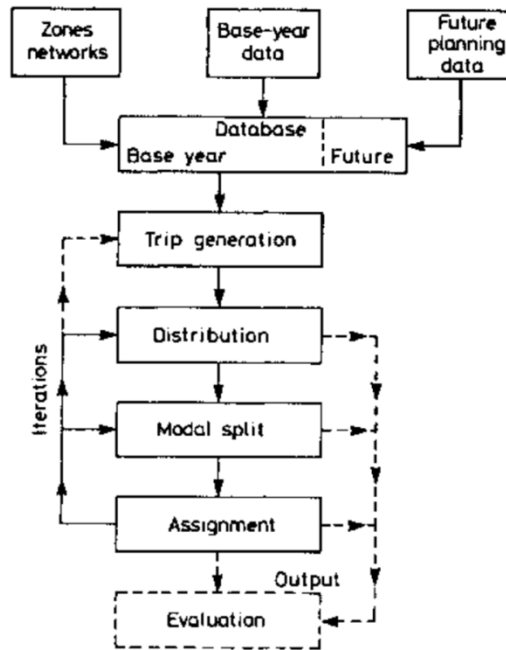


Figure 1: The classic four-stage transport model

This four-stage transport model isn't the only option, but it is known as a solid point of reference.

In the case of interurban movements, stochastic assignment models are often used. However, each vehicle type can have specific modeling drawbacks.

Vans might not be sensitive to the slopes of routes whereas lorries are. Vehicles carrying perishable goods should be prioritized against those carrying sand or other non-perishable goods. The use of multi-class assignment methods may then be warranted to cope with this variety of cost concepts. Lorry drivers often prefer to drive at night to avoid traffic jams, but they are sometimes constrained to drive at rush hours due to short delivery times.

The use of the shortest path algorithm is likely to require incremental path searches. The algorithm should somehow take into consideration different train types (slow local trains, faster direct trains) and contain realistic penalties for delaying operations at transfer locations (Ortúzar & Willumsen, 2011).

2.5.1 First Stage: Trip Generation

The trip generation stage predicts the number of trips generated and attracted to each zone (Mathew & Krishna Rao, 2007).

The Freight and Service Activity (FSA) can be divided into two processes: the process of attraction (the flow that arrives to perform an activity) and the process of production (the flow that leaves the establishment to perform activity elsewhere). Hence the following metrics can be divided into attraction and production:

Freight Generation (FG): the amount of cargo generated by an establishment usually calculated by weight/day (tons/day). It is the “economic activity performed at a business establishment by which input materials are processed and transformed generating an output that, in most cases, is transported elsewhere for further processing, storage, distribution, or consumption” (Holguín-Veras et al., 2016, p. 2). It is the sum of:

- **Freight Attraction (FA):** amount of cargo brought to the establishment to be processed, stored or sold.
- **Freight Production (FP):** amount of cargo sent out. Establishments that sell the final product to the customer have an $FP=0$, if an establishment conducts intermediate processing, $FP>0$.

Freight Trip Generation (FTG): the number of freight generated trips by an establishment. Hence any flow creates two trips: inbound and outbound. FTG is defined as “the result of the logistics decisions concerning how best to transport the FG in terms of shipment size, frequency of deliveries, and vehicle/mode used” (Holguín-Veras et al., 2016, p. 2). It is the sum of:

- **Freight Trip Attraction (FTA):** number of trips arriving to the establishment.
- **Freight Trip Production (FTP):** number of trips departing from the establishment. Establishments that don't have FP will generate no FTP.

FTG is not necessarily proportional to the size of the company. The shipper can change the shipment size without affecting the number of trips.

The models using FG/FTG work best if the internal economic occupation is as homogeneous as possible. Industrial classification systems, such as the Standard Industrial Code (SIC) can be used to make classifications as homogeneous as possible.

Holguín-Veras et al. (2016) also stressed that FTG and FG should be seen as separate. “The amount and nature of the incoming and outgoing Freight Generation depends on the type of business, and its size. In contrast, the FTG depends on the corresponding shipment sizes and the ability of the carriers to consolidate their shipments” (Holguín-Veras et al., 2016, p. 13).

The economic nature¹ of the models reveals the link between FG and FTG: FG is the output of an economic process, FTG is the output of logistic decisions on the size and frequency of the shipments. But there is no one-to-one relationship between both: increasing shipment size could lower the number of trips without affecting the number of commodities (Holguín-Veras et al., 2016).

2.5.1.1 Home-Based and Non-Home-Based Trips

Trips are also often classified by their source and purpose. In home-based (HB) trips, the home of the trip initiator is either the origin or the destination of the trip. Non-home-based (NHB) trips are journeys that don’t involve the home of the initiator. NHB amount to only 15 to 20% of all trips.

“**Trip Production** is the home end of an HB trip or the origin of an NHB trip. **Trip Attraction** is the non-home end of an HB trip or the destination of an NHB trip.” This is explained in Figure 2 (Ortúzar & Willumsen, 2011, p. 140).

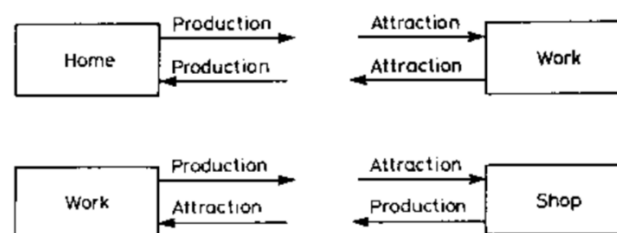


Figure 2: Trip productions and attractions

(Ortúzar & Willumsen, 2011)

¹ Reminder: the amount of cargo produced and consumed are bound to economic transactions

Freight trip productions and attractions only amount to 20% of all the trips, yet they can be of significant importance to traffic congestion.

Important variables include:

- number of employees,
- number of sales,
- roofed area of the firm,
- total area of the firm.

2.5.1.2 Production-Consumption Link and the Importance of Service

The concept of a Production-Consumption (PC) link is the transaction that binds a producer to its customer. There can be simple supply chains with few PC links, or very complex ones with numerous of PC links corresponding to the stages of the production process. PC links have great importance in the understanding of freight demand. Transportation flows materialize into freight traffic. Hence, the analysis of Freight Trip Generation must consider:

1. Production sites and shippers
2. Intermediate processing points
3. Intermediate and end consumer sites

Service should not be overlooked: “As an example (...) the 10,000 restaurants and bars in Manhattan, New York, produce more truck traffic than do the Port Authority of New York and New Jersey terminals combined.” (Holguín-Veras et al., 2016, p. 10)

Service trips are generated by technicians, service providers, who visit an establishment to perform various services. Hence, services greatly contribute to congestion. This should ideally be considered in the making of a freight demand forecast model.

2.5.1.3 Trip Generation Modeling

Trip generation models are often developed using regression analysis. Generally, the functional form will be a multiple linear regression model whose equation resembles the following:

$$y = a_0 + (a_1 x_1) + (a_2 x_2) + (a_3 x_3) + \dots + (a_n x_n) + e$$

Equation 1

In case of a one-variable model, the equation would be:

$$y = a_0 + (a_1 x_1) + e$$

Equation 2

Where y = dependent variable

x_i = independent variable (i = 1, 2, 3...n)

a₀ = constant term

a_i = coefficient of independent variables (i = 1, 2, 3...n)

e = error term

n = number of independent variables (Modi et al., 2011)

The category analysis (or cross classification analysis) is a frequent alternative to regression methods. It is commonly used to determine the number of trips generated using criteria such as household groups, the range of income and car ownership. It is, however, mostly relevant for passenger trip generation. Therefore, it will not be detailed in this paper. The growth factor model (developed further, in section 2.5.2.1) can also be applied to trip generation. (Sarkar et al., 2017).

2.5.2 Second Stage: Trip Distribution

In order to obtain a better representation of trips, they can be represented as “trip matrix” or “trip tables”. This consists in storing trips depending on their origin and destination within a period. Another way to represent trip patterns is by assessing the factors that generate and attract trips.

These matrices are two-dimensional array of cells (Table 1), rows represent origins, columns are the destinations. The diagonal represents intro-zonal trips.

Origins	Destinations					$\sum_i T_{ij}$
	1	2	3	...j	...z	
1	T_{11}	T_{12}	T_{13}	... T_{1j}	... T_{1z}	O_1
2	T_{21}	T_{22}	T_{23}	... T_{2j}	... T_{2z}	O_2
3	T_{31}	T_{32}	T_{33}	... T_{3j}	... T_{3z}	O_3
...						
i	T_{i1}	T_{i2}	T_{i3}	... T_{ij}	... T_{iz}	O_i
...						
Z	T_{z1}	T_{z2}	T_{z3}	... T_{zj}	... T_{zz}	O_z
$\sum_i T_{ij}$	D_1	D_2	D_3	... D_j	... D_z	$\sum_{ij} T_{ij} = T$

Table 1: A general form of a two-dimensional trip matrix

T_{ij} is the number of trips between origin i and destination j.

Origin-Destination (O-D) matrices can be further disaggregated, for instance, by transportation modes and by commodity type.

The sum of the trips in a row/column should be equal to the number of trips produced from/attracted to that zone (Ortúzar & Willumsen, 2011).

$$\sum_j T_{ij} = O_i$$

$$\sum_i T_{ij} = O_j$$

2.5.2.1 Trip Distribution Modeling

Linear Programming

A first approach to freight distribution modelling is **Linear Programming** (LP), which tries to minimize haulage costs subject to supply and demand constraints.

$$\text{Minimize } Z = \sum_{ij} T_{ij} C_{ij}$$

Equation 3

Subject to

$$\sum_i T_{ij} = D_j$$

$$\sum_j T_{ij} = O_i$$

Where T_{ij} are tons of the product moved from i to j ;

C_{ij} are generalized transport costs per ton of products between zones i and j ;

O_i^k, D_j^k are supply and demand at zone i (or j) (Ortúzar & Willumsen, 2011, p.467).

The LP approach gives more realistic results when:

- the industry contains few firms;
- it deals with low-value goods and high transportation costs;
- there are few demand points (buyers)

LP gives satisfying results of the behavior of a single client or industrial firm, but it won't truthfully represent aggregate behavior for various commodities (Ortúzar & Willumsen, 2011).

Growth factor models

Growth factor models can be used to simply determine a future number of journeys. They are based on the assumption that future transport pattern can be forecasted simply by taking the current patterns and multiplying them by a growth factor.

The **Uniform Growth Factor model** is a crude and very simplistic method. It should only be used if the only information available is a general fluctuation rate on studied zones.

Its basic equation is:

$$T_{ij} = F * t_{ij}$$

Equation 4

Where T_{ij} = future number of trips from zone i to zone j;

t_{ij} = current number of trips from zone i to zone j;

F = The uniform growth factor.

This model can be useful for short-term planning but the main drawback is that the same growth factor will be used on every zone. Slightly improved versions of this model are the **Average Growth Factor model** and the **Fratar model**. They take into account multiple growth factors, but can still cause major inaccuracies (Modi et al., 2011).

The **Doubly Constrained Growth Factor model** performs better than the ones quoted above but requires more data. It takes two sets of growth factors for each zone studied: one for the trips entering the zone, and one for the trips leaving the zone (Ortúzar & Willumsen, 2011).

The Gravity Model

The gravity model is linked to Newton's gravitational law. It reveals macroscopic relationships between zones. "The gravity model assumes that the trips produced at an origin and attracted to a destination are directly proportional to the total trip productions at the origin and the total attractions at the destination." (University of Idaho, n.d., p.1)

The model considers the likeliness of a trip to be made according to its duration, distance (University of Idaho, n.d., p.1) and supply and demand constraints (Ortúzar & Willumsen, 2011). This likeliness lies in the calibration parameters β^k (also known as the friction factor). The calibration of the model is a prerequisite step to use the model for demand forecasts.

The **Gravity Model**'s functional form applied to freight forecast is:

$$T_{ij}^k = A_i^k B_j^k O_i^k D_j^k \exp(-\beta^k C_{ij}^k)$$

Equation 5

Where k is a commodity type index,

T_{ij}^k are tons of product k moved from i to j;

A^k, B^k are balancing factors with their usual interpretation;

O_i^k, D_j^k are supply and demand for product k at zone i (or j);

β^k are calibration parameters, one per product k;

C_{ij}^k are generalized transport costs per ton of product k between zones i and j.

(Ortúzar & Willumsen, 2011, p.466).

A generalized cost function can be drawn out from that: (omitting the suffix k for simplicity)

$$C_{ij} = f_{ij} + b_1 s_{ij} + b_2 \sigma s_{ij} + b_3 w_{ij} + b_4 p_{ij}$$

Equation 6

where f_{ij} is the out-of-pocket charge for using a service from i to j;

s_{ij} is a door-to-door travel time between i and j;

σs_{ij} is the variability of travel time s;

w_{ij} is the waiting time or delay from requests for service to actual delivery – it may be a long time for maritime transport for example,

p_{ij} is the probability of loss or damage to goods in transit. (Ortúzar & Willumsen, 2011, p. 467)

All these factors depend on the mode used and the commodity transported. The constants b_n are at least proportional to the value of the goods. For instance, in case of loss, the cost will likely be more than the good's value, due to penalties for delays in delivery.

In general terms, it is important to consider the contribution of transport costs to the final cost of a commodity. For example, in the case of wheat, coal, cement and bricks, transport costs are a main element in their final price; however, in the case of convenience foods, consumer goods, chocolates or electronics, transport costs have a low (direct) contribution to price (Ortúzar & Willumsen, 2011).

The gravity model is the most flexible and exhaustive model presented in this section.

The relationship between LP and gravity models has been explored by Evans (1973). When β approaches 0, (in equation 5) transport costs will tend to negligible; while a very large value for β generates a solution closer to that of an LP model. Hence, the gravity model can engulf the LP model and give accurate results for a wide range of commodities, regardless of the importance of the costs in transportation.

2.5.3 Third Stage: Modal choice

The aim of this stage is to determine the separation between transport modes. It is considered one of the most important stages of the four-stage model. It enables the development of policies to regulate the traffic and road congestion. Mode choice greatly affects efficiency.

The freight mode choice is made by the shipper who decides which mode or combination of modes should be used to send its goods. Each mode has a set of benefits compared to the others, the key is to compare their costs to the efficiency researched (Ortúzar & Willumsen, 2011).

Inventory management can have a great impact on the modal choice. Low inventories imply relying strongly on transportation. The mode choice has to address many concerns to deliver reliable, flexible, time-effective and cost-effective services (CUTR, n.d.)

A very common mode choice modelling method is the Multinomial Logit (MNL) formulation (based on generalized costs). The main drawback is that the formulation is built upon the generalized costs solely, which likely creates approximate results (Ortúzar & Willumsen, 2011).

The following table gathers the variables explaining mode choice for freight demand forecasting. They were separated in four categories, as suggested by the Center for Urban Transportation Research at the University of South Florida (CUTR, n.d.). The factors were extracted from the papers of Cook et al. (1999), Cullinane & Toy (2000) and Jiang et al. (1999).

Total Logistics Costs	Order and handling costs
	Transportation charges
	Loss and damage costs
	Capital carrying cost in transit
	Inventory carrying cost at destination
	Unavailability of equipment costs
	Service reliability costs
	Intangible service costs e.g. Billing processes
Physical Attributes of Goods	Shipment size
	Package characteristics
	Shipment shelf life
	Shipment value
	Shipment density
Flow and Spatial Distribution of Shipments	Shipment frequency
	Distance of Shipment
Modal Characteristics	Capacity
	Trip time and reliability
	Equipment availability
	Customer Service
	Handling Quality – Damage Loss Reputation

Table 2: Variables explaining freight mode choice

(CUTR, n.d.)

According to the Center for Urban Transportation Research at the University of South Florida (n.d.), the factors in Table 2 hold the highest potential for influence through policy intervention.

However, their influence might vary upon different criteria. As an example, the type of commodity can be one of the main concerns as it can influence the importance of other factors. “Price is the most important factor for coal, while transit time is the most important factor for food grains” (CUTR, n.d., p.15).

Cook et al. (1999) also stated that policy interventions can shift the balance among these factors. For sake of clarity, Cullinane & Toy (2000)'s analysis ranked the five most important factors as the following:

1. Cost/Price/Rate, Speed
2. Transit time
3. Reliability
4. Characteristics of the goods
5. Service (unspecified).

A table was developed by the Center for Urban Transportation Research at the University of South Florida to identify at which point in time the factors from Table 2 come into play. This table is covered in Appendix 3.

2.5.4 Limits of the Four-Stage Model

The four-stage model was initially developed for passenger transport. Several authors (such as Shankar, Pendyala and D'Este) acknowledged that the structure can be applied to freight transport as well. However, the models within of these four stages can be much more diverse than those for passenger transport. Indeed, there are many more decision makers (shippers, carriers, intermediates...) and the diversity of the type items represents different and more numerous requirements than for passenger transport (de Jong et al, 2004).

Aside from these four stages, transformation modules are often required (to convert physical flows into tons, and tons into vehicle units, for instance). They are of paramount importance because the converted variables may impact the final predictions greatly. Transformation modules are influenced by decisions on shipment size and frequency, returns and vehicle utilization rates. These decisions could be modelled explicitly in additional logistic modules (for instance, in the FAME model, analyzed further in this paper) or the conversion rates could be fixed in the assumptions of the model.

Inaccuracies could also be engendered for the following reasons:

- The zones' characteristics and compositions are assumed constant throughout the forecast horizon. The development of the transport modes isn't considered either. For instance, the introduction of new trains on the network, which would slightly increase delays and congestion, would have no influence on the frequency of trips (de Jong et al, 2004).
- For trip distribution, the average travel time is assumed to remain the same throughout any day and in the future. There is no consideration of peak/off-peak hours in the original four-stage model. The utility function for the modal choice is also considered constant no matter what time of the day/day of the week it is.
- Finally, the model always represents the average behavior of a group of actors. Analyzing the network in zones only does not consider the shipper/traveler's value of time and modal preferences. The model cannot forecast the behavior of an individual actor accurately (Mladenovic, 2014).

2.6 Activity-Based Models and Multi-Agent Systems

Most of the demand generation models, as shown by Ortúzar & Willumsen (2011), have the drawback of being *trip-based* (which means they assume the transport demand as a sum of independent and isolated trips) and have no regard to the time at which the trips are made. Both criteria are of paramount importance: trips are interconnected in the supply chain and sending truck off-peak hours can drastically improve the journey performance.

Consequently, *activity-based* models should be emphasized. They consider time, space, the different actors involved in the transport and the type of activity. They emphasize the generation of activity/logistic chains. These models were developed much later than the trip based models (Hägerstrand, 1970; Chapin, 1974; Fried, et al., 1977; Jones, et al., 1983) (Piotte, 2013).

In his report, Ferber (1995) explains the intricacies of the methodology of the multi-agent systems.

The *multi-agent system* paradigm was designed to simulate the behavior and interactions between different actors of a system according to their needs and objectives. A multi-agent system is a computerized system in which actors are represented as autonomous virtual entities called *agents*. This concept can approach *activity-based* problems and reflect the demand more realistically than simple *trip-based* models. The initiator of multi-agent systems lets agents meticulously modeled interact to further analyze the macroscopic results.

According to Ferber (1995), physical or virtual entities:

- are able to act in an environment,
- are able to communicate directly with other agents,
- are powered by individual objectives or a satisfaction/survival function that they try to optimize,
- possess their own resources,
- are able to perceive their environment (to a limited extent),
- dispose of a limited representation of the environment,
- possess skills and provide services,
- can potentially reproduce,
- behave to satisfy its objectives, taking into account resources and skills available, and according to their perception, representations and interactions.

He defines a multi-agent system as a system composed of:

- an environment, a space that generally hosts a metric,
- a set of objects, whose position is recorded. These objects are passive, which means they can be perceived, created, destroyed and modified by agents,
- a set of agents,
- a set of relationships, joining objects and agents altogether,
- a set of operations allowing agents to perceive, create, destroy and modify objects,
- operators that represent the application of these operations and the reaction to the world.

Multi-agent systems are used in various sectors such as urban design (Valbuena et al., 2010), ecology (Saqalli et al., 2010), social sciences (Pavon et al., 2008) and studies of epidemic propagation. Davidson, et al. (2005) conducted a specific inquiry on the use of multi-agent systems in the transport field. The paper gathers and criticizes the results of existing methods and models (Piotte, 2013).

One of the main difficulties of the multi-agent systems is the validation of the simulation. Louie & Carley (2008) suggested solutions to validate such models. They also emphasize the fact that multi-agent models require a significant amount of disaggregate data. Gathering every piece of data necessary is often impossible. Hence, Louie & Carley (2008)'s paper tries to advise on data representativeness and exhaustiveness. It also explains the concept of macroscopic validation which consists in analyzing the coherence of the whole multi-agent system and comparing the system results with data drawn from the real world.

Multi-agent systems are able to simulate freight transport generation and incorporate time series. Piotte's doctoral thesis (2013) covers this intricate topic and applies it to the Walloon region in Belgium.

3 Classifying Freight Demand Models

Different models are used to effectively reflect the demand in the short run and in the long run. As of 2016, the freight demand forecasting models are not as developed and accurate compared to passenger transportation forecasting. One of the main issues is that more parties are involved in freight transportation. The abundance of variables also explains the complexity to obtain accurate and generalizable results (Doustmohammadi et al., 2016).

The variables of a chosen model should correctly measure the intensity of freight and service activity. In other words, the variables should somehow be correlated to the economic growth (Holguín-Veras et al., 2016).

3.1 Aggregate vs. Disaggregate Data

Aggregate data is sometimes the only type of data available. It is only relevant to use aggregate data when its components are assumed endogenous. Aggregate freight demand models represent a vast majority of the freight demand models. However, aggregate models can be inaccurate, costly and inflexible.

Disaggregate approaches can be more accurate but require higher statistical and econometric skills.

[Disaggregate] models are very efficient because they (1) require smaller samples than their aggregate counterparts (2) establish a direct connection between the attributes of the establishment and the measure of FSA [Freight Service Activity] that is being estimated, and (3) could be aggregated to any level of geography (Holguín-Veras et al., 2016, p. 20).

However, this type of approach has found limited application. The main reasons for this are:

- the limited understanding of all the elements involved in developing these utility functions and
- the very demanding data-collection efforts required to estimate this type of model.

Data collection for disaggregate approaches suffers from confidentiality and reliability problems.

Yet there is no definitive approach, aggregate and disaggregate models both have their pros and cons. Even collecting data for aggregate modelling represents a much greater effort than that for passenger movements: great dispersion of firms, important daily and seasonal variations, and so on (Ortúzar & Willumsen, 2011).

3.2 Attributes Influencing Freight Trip Generation

An intuitive factor that influences the amount of freight generated is the size of the business (the larger the business, the larger the freight generation). Nonetheless, freight trip generation relies upon many more factors such as inventories, logistic costs and the number of establishments. Table 3, provided by the National Academies of Sciences, Engineering, and Medicine (2012), lists factors that influence freight trip generation.

These factors will be used to analyze and compare the practical models in section 4.5 of this paper. The National Academies of Sciences, Engineering, and Medicine also provided a list of important factors to take under consideration to analyze or develop freight modeling techniques. These are listed in Table 4.

Stratification	Factor	Statistically		Source
		Tested	Significant	
Land Use	Employment	Yes	Yes	Brogan, 1980; Lawson et al. 2012; ITE 2008
	Building area	Yes	Yes	Tadi and Baldach, 1994; ITE 2008
	Establishments	No		Iding et al., 2002; Lawson et al. 2012
Trip Purpose	Employment	Yes	Yes	Brogan, 1980
Vehicle Type	Employment	Yes	Yes	Zavattero and Weseman, 1981; Brogan, 1980
	Land use	Yes	Yes	Zavattero and Weseman, 1981
	Vehicle Type	No		Middleton, 1986
Industry Sector	Employment	Yes	Yes	Iding et al., 2002; Bartlett and Newton, 1982
	Business area	Yes	Yes	Iding et al., 2002; Bartlett and Newton, 1982
	Establishments	No		Iding et al., 2002
	Gross floor area	Yes	Yes	Bartlett and Newton, 1982
	Non-office employ.	Yes	Yes	Bartlett and Newton, 1982
Employment	Commodity type	Yes	Yes	Bastida and Holguín-Veras, 2009
	Type of Business	Yes	Yes	Bastida and Holguín-Veras, 2009
	Industry Segment	Yes	Yes	Bastida and Holguín-Veras, 2009; Holguín-Veras et al., 2011
Number of Truck Drivers	Industry Segment	Yes	Yes	Bastida and Holguín-Veras, 2009
	Type of Business	Yes	Yes	Bastida and Holguín-Veras, 2009
Fleet size	Company SIC	Yes	Yes	Bastida and Holguín-Veras, 2009
Ports	Area	Yes	Yes	Holguín-Veras et al., 2002
	TEUS	Yes	Yes	Holguín-Veras et al., 2002
	Boxes	Yes	Yes	Holguín-Veras et al., 2002
	Daily total vessels	Yes	No	Al-Deek et al., 2000
	Gross tons	Yes	No	Al-Deek et al., 2000
	Gantry crane activity	Yes	Yes	Al-Deek et al., 2000
	Day of the week	Yes	No	Al-Deek et al., 2000
	Imported freight units	Yes	Yes	Al-Deek et al., 2000; Al-Deek et al., 2001
	Exported freight units	Yes	Yes	Al-Deek et al., 2000; Al-Deek et al., 2001
	Exported commodity type	Yes	Yes	Al-Deek et al., 2001
	Imported commodity type	Yes	Yes	Al-Deek et al., 2001
	Exported com. tonnage	Yes	Yes	Al-Deek et al., 2001
	Imported com. tonnage	Yes	Yes	Al-Deek et al., 2001
Carrier vs. Receiver, Attraction vs. Production	Employment	Yes	Yes	Bastida and Holguín-Veras, 2009; Holguín-Veras et al., 2011
	Commodity type	Yes	Yes	Bastida and Holguín-Veras, 2009
	Sales	Yes	Yes	Bastida and Holguín-Veras, 2009
	Industry Segment	Yes	Yes	Bastida and Holguín-Veras, 2009; Holguín-Veras et al., 2011

Table 3: Summary of findings concerning attributes influencing freight trip generation

(National Academies of Sciences, Engineering, and Medicine, 2012)

Dependent variable	Vehicle-trip
	Cargo weight
	Value
Independent variable	Employment
	Building area (square footage)
	Land Use
	Other economic data (sales, establishments, industry segments, type of business)
	Vehicle type
	Commodity type
	Other
Level of aggregation	Disaggregate
	Aggregate
Level of geography	Metropolitan
	Statewide
	National
	Corridor
	Special Facility (e.g. Ports)

Table 4: Review factors for freight modeling techniques

(National Academies of Sciences, Engineering, and Medicine, 2012)

3.2.1 Example of attribute: linking the forecasts to employment

This section aims to show the way to use an attribute of Table 3 to forecast freight demand.

3.2.1.1 Why the employment

The National Academies of Sciences, Engineering, and Medicine (2012) defined two sets of models:

- the first set uses a traditional independent set of variables such as employment,
- the second set uses spatial variables such as proximity to similar businesses.

Although the use of spatial proxies leads to better results, the data remains quite out of reach for most analysts. With employment data being publicly available, it is often used as the only independent variable. Employment statistics are periodically released. Furthermore, it gives a fair representation of the economy (National Academies of Sciences, Engineering, and Medicine, 2012).

3.2.1.2 Econometric Models

Holguín-Veras, J. et al. (2016) used rather simple econometric models to forecast freight demand with different variables. They used either linear or non-linear models. The linear models are variants of the equation:

$$f_i = \alpha + \beta E_i$$

Equation 7

Where f_i = Freight Service Activity (FSA) metric for establishment

E_i = Employment at establishment i

Three linear model types were used:

- **Constant:** Used when the FSA metric does not depend on the establishment employment.

$$f_i = \alpha$$

Equation 8

- **Employment Rate:** FSA metric correlates to establishment employment.

$$f_i = \beta E_i$$

Equation 9

- **Constant and Employment Rate:** FSA metric correlates to establishment employment and a constant.

$$f_i = \alpha + \beta E_i \quad (\text{Equation 7})$$

As for non-linear functions, they rely on the power function because it is consistent with the Economy Order Quantity (EOQ)² model and it is able to accommodate various data patterns.

$$f_i = \phi E_i^\gamma$$

Equation 10

Most of the time, the power function model is the most accurate, but no model can be stated as the best one or the only model to use. The following graph (Figure 3) displays all the basic function developed above (Holguín-Veras et al., 2016):

² “The economic order quantity (EOQ) is the order quantity that minimizes total holding and ordering costs for the year” (NC State University, 2011, p.1)

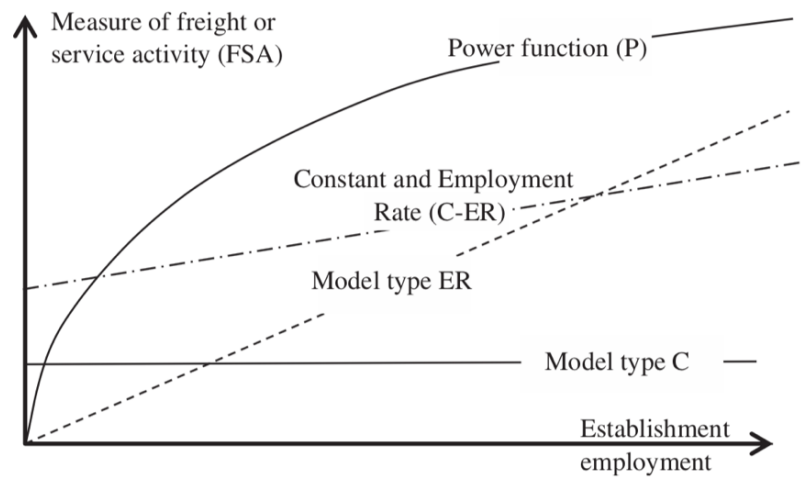


Figure 3: Functions to measure FSA

(Holguín-Veras et al., 2016)

3.2.1.3 Application of the models, using scenarios

It is important to first notice if the goal is to forecast the activity of one or multiple establishments. Next, the type of available data holds an important part in the choice of models. The scenarios shown in the next table (Table 5) derived from these prerequisite decisions. The scenarios are explained further in Appendix 1.

Number of Establishments	Type of Data Available	
	Establishment-Level Data Are Available	Only Aggregate Data Are Available
Single	<u>Scenario 1: Single Establishment</u> - Decide on the desired metric(s) of FSA. - Select the model(s) of interest for the establishment's industry sector. - Run the model(s) with the establishment data. <u>Applicability:</u> Linear and non-linear models	(This scenario is the same as Scenario 1.)
Multiple	<u>Scenario 2: Complete Enumeration</u> - Decide on the desired metric(s) of FSA. - Run model(s) with data for each establishment. - Aggregate results. <u>Applicability:</u> Linear and non-linear models <u>Scenario 3: Sample Estimation</u> - Decide on the desired metric(s) of FSA. - Run the model(s) with the data for each establishment in the sample. - Compute mean values of FSA. - Expand results to entire population. <u>Applicability:</u> Linear and non-linear models	<u>Scenario 4: Only Aggregate Data</u> - Decide on the desired metric(s) of FSA. For each industry sector: - Select the correct aggregation procedure. - Apply the aggregation procedure to obtain aggregate results. <u>Applicability:</u> Only linear models

Table 5: Outline of estimation process for typical application scenarios

(Holguín-Veras et al., 2016).

3.3 History of model classifications

Several classifications of freight demand forecast models have been made over the past 35 years. The major part of the literature focuses on modal choice or differentiating aggregate and disaggregate models. Winston (1985) opposed aggregate and disaggregate models based on the nature of the data used.

Oum (1989) furthered this opposition stating that the model choice depended on the study purpose and the availability of data. He also came to the conclusion that this opposition wasn't a good classification for freight demand forecast models, because long-term freight demand forecast models are mostly aggregate models.

Some papers suggest other considerations than modal choice: Harker (1987) distinguishes econometric models and network balance models. The econometric models he referred to are in fact the aggregate and disaggregate models that have the end goal of modal choice. The network balance models combine route choice and modal choice.

Next, Regan and Garrido (2002) distinguished models according to their geographical spread (international transport, inter-urban transport, transport within a city...).

3.4 The most recent classification system: Freight Demand Forecasting in Class

In 2008, the National Cooperative Highway Research Program (NCHRP, 2008) listed 5 traditional approaches applied in the United States, which are classified from class A to E. Here is a brief explanation of each class:

- Class A, the *Direct facility flow factoring method*, estimates growth in truck transport and applies it to truck traffic volumes. It, however, implies many assumptions and is infrequently applicable.
- Class B, the *O-D factoring method*, is the same as Class A, except the growth is applied in the Origin-Destination table of the truck transport. Yet, data acquisition availability for this class represents an important problem for the implementation.
- Class C, the *Truck model*, aggregates truck trips per zone, distinguishing trips between or into internal and external geographical zones. The main drawback is that the analysis is limited to the freight truck mode, these models are unable to analyze shifts between modes.
- Class D, the *Four-stage commodity model*, is a commodity-based version of class C. It uses forecasts of commodities and applies them to the daily trips. It supports multiple transport modes. It allows to conduct impact analyses, but they do not explain the interaction between industries.
- Class E, the *Economic activity model*, gathers relationships between two relevant models in order to develop a more robust model (for instance, by fusing an economic/land model with a freight demand forecast model). Nonetheless, they do not take into account the performance of transportation facilities (Doustmohammadi et al., 2016).

However, “No model in class A to E alone is capable of addressing all freight policy and analytical needs. Also, the models pay little attention to truck types, trip patterns, trip lengths, or comparisons with truck counts (which are very difficult to obtain).” (Doustmohammadi et al., 2016, p. 20).

Two classifications should be added to these five traditional ones, according to Doustmohammadi et al. (2016). Classes F and G go further and refer to advanced freight demand forecasting models facing current or up-coming problems.

- Class F, the *Logistics Chain Models*, emerged from the need for sensitivity in the economics of commodities.
- Class G, the *Vehicle Touring Models*, tries to truthfully represent trucks motion for impact assessment.

These two models have the drawbacks that the data isn't readily available and that the models are quite costly and hard to put into practice.

Table 6 shows the advantages and disadvantages of using methods of each class.

Class	Model	Disadvantages	Advantages
A	Direct facility factoring	No supply/demand, Not network based	Multi-variable; Corridor and mode specific
B	O-D factoring	Local & state data are proprietary or estimated; Not directly integrated with economic census	Available national data; Adaptable to state & local scales; Multimodal commodity flows; Relatively low cost
C	Three-step truck model	High data collection or purchase cost; Long development time	Predictive model; Multimodal commodity flows; Detailed level of analysis
D	Four-step method	Same as model class C	Same as model class C
E	Economic activity model	Linear relationships between economic activity & freight flow	Economic & land use data; forecasts integrated with the three- or four-step methods; Uses data from local, state & national sources
F	Logistic chain model	Significant data requirements from different sources which are difficult to acquire	Improves the sensitivity to economics of commodities for policymaking; Involves details on the movements of raw goods and finished products
G	Tour-based model	Significant data requirements such as truck load data, intermodal facilities, firm shipment sizes and distributions, and truck activity diaries	Similar to the model F; Captured movements of vehicles and decisions of carriers more realistically; May allow truck tour-based microsimulation to more accurately forecast truck movements at local level

Table 6: Class advantages and disadvantages

(Doustmohammadi et al., 2016)

Overall, the published literature on the application of tour-based models (Class G) suggests that such models are more realistic than traditional freight demand forecasting models and best suited for impact assessment and vehicle-based policy studies, such as pricing strategies and restricted lanes (Doustmohammadi et al., 2016).

Table 7 shows the relationship between each class and its inputs/outputs. It can be very useful to observe which type of model to reference depending on the available data.

Table 4. Freight Model Characteristics by Input/Output

Output Input	Commodity O-D	Mode Choice	Supply Chain	Truck O-D	Truck Route Flows
Socioeconomic data	Class D, E	Class D, E, F	Class F	Class B, C, D, E, F	Class A, B, C, D, E, G
Land use data	Class E	Class F	Class F	Class E	Class E, G
Transport supply/demand	Class D, E	Class D, E, F	Class F	Class C, D, E, F	Class A, B, C, D, E, G
Commodity flow data	Class D	Class D, F	Class F	Class D, F	Class D, G
Truck O-D				Class B	Class B, C, G
Shipment characteristics		Class F	Class F	Class F	
Trans-shipment Point data		Class F	Class F	Class F	
Logistics costs		Class F	Class F	Class F	Class G
Vehicle tour characteristics					Class G

Table 7: Freight model characteristics by Input/output

(Doustmohammadi et al., 2016)

These last two tables will be helpful to classify and compare the real-world models analyzed in the next part of this paper.

4 Real-world models

This section deals with macro-models that have been developed by public entities to forecast freight demand in different zones of the world. They all use concepts and frameworks explained in sections 2 and 3 to forecast freight demand in Belgium, Europe, the United States and western Asia in practice.

4.1 PLANET Model

4.1.1 Description of the model

The PLANET model is a long-run model oriented towards transport in Belgium. It was developed by the Federal Planning Bureau in collaboration with the SPF Mobility and Transport. The bureau comes up with long run outlooks on the freight and passenger demand in Belgium every three years. The model also aims to:

- develop long run perspectives on the freight and passenger demand in Belgium,
- evaluate the effects of transport policies, pollution costs and congestion costs on the actual transport activity,
- carry out profit analysis of these transport policies.

The PLANET model is composed of seven interconnected modules:

1. The Macro module
2. The Transport Generation module
3. The Trip Distribution module
4. The Modal and Time Choice module
5. The Vehicle Stock module
6. The Welfare module
7. The Policy module

These 7 interconnected modules are illustrated in Figure 4, their relationships will be explained further in the next pages (Vandresse, 2016)

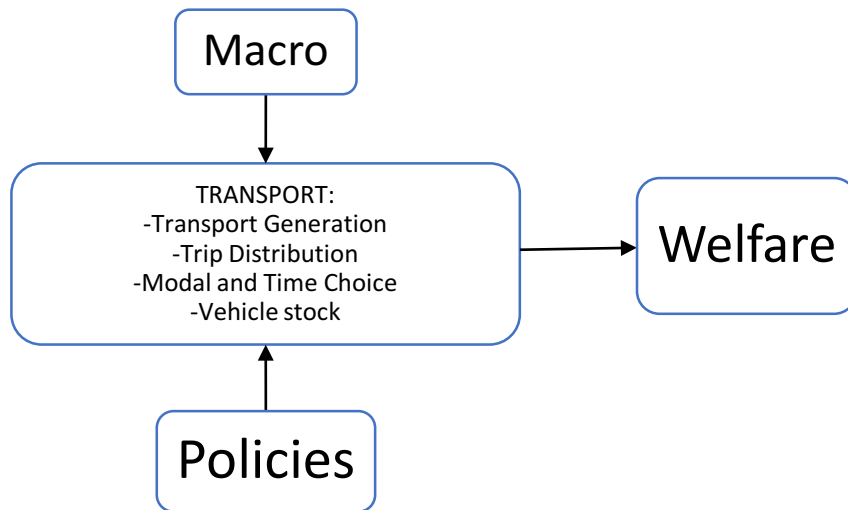


Figure 4: Modules of the PLANET model

(Vandresse, 2016)

The *Macro* module sets macroeconomic projections at the level of the “arrondissements” (NUTS3³ zones) of Belgium.

The results from the Macro module are considered as exogenous to the rest of the model. Only the impact of the economic and demographic variables is examined.

The *Policy* module outlines the transport policies used in usual and alternative scenarios. These consist in transport instruments (such as fuel taxes, vehicle purchase taxes or road pricing) but also encompass more general instruments such as labor taxation.

The part of PLANET focused on transport gathers four modules:

- The *Transport Generation* module determines the number of journeys between home and workplace for each zone, the number of other passenger trips and the total tons of freight transported nationally and internationally.

³ NUTS3 stands for the third level of the *Nomenclature of Territorial Units for Statistics*. It is a territorial division to ease comparison between countries or regions. The level depends on the number of inhabitants of the region. (Eurostat, 2017)

- The results of the Transport Generation module are provided to the *Trip Distribution* module which determines the number of trips taking place between each of the zones.
- Next, the *Modal and Time Choice* module aligns the modes by which the trips are made and the timeframe in the case of road transport (peak hours or off-peak hours). The choice of mode and time period depends on the costs and length of the different options. Travel time for the road modes is determined by means of the speed-flow function that links the average speed of transport modes to the traffic flow.
- The road transport demand obtained from that function is used as an input for the *Vehicle Stock* module. This module reveals the composition of the stock of vehicles registered in Belgium according to their size, the current EURO norm, and the cubic capacity and type of their engine. The *Modal and Time Choice* and *Vehicle Stock* modules disclose information on environmental costs of transports and public revenues generated from transports.

In some cases, results for a given year (t) determine the transport demand for the following year ($t+1$). Hence, the transport demand can depend on the average generalized cost of the transport flux of the previous year. Furthermore, the generalized cost of transport on a given year influences the following year's trip distribution. Finally, the Vehicle Stock has an impact on the following year's road transport costs. This is illustrated in Figure 5.

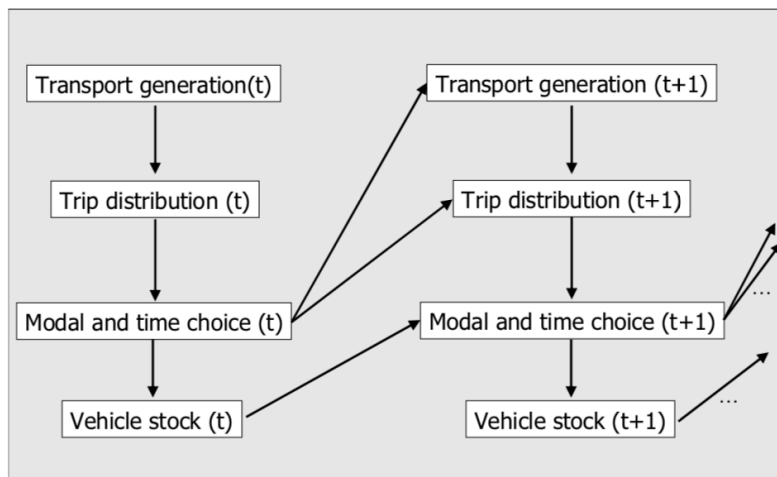


Figure 5: Links between transport modules in PLANET

(Vandresse, 2016)

The *Welfare* module calculates the effects of transport policies on welfare. The module takes into account effects on consumers, producers, public authorities and the environment.

4.1.2 Methodology

4.1.2.1 *Calibration of the model*

The exogenous parameters of the model are determined in the calibration step in order to better simulate the behavior of modal choice. The reference year for calibration is 2008 due to the fact that:

- All the necessary data is available for 2008 (costs, transport statistics...). It also follows the introduction of the latest uniform freight classification (NST 2007).
- No particular crisis or biasing events occurred in 2008. Calibrating on an atypical year wouldn't yield good results. The model, however, takes the economic crisis into account using macroeconomic forecasts conducted in 2011.

Following the introduction of the NST 2007, an origin-destination matrix has been built for the year 2008. The input data came from the Directorate General Statistics and Economic Information (DGSIE) for road and marine transport. The rail data was retrieved from the SNCB.

On a macroeconomic level, three main factors influence transport demand: employment, production and international trade. Hence, the evolution of the gross domestic product, the economic and financial crisis, the geographical spreading of employment and the changes of industry sector are some of the inputs of the PLANET model (Vandresse, 2016).

4.1.2.2 *Assumptions*

Three categories of costs are considered in the model:

- The monetary costs supported by the user (the acquisition and utilization of the vehicle),
- The time costs, the monetary value attributed to the journey time,
- The environmental costs, even though they often aren't the first concern of the individual choosing a transport mode.

The determination of forecasts of these costs requires several assumptions (considering only freight demand forecasting):

Monetary costs

- Motorization doesn't affect road freight transportation (all trucks and lorries are assumed to run on diesel). But diesel consumption is assumed to reduce over the years thanks to an increase of engine efficiency. Other costs are expected to remain constant.
- Regarding rail and marine freight transport, all costs should remain constant, except fuel and electricity costs, which will also benefit from enhanced energy efficiency.

Time costs

- The time spent in the vehicle is the only criterion assumed variable over the forecasted period (from 2008 to 2030).
- Roads are becoming wider and allow a greater capacity of vehicles, but the number of vehicle continues to increase. Hence, the road lead time is assumed to remain constant over the forecasted period.

Environmental costs

- Pollution emissions are bound to decrease over the forecasted period due to legal obligations to vehicle manufacturers to reduce the emissions of their vehicles.
- Regarding indirect emission costs, the production of electricity for trains is considered constant over the forecasted period. However, a consistent but moderate increase for hybrid and electric vehicles has been determined.
- The fall of nuclear power for alternative energy sources is also assumed.
- As the scarcity of fossil fuels increases, the indirect costs for transportation of petrol are predicted to increase. Fossil fuels will be collected farther away from their distribution points, which would require additional transportation (Vandresse, 2016).

4.1.2.3 Indicators used for the freight demand forecast

The different indicators are:

- The transported tonnage,
- The geographical spreading of the transport,
- The number of tons-kilometers,
- The modal and time choice.

Transported Tonnage

The transported tonnage is calculated from the evolution of economic activity (production, import and exports) and the average value of the transport of a ton. The latter depends on the type of freight transported and the distribution of the categories of freight within the transport streams. Marine, air and pipeline transport is determined exogenously. The tonnage for 2008 was determined statistically. For the years 2008 to 2030, the forecasts were made upon the annual growth rate of imports and exports per freight type (using the NST 2007 classification).

Geographical Spreading of the Transport

Origin-destination matrices are calculated endogenously to the model. The calculus is made upon a linked model established for each freight type (NST 2007) and is estimated from flows of goods (or freight flows) studied in 2008. The freight flows between geographic areas (“arrondissements” in this case) depend on the generalized cost of transport and the characteristics of the area (population, production, presence of port or airport...). Public regulations and other commercial barriers are also taken into account.

The Modal and Time Choice

Four transport modes were considered: lorries, trucks, trains and barges. Regarding road freight transport, transporters could choose their time period of travel (peak hours or off-peak hours). The modal and time choice is determined from the generalized cost of transport and the freight demand characteristics (for instance, the type of freight) (Vandresse, 2016).

4.1.2.4 *Alternative Scenarios*

This model strongly relies on the assumptions detailed in section 4.1.2.2. It was built around the policies under application in 2008. Yet, government representatives constantly work on strengthening rules against the negative effects of transport. Hence, some alternative scenarios were explored for this model:

- Hybrid and electric cars might develop and sell more than what has been predicted by 2030
- The introduction of stronger taxes on kilometers is likely to be applied to lorries and trucks.

Additional scenarios might occur, but this last part of the report reveals the sensibility of the model and stresses the importance of the assumptions it is built on (Vandresse, 2016).

4.2 TRANS-TOOLS model

The TRANS-TOOLS (TOOLS for TRansport Forecasting ANd Scenario testing) model was funded by the European Commission and DG MOVE (De Jong et al., 2016).

“The TRANS-TOOLS model is the most recent state-of-the-practice transport-oriented 4-steps forecast model available at EU level.” (Johansen, 2016, p.1). Its goal is to produce a European transport network model for both passenger and freight. It also deals with intermodal transport, which is a first for European transport network models.

It produces results such as:

- *Transport performance (passenger-kilometers, ton-kilometers, vehicle kilometers, etc.) at EU, national and NUTS⁴ II level, on an annual basis or more detailed,*

⁴ NUTS1: 3 to 7 million inhabitant zones

NUTS2: 800 000 to 3 million inhabitant zones

NUTS3: 150 000 to 800 000 inhabitant zones

- *Modal split (share of demand using road modes, rail, air, etc.) both with reference to passengers and tons and to passenger-kilometers and ton-kilometers, thus also average distance of transport can be assessed,*
- *Load on corridors (passengers and freight vehicles) and, therefore, levels of congestion on road infrastructure (TEN-T and main national links), and also, annual averages, and if available, daily peaks (Johansen, 2016, p.1).*

The model sorts the flow of goods per commodity type. The freight and logistics model of TRANS-TOOL 3 takes scenarios as input and uses the pivoting approach to generate origin-destination matrices.

The TRANS-TOOLS 3 freight model is composed of:

- *A trade model to deliver zone-to-zone production-consumption matrices in tons;*
- *Single modal route choice models to generate transport costs and times, and chain route choice models to calculate overall costs and times for a number of multimodal transport chain types;*
- *A disaggregate logistics model for transport chain choice (encompassing mode choice);*
- *Aggregation to zone-to-zone OD flows (tons);*
- *Chain route choice and network loading (assignment of tonnes);*
- *Pivot-point procedure to pivot predicted OD flow changes on a base OD matrix;*
- *Unimodal Assignment of flows onto the networks. (De Jong et al., 2016, p.13)*

The specificity of the TRANS-TOOLS sub-models is that they rely on the Aggregate-Disaggregate-Aggregate (ADA) model. These ADA models have the characteristic of including logistic decisions such as shipment size and transport chains.

Transport flows between NUTS3 zones are covered in the Trans-Tools trade models. They rely on the gravity model explained in section 2.5.2.1.

The main freight model links the demand and supply of transport. The model can thus predict changes due to user input/scenarios regarding the supply and demand.

Figure 6 groups the sub-models into the 3 model components: the level of service calculation, the demand response calculation, and the route choice and network loading. The next section will briefly explain how the model is used as a whole.

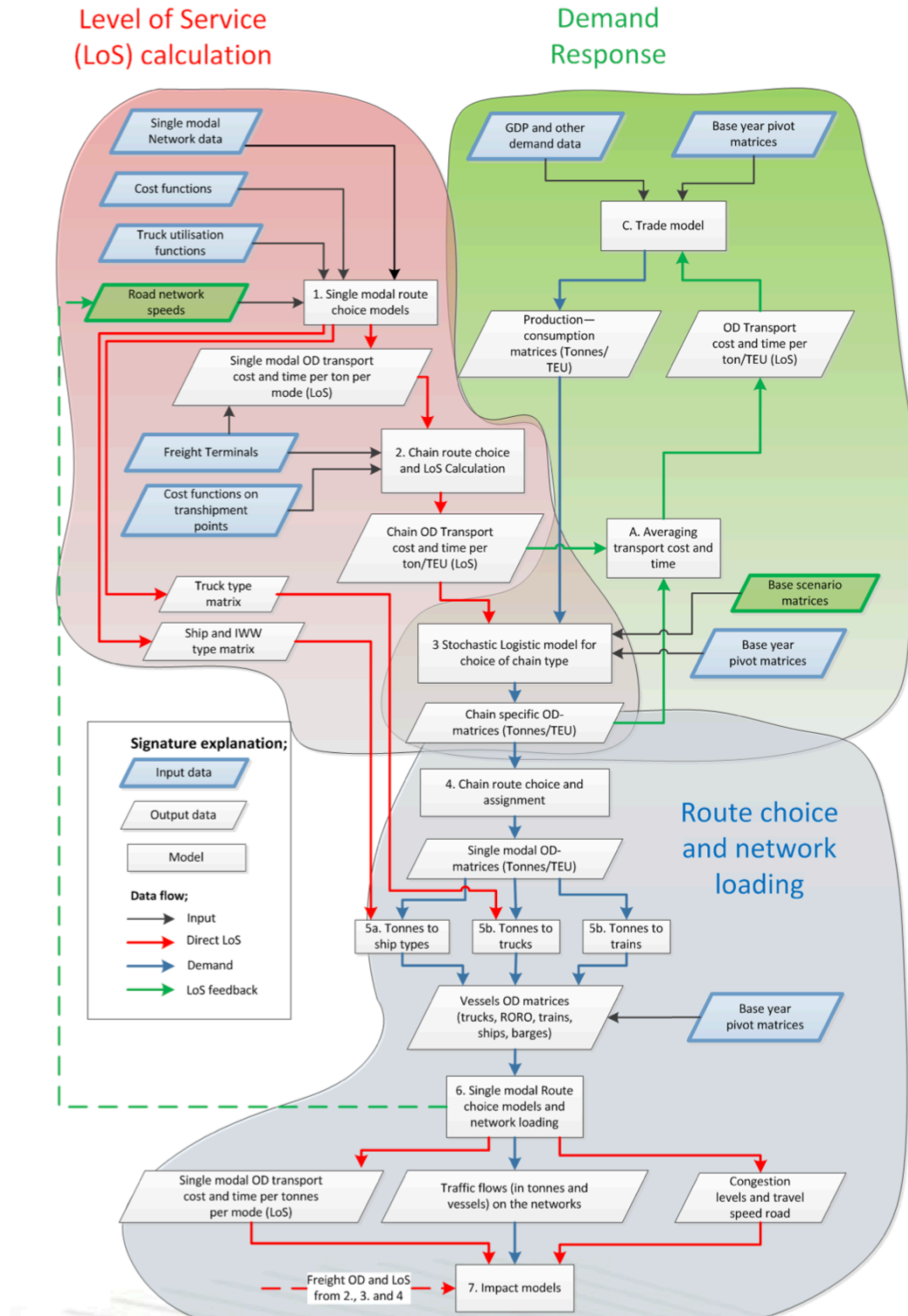


Figure 6: Main model-groups in the Freight and Logistics Model flow

(De Jong et al., 2016, p.20)

4.2.1 Level of Service calculation

“The Level of Service (LoS) is a mechanism used to determine how well a transportation facility is operating from a traveler’s perspective” (Thurston Regional Planning Council, 2016)

1 – Single modal route choice models⁵

The first step of the Level of Service calculation is the use of single modal *route* choice models. These models calculate level of service (which refers to the time and cost per container or tons) between zones and terminals for road transport, and other modes (rail, inland waterways, sea). This first step considers trucks separately because they can drive directly to the end users, whereas trains and boats require terminals.

The first type of network is the road route. It uses inputs such as length, modelled congested speed, tolls (evaluating time lost at tolls), and time loss at certain border crossings.

The rail route choice sub-models recognize 3 possible combinations of traction;

- . Diesel the entirety of the journey,
- . Switching of locomotive from diesel to electric or vice versa,
- . Electric the entirety of the journey.

Diesel is the most expensive method and changing locomotive will result in switching costs (both in time and capital).

Hence, the rail route choice model considers travel speed, travel length (with separate cost functions for diesel/electric trains), types of goods, infrastructure charges, various border crossing time penalties and change of traction type.

The inland waterway freight route choice calculates the cheapest cost routes between possible pairs of Inland Waterway Terminals. The input criteria are the commodity type, the length of the link, and the speed of operations.

⁵ The main models are numbered from **1** to **7** and will be in split in the following sections

The sea freight transport route choice model finds the cheapest cost routes between possible pairs of harbors depending on the type of freight transport and what the harbor can host. The criteria are shipping size and ship type. Calculations are made separately for containers, dry and liquid bulk transport (de Jong et al., 2016).

2 – Chain route choice model and level of service calculation

In this second step, the model determines the combinations (chains) of modes and calculates the costs and time of the optimal chain.

3 – Stochastic Logistic model for choice of chain type

The third step collects the results from the Level of Service and the demand response groupings. It also takes O-D matrices as an input. “Since this has not (yet) been calculated for the scenario that a user is evaluating, there is a need for a user-defined base scenario matrix as input. [It also requires a base pivot year matrix to] ensure that the model fits the observed 2010 base matrix if the model is run for the 2010 base year” (De Jong et al., 2016, p.25).

It then calculates the probability for choosing chain type for freight category. The output is chain type specific O-D matrices (measured in tons or containers).

4.2.2 Demand response

The demand response grouping is a loop. It models how demand for freight transport reacts to change in the level of supply. The Level of Service is the input of the freight demand modeling.

The first step in the demand loop is to calculate the average transport cost and time.

Then, the unit adjustment model converts the units of cost and time per ton/container to monetary values.

Next, the trade model intervenes and calculates the change of trade patterns. The trade model takes commodity groups and outputs production-consumption matrices (measured in tons).

And finally, the Stochastic Logistic model for choice of chain type is rerun with the calculated Production-Consumption matrices for the scenario year retrieved from the trade model instead of the base scenario matrices.

4.2.3 Route choice and network loading

“The last sequence of models calculates final routes for freight transport and assigns (loads) the flows onto the networks.” (De Jong et al., 2016, p.26).

4 – Chain route choice and assignment

The first model is the same as the chain route choice model (2), except that here, the flows are also loaded onto the network. The output is multimodal network flows in tons/containers for the four freight categories. Both single-modal and multimodal O-D matrices can be directly drawn from the output network.

5 – Calculation of vessel flows

This step computes the number of vessels between a given origin-destination pair given the freight flow and size.

The unit of calculation is decimal. For instance, the model may calculate that a given transport is using 0.5 trucks. The reason for this is that the model is a weekday model. So, this basically states that the likelihood for a given truck transport on a given day is 50% (De Jong et al., 2016, p.27).

There are specific calculations of vessel flows for each type of route.

6 – Single model route choice and network loading

Similarly to the first step (1), this model calculates the flows on the network, but both freight flows (tons and containers) and vessels are considered on the network.

7 – Impact models

This last model shows the impacts of a given scenario run compared to a base scenario.

The impacts are calculated by the means of the following outputs from the freight models:

- *Network level: the output is vessel flow and speeds on the transport network level. This is used for impact calculations of safety, energy use, CO2 and other emissions.*
- *Unimodal O-D level: the outputs are freight flows (tons or containers) and vessels and level of service (time and cost) for unimodal transports between terminals and end zones (points of production and consumption).*
- *Chain O-D level: this calculates O-D Freight flows and level of service per chain type per freight category for an entire chain type for a given O-D pair.*
- *Production-Consumption freight level: this provides the overall PC matrices and level of service for each of the four freight categories*
- *Production-Consumption economic level: Overall level of service for each of the four freight categories (De Jong et al., 2016, p.29).*

4.2.4 ETISplus Database

ETISplus was a project run by the Institute of Transport and Communications that aimed to design and build a European transport information system. The project was launched in 2009 and was finished in 2012. It was implemented by 18 transport consultants from 11 European countries. It was built on the conclusions of the ETIS project, which was completed in 2005. The project was financed by the European Commission. The TRANS-TOOLS model drew its data from the ETISplus Database.

ETIS stands for European Transport policy Information System. Its goal is to create a database of harmonized European data. It intends to link official statistics to applications in the transport policy theme.

“One of the goals in ETISplus, under this thematic area, was to take all of the data that is publicly available, and use it to generate a best possible estimation at the disaggregate level” (Szimba et al., 2013, p. 21.1)

The main data collection occurred in 2010. However, policy changes can have a drastic importance on the accuracy of the forecasts inducted from this data. These changes are certain, yet their timings are unpredictable. Which is why one of the focuses of the project was to integrate the change in the database updating strategy (Szimba et al., 2013).

4.2.4.1 Methodology: Data Harmonization

Harmonizing data from different countries requires agreements on common definitions. For instance, the definition of the geographic areas within countries must be decided. The mode of transportation and what they encompass has to be defined thoroughly. These definitions might also change over time. Hence, compromises had to be made beforehand in sake of data integrity.

In ETISplus, the data is always matched to a year with its corresponding definition. If a major new collection of data is proceeded, it will come with its matching current definitions. If the user wants to use data from two different periods, the data has to be converted.

The model contains road, rail, waterway, air networks converged in GIS (Geographical Information System) database sets. It gathers information on the shortest paths, road types, transport modes that can be used on the roads... Other linked database sets regroup the data on the terminals and other facilities for modal change (Szimba et al., 2013).

4.2.4.2 Methodology: Freight Matrices

ETISplus created freight matrices for each transport mode. Data availability varies according to the mode; it was retrieved through different sources. Different approaches were taken for each mode.

Road Freight Matrix

Road data was provided by Eurostat and available at the NUTS3 level.

The Origin Destination matrix is made in two steps:

First, the domestic O-D matrix is calculated for each country, since these each relate to a single data source.

Then the international O-D matrix is made for all country pairs together (Szimba et al., 2013, p. 19.5).

The advantage for domestic matrices is that the Eurostat provides a thorough NUTS3 to NUTS3 matrix. “The international matrix is more complicated. The incoming and outgoing international transport per zone has to be derived from the total transport and domestic transport per zone. By subtracting the domestic from the total, the international part can be estimated.” (Szimba et al., 2013, p. 19.6)

Rail Freight Matrix

Like the road freight matrix, the rail freight matrix was done in stages. The main problem is that no data was available at the NUTS3 level. The data was only available at the NUTS2 or NUTS1 levels, hence data had to be induced. However, the availability of train freight data is relatively low compared to other transport modes. Furthermore, the amount of released per country is inconsistent. For example, Eurostat published no information the domestic traffic of Germany by 2010, despite the country being one of the most important in freight traffic in Europe.

As no data was available for 2010, data from 2005 from Eurostat or national data files was used. Then, information was induced and harmonized to the 2010 ETISplus standards (Szimba et al., 2013).

Waterway Freight Matrix

Data for waterway freight is also quite limited. It is not available at the NUTS3 level, but for the most part, Eurostat was able to provide it at the NUTS2 level. The missing data was retrieved from national sources then converted and harmonized. Once again, the lack of information and consistency for international data was challenging. The waterway freight matrix required the most national sources and the most numerous conversions and induction (Szimba et al., 2013).

Air Freight Matrix

Air cargo represents a lesser share of freight transport (about 2% of the freight transported globally). The freight was assigned at an aggregate level (per country) which had to be disaggregated at a NUTS3 level.

Air cargo focuses on products that have a much shorter life cycle or which have specific characteristics (such as perishable goods). It is ten times more expensive than surface transport. Additional information is present in this matrix given the level of detail. The data was retrieved from WorldNet as a starting point, then induced and harmonized (Szimba et al., 2013).

4.3 FAME

The Freight Activity Micro-simulation Estimator (FAME) is a freight activity-based modeling framework. According to Pourabdollahi et al. (2012), it is the first comprehensive nationwide freight demand microsimulation model in the United States. It is composed of five modules (illustrated in Figure 7). It forecasts freight movements at the disaggregate level (firm-to-firm).

The model was developed based upon publicly available datasets because “states and planning agencies will more likely adopt models that can be developed using only widely available data” (Kawamura et al., 2012, p. 26). The data is mostly retrieved from the County Business Patterns dataset.

Four categories of data are required for developing FAME:

- *information on business establishments for the first module,*
 - *aggregate freight flows (OD matrix) for the second module,*
 - *detailed information on a sample of individual shipments and supply chains for the third and fourth tasks,*
 - *and specifications of the transportation networks for the last module.*
- (Pourabdollahi et al., 2012, p.2)

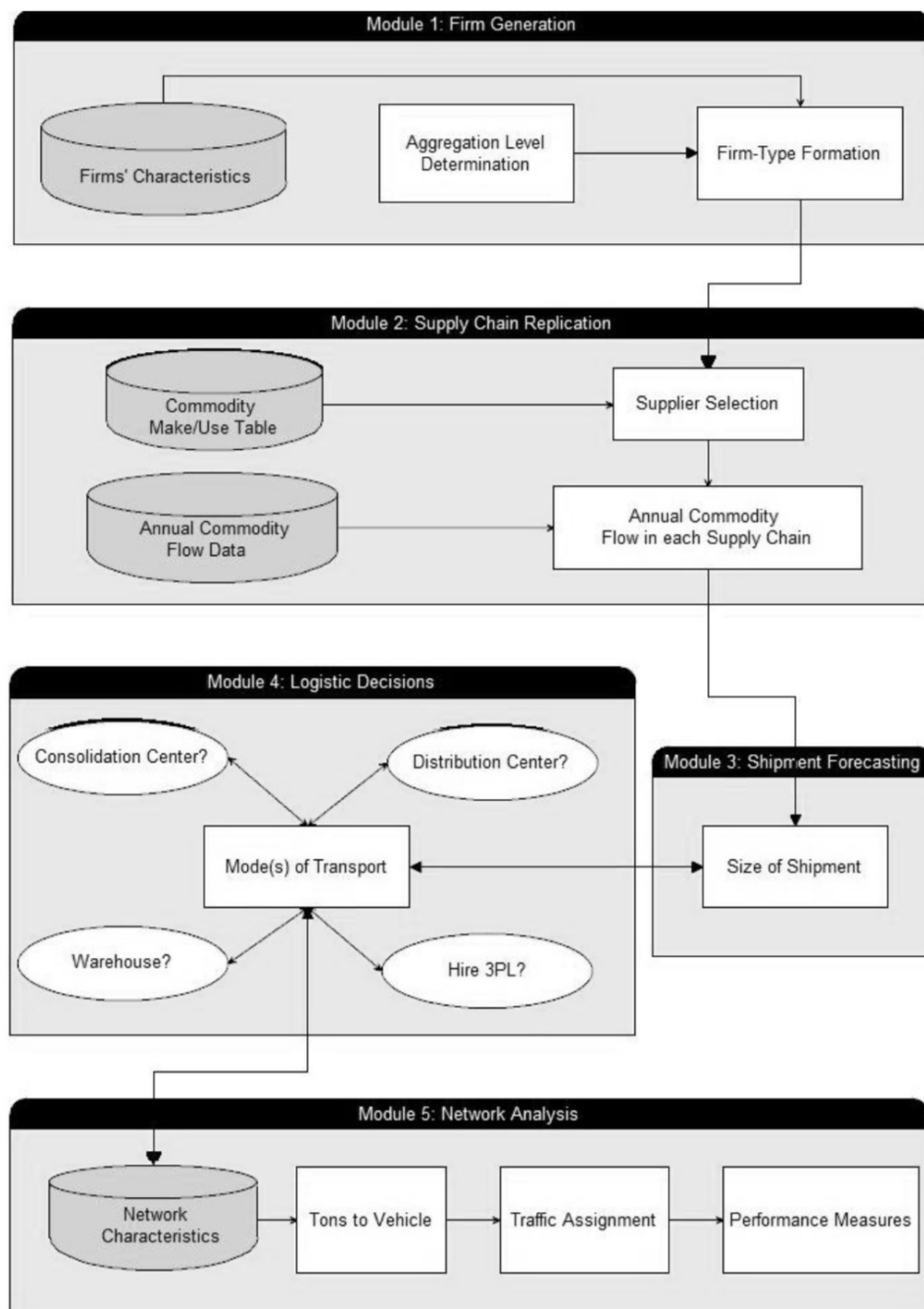


Figure 7: Framework of the Freight Activity Microsimulation Estimator

(Kawamura et al., 2012)

4.3.1 Firm-type Generation module

This first module identifies all the firms concerned by the study and collects their relevant characteristics. The characteristics include firm type, industry, location and number of employees (Kawamura et al., 2012).

In 2009, the County Business Patterns dataset had information on over 239 000 firms. In theory, the FAME model can consider all these companies at once, but the computation costs and inputs required would be considerable. Hence, the model aggregates the firms based on firm types. “A firm-type is a collection of firms with similar location, industry type, and establishment size” (Kawamura et al., 2012, p. 25). The type of industry is defined by the 2002 North American Industry Classification System (NAICS), which contains around 1200 categories (U.S. Census Bureau, 2002). An important assumption of the FAME model is that firms with similar characteristics have similar behavior in the freight decision-making process (Kawamura et al., 2012).

4.3.2 Supply Chain Replication module

The second module investigates the type and the amount of goods entering and leaving each firm, depending on each firm’s characteristics.

The supply chains are replicated by matching suppliers and buyers of goods. The module is divided in two steps.

The **first step** determines all the potential suppliers for a given firm type. It lists all the firm types that could provide a commodity to a selected firm-type.

The **second step** uses probabilities (based on a fuzzy expert system detailed by Kawamura et al. (2012) in their appendices) to determine supplier suitability. These probabilities are namely built upon the commodity researched, the commodities characteristics, the distance between the firms, the characteristics of the firms and the financial situation of the firm (using the number of employees and the growth rate) (Kawamura et al., 2012).

4.3.3 Shipment Forecasting module

This module uses the data previously collected on firms to determine the size of the shipment. It also considers the firms’ trade habits.

This module breaks down annual flows of a given commodity between two firms into single shipments. The output is expressed in terms of weight and distance.

Shipping distance is provided in nine categories (<50 miles, 50-99, 100-249, 250-499, 500-749, 750-999, 1000-1499, 1500-2000, >2000), and shipment size is also given in nine categories (<50 lbs., 50-99, 100-499, 500-749, 750-999, 1000-9999, 10000-49999, 50000-99999, 100,000<) (Kawamura et al., 2012, p. 40)

However, the next module requires precise weights and distances and won't accept categories. Hence conversions and rough estimations can sometimes be performed in case of lack of data (Kawamura et al., 2012).

4.3.4 Logistic Decision module

The mode choice engulfs decisions on shipping mode, haul time, shipping costs, warehousing, etc. The mode split model used in this logistic decision module is the only module that uses not publicly available data. The information was drawn from a survey of businesses (The UIC National Freight Survey) carried out across the USA.

*To achieve the goal of developing a model that can only be used only with publicly available data, FAME includes **two freight mode choice models** that were calibrated based on the UIC National Freight Survey. Depending on the availability of input variables, agencies will be able to select the powerful yet data-hungry **explanatory model** or the **parsimonious model** that requires a limited amount of input variables (Kawamura et al., 2012, p. 41).*

The variables for each sub-model are listed in Table 8.

The Explanatory Model	The Parsimonious Model
Mode (rail and/or road)	Mode (rail and/or road)
Distance	Great circle distance
Weight of the shipment	Weight of the shipment
Monetary value of the shipment	Highway and Rail Impedance
Shipping costs (for rail and road)	Containerized shipment (Boolean)
Shipping time (for rail and road)	Commodity type
Access to intermodal facilities (Boolean)	
Perishable Goods (Boolean)	
Consolidation center (Boolean)	
Distribution center (Boolean)	
Warehouse (Boolean)	

Table 8: Input variables for the explanatory and parsimonious models

(Kawamura et al., 2012).

4.3.5 Network analysis module

Finally, this module finds the impacts of the goods movement on the transportation network.

The general cost of commodity transportation obtained in the network analysis can be fed back into the second modules for multiple iterations, until stabilized commodity flows and costs are achieved.

4.4 The Integrated transport system in the Arab Mashreq (ITSAM)

4.4.1 Description of the model

The ITSAM was created in an effort to counteract an export fall in the 1990s. Intraregional trade was very low within the member countries of the United Nations Economic and Social Commission for Western Asia (ESCWA). The model simultaneously predicts trip generation, trip distribution, modal split, and trip assignment.

ITSAM contains three components:

- *ITSAM-NETWORK, an integrated (multimodal) transport network*
- *ITASM-INFOSYS, an associated information system*
- *ITSAM-FRAMEWORK, a methodological framework for issue analysis and policy formulation* (Nabil et al., 2004, p.129)

The goal was to create a model to predict freight flow patterns (time and cost) that can describe the behavior of importers and exporters for given countries and commodity sets. The model aims to facilitate intraregional trade and to integrate the countries of ESCWA into the global economy (Nabil et al., 2004).

4.4.2 Development of the model

The Simultaneous Transportation Equilibrium Model (STEM) was introduced in 1982 by Safwat. This model links transportation costs to the trip generation step of the 4-step model (Safwat, 1988). The STEM was then applied to the passenger and freight intercity transport system in Egypt, to the urban transportation network in Texas, to the urban transportation network in Saudi Arabia, etc.

The STEM was then improved to become the International Freight Simultaneous Transportation Equilibrium Model (IFSTEM) in 1999. IFSTEM is a core component of the ITSAM-FRAMEWORK.

“IFSTEM supports the transportation of several products par several modes” (Nabil et al., 2004, p.130). In this model, a product is a category of commodity or goods.

The network infrastructure represented by IFSTEM gathers cities, border points, rail/road terminals, seaports, airports and intermodal terminals. The model also considers origins, destinations, import and export procedures, transit procedures, and transfer operations. All these are represented by either physical nodes (ports, terminals...) or artificial nodes (transit procedures, transfers...). These artificial nodes are also known as Administrative and Logistical Operations (ALOs).

“Since each commodity can be transported by a specific mode or set of modes, depending on commodity characteristics, a network for each commodity type has been created under IFSTEM” (Nabil et al., 2004, p.130). The network is next decomposed in multimodal origin-destination matrices. The model then maximizes the exporter’s (or importer’s) utility function.

4.4.3 Assumptions

“IFSTEM is constructed in such a way that commodity exporters make decisions about where and how to transport their freight; choices are made regarding destination, mode, transshipment, and routing” (Nabil et al., 2004, p.132).

The delivery cost/price is evaluated with the value of time, the total time, the unit price of a commodity, ALOs costs on the current multimodal path, the tariff cost and the transportation cost.

“The measured utility is a function of the socioeconomic characteristics of the destination (such as consumption levels, commodity deficit, population, and selling prices)” (Nabil et al., 2004, p.134).

The model is influenced by how the exporters value time, the Electronic Data Interchange (EDI, level of implementation used to perform operation a for commodity r) and the transport system performance.

The demand for the transport of one commodity is independent to any other commodity's demand. Which is why they are modeled separately, and IFSTEM may therefore be decomposed by commodity type.

Additionally, since capacity issues are generally not a principal concern in regional or international freight transportation planning, it is not necessary to assign multicommodity flows to this international network simultaneously; a simplified separation of freight into commodity groupings is sufficiently relevant. Each commodity or sector becomes a layer, and together, all that is useful for planning. (Nabil et al., 2004, p.135).

The real application of IFSTEM to the ESCWA countries' transportation systems requires a calibration process for both demand and supply models to estimate the parameters. The inputs necessary for the ITSAM model are listed in Table 9.

The variables of the ITSAM model
The modes to consider (road, rail, sea, air)
Commodity types
ALOs
OD matrices
Multimodal Network
Generalized cost per unit
Number of tons of any commodity
Consumption levels
Commodity deficit
Population
Selling prices

Table 9: ITSAM model variables

4.5 Comparison of the real-world models

	<u>PLANET</u>	<u>TRANS- TOOLS</u>	<u>FAME</u>	<u>ITSAM</u>
Smallest level of results	NUTS3 ⁶ in Belgium	NUTS3	Type of firm	City or NUTS2
Calibration reference year	2008	2010	2009	2001
Modes	Road, rail and inland water, sea, air	Road, rail and inland water, sea, air	Road, rail	Road, rail, sea, air
Number of modules	7	7	5	4
Time sensitive (peak hours)	Yes	Yes	Yes	Yes
Time necessary as input	Yes	Yes	Yes/no ⁷	Yes
Environmental and Energy concerns	Yes	Yes	-	-
Presence of terminals considered	Yes	Yes	Yes/no ⁷	Yes
Passenger or Freight forecasts	Both	Both	Freight	Freight
Type of data	Aggregate	Aggregate	Disaggregate	Aggregate
Major variables	Cargo weight, Vehicle type, Employment	Cargo weight, Vehicle type, commodity type	Cargo weight, Type of business, Commodity Type, industry segments	Cargo weight, Commodity type, sales
Output Level of geography	NUTS3, National	National	Firms, National	Metropolitan, National

Table 10: Comparison of the real-world models

⁶ NUTS1: 3 to 7 million inhabitant zones

NUTS2: 800 000 to 3 million inhabitant zones

NUTS3: 150 000 to 800 000 inhabitant zones

⁷ Yes for the explanatory model, no for the parsimonious model

Table 10 gathers the factors relevant to compare between the four real-world model analyzed.

Each macro-model takes different variables as input. Some variables are common for all models (mode, commodity type, weight) while others are specific (employment, sales, type of business).

As it was suggested in section 2.6, all these macro-models can solve activity-based problems. They all consider temporal and spatial factors, but also transport actors and logistics⁸.

As de Jong et al. (2012) mentioned, the inclusion of logistics into these models has been a recurring theme since 2004. The behavior of the shippers, freight forwarders, carrier and drivers are somehow integrated in all four of these models (which was a shortcoming of the classic four-stage model).

Nowadays, the inclusion of logistics is recommended. The last decade's trends revealed the need of advanced inventory management to face, inter alia, returns and exchange policies. Anyhow, the consideration of transport costs to the final price of a commodity remains of paramount importance (For instance, the share of transport costs in the final price of bricks is much more important than the share of transport costs in the price of electronics).

Production-Consumption links are also taken into account in all these models. TRANS-TOOLS even explicitly utilizes PC matrices. All four macro-models somehow consider production sites and shippers, intermediate processing points and intermediate/end consumer sites.

⁸ *The logistics models typically include all logistics costs and aim to describe the trade-offs between inventory and transport and the routing via "intermediate points". These points can be transshipment terminals (road terminals, rail terminals, ports, airports) that are used for transfers between modes or different vehicle types of the same mode (allowing for consolidation and deconsolidation), or distribution centres (DC), warehouses where goods are stored, not necessarily labelled for their destination and where goods typically change consignment sizes (de Jong, 2012, p.14).*

PLANET, TRANS-TOOLS and ITSAM rely on aggregate data for input, which is coherent to the contents of section 3.1. The availability of disaggregate data being one of the major obstacles, most models take aggregate data as input. One of the major preoccupations of the development of these macro-models was to build non-confidential databases, which could only be done with publicly available (and frequently, aggregate) data.

FAME and ITSAM are firm centered, they provide forecasts at the establishment level. Moreover, they take into account actors' value of time. FAME only considers firms compatible with the type of goods sent. ITSAM considers more socioeconomic variables than the others (consumption levels, commodity deficit, population, and selling prices, sales).

The PLANET model was the only one to address new technology concerns, such as the arrival of electric vehicles. Alternative future transport modes shouldn't be neglected (for instance, drones and autonomous robots might be tomorrow's postmen).

Environment and pollution concerns were only recalled by PLANET and TRANS-TOOLS. These models are quite close. They provide the same types of results, only the inputs vary.

All models are activity-based, but FAME's parsimonious model is the only one which doesn't require shipping time or any other time value to predict demand. The parsimonious model is quite powerful because it only requires a portion of the variables to forecast demand. Although, it is accepted that it might give less accurate results.

It is important to note that the parsimonious and explanatory models may give different results if they are both applied for the same scenario (having all the inputs available).

Referring to the class approached of section 3.4, PLANET, TRANS-TOOLS and ISTAM would be class F because they don't fully support tours, but have logistic information integrated. The FAME model can however solve tour-based problems (as Kawamura et.al (2011) demonstrated), and is then a part of class G.

All models have their pros and cons, the key decision elements as to which one to use are:

- The data available
- The (computing) time available
- The type of commodity to model
- The calibration reference year
- The level of aggregation desired
- The modes to consider
- The outputs desired

These models provide different approaches to forecast demand under different contexts and using different variables. There will forever remain room for improvement in terms of accuracy, performance and relevance of the variable. Further project could always improve transport planning decisions by suggesting more dependent variables, simulating the behavior of all actors, including upcoming trends and technologies... Future models could also consider the jointure of freight and passenger in one transport mode as part of the forecasts (as some ferries transport trucks, vans and passenger vehicles altogether).

5 Conclusions and Managerial Implications

This paper began with an explanation of the basic concepts of transport planning. Though it has limitations and was at first applied for passenger planning, the four-stage model was approached as a point of reference. To understand how to predict freight movements, the literature review focused on the first three stages (trip generation, distribution and modal split), on the modelling of these stages and the variables that come into play.

Then, the importance of these models to incorporate time series was highlighted. Activity-based models should be emphasized, and they should express an intuitive link to a macroeconomic metric such as the gross domestic product. Multi-agent systems are able to simulate freight transport generation and incorporate time series.

Next, the classification systems available for freight forecasting models were discussed. The employment was used as an example of a variable that can be used to predict commodity flows.

Section 4 focused on the methodology of macro-models developed across the globe. The aim was to present the methodology used in the making of the macro-models and understand their goals. The scope of these models is different, they were all relevant for different scales and different regions. They were chosen specifically for their diversity. Other studies, such as the paper provided by de Jong et. Al (2004), list macro-models developed in Europe.

Although the literature is still less abundant than for passenger transport planning, the importance of the freight has been illustrated many times.

Economic growth is impacted by transport problems. For instance, the freight supply of an area can be jeopardized by the lack of new transport tracks. “A heavily congested or poorly connected system restricts options and *limits* economic and social development” (Ortúzar & Willumsen, 2011, p. 3).

It’s important to notice the link between logistics and economy: monetary transactions are very often accompanied by a commodity flow (and vice versa). The amount of cargo produced and consumed are bound to economic transactions.

The freight system is complex, multidimensional, highly heterogeneous and depends on many economic agents. Classification systems are required to approach such a wide concept.

As reminded by Ortúzar & Willumsen (2011), models are simplifications of the reality based on assumptions to obtain satisfactory results. They will never provide entirely accurate results, and they will always require some interpretation.

It should also be considered that errors can occur all along the forecasting process. The chosen model could be inadequate, it could neglect a capital factor, the chosen year for calibration could be deficient, ... Let alone uncertainty and inaccuracies have a tendency to grow over time.

The models presented in the paper are primarily tools to aid decision-making. Some allow the calculation of profitability of infrastructure and help to allocate public funding accordingly. They help to test schemes and policies. Transportation policies aim to make sure that the public resources are being shared to sustain the economy and the society. They can also assess impacts on finances, the environment, safety, performance, and on traffic and congestion. Better models provide better forecasts and lead to better decisions.

Industry sectors, international trade networks, public agencies, and other policy makers need accurate information about national freight movements to continue efficiently delivering various goods within and among consumer markets. They need to plan for future freight traffic impacts and evaluate the effectiveness of policies and projects designed to alleviate freight congestion problems.

Finally, the resources allocated to models for analysis should be aligned with the cost of poor decisions. Transport investments and policies are extremely time-consuming, requiring models to predict scenarios that do not currently exist. Most freight forecasting models are expensive, but they can prevent much higher costs that could occur in the case of unforeseen circumstances (such as infrastructure costs).

6 Bibliography

- Brunel, J. (2007). *Prévoir la demande de transport de marchandises à long terme*. (Doctoral Thesis) Université Lumière Lyon 2, Lyon.
- Chapin F.S. (1974). *Human activity pattern in the city*. New York: Wiley.
- Cook, P., Das, S., Aeppli, A. & Martland, C. (1999). Key Factors in Road-Rail Mode Choice in India: Applying the Logistics Cost Approach. *Proceedings of the 1999 Winter Simulation Conference*.
- Cullinane, K. & Toy, N. (2000). Identifying Influential Attributes in Freight Route/Mode Choice decisions: A Content Analysis. *Transportation Research Part E* 36, 41-53.
- CUTR - The Center for Urban Transportation Research at the University of South Florida. (n.d.). *Analysis of Freight Movement Mode Choice Factors*, Online <http://www.fdot.gov/rail/Publications/Studies/Planning/ModeChoiceFactors.pdf> consulted on 18/05/2018
- Davidson W., Donnelly R., Vovsha P., Freedman J., Ruegg S., Hicks J., Castiglione J. & Picado R. (2007). Synthesis of first practices and operational research approaches in activity-based travel demand modeling, *Transportation Research Part A*, 41, 464–488.
- de Jong, G., Gunn, H.F. and Walker, W. (2004) National and international freight transport models: overview and ideas for further development. *Transport Reviews*, 24, 103-124. Online <https://doi.org/10.1080/0144164032000080494>
- de Jong, G., Nielsen, O. A., Anders Fjendbo Jensen, M. T., Henriques, M., Tanner, R., Johnson, D., Dekker, T., Sowa, V., Vierth, I., Bates, J., & Jeppe, R. (2016). Final report on the freight and logistics model. *Research and development of the European Transport Network Model – Transtools Version 3*. Denmark: Lyngby
- Doustmohammadi, E., Sisiopiku, V.P., Anderson, M.D., Doustmohammadi, M., & Sullivan, A. (2016). Comparison of Freight Demand Forecasting Models. *International Journal of Traffic and Transportation Engineering*, 5(1), 19-26. doi: 10.5923/j.ijtte.20160501.03
- Eurostat. (2017). *NUTS – Nomenclature des unités territoriales statistiques – Contexte*. Online <http://ec.europa.eu/eurostat/fr/web/nuts/background> Consulted on 20/04/2018

- Evans, S.P. (1973). A relationship between the gravity model for trip distribution and the transportation problem in linear programming. *Transportation Research* 7, 39–61.
- Ferber, J. (1995). Les systèmes multi-agents. *Vers une intelligence collective*, Interéditions, France: Paris.
- Fischer, M.J., Outwater, M.L., Cheng, L.L., Ahanotu, D.N., & Calix, R. (2005). An innovative framework for modeling freight transportation in Los Angeles County, *Los Angeles County Metropolitan Transportation Authority by Cambridge Systematics*
- Fried L., Havens J. & Thall M. (1977). Travel Behavior – a synthesized theory, *NCHRP Transportation Research Board*, USA: Washington DC
- Hägerstrand T. (1970). What about people in regional science? *Papers of the Regional Science Association*, 24, 7-21.
- Harker, P. T. (1987). Predicting intercity freight flows. *VNU Science press*. Netherlands: Utrecht.
- Holguín-Veras, J., Lawson, C., Wang, C., Jaller, M., González-Calderón, C., Campbell, S., & Ramirez, D. (2016). *Using Commodity Flow Survey Microdata and Other Establishment Data to Estimate the Generation of Freight, Freight Trips, and Service Trips: Guidebook*. Washington, DC. doi:10.17226/24602
- Jiang, F., Hohnson, P., & Calzada, C. (1999). Freight Demand Characteristics and Mode Choice: An Analysis of the Results of Modeling with Disaggregate Revealed Preference Data. *Journal of Transportation and Statistics*, 2, 149-158.
- Johansen, H. M. (2016). TRANS-TOOLS: A transport-oriented 4-steps forecast model. Online on the Tetraplan website <http://www.tetraplan.dk/kompetencer/forskningsprojekter/trans-tools.aspx>
Consulted on 5/05/2018
- Jones P.M., Dix M.C., Clarke M.I. & Heggie I.G. (1983). *Understanding travel behavior*. Aldershot: Gower.
- Kawamura, K., Mohammadian, A., Samimi, A., Pourabdollahi, Z. & Klekotka, M. (2011). *Development of a Freight Policy Analysis Tool for Northeastern Illinois and the United State - Freight Activity Microsimulation Estimator*. USA, Chicago: Urban Transportation Center
- Louie M.A. & Carley K.M. (2008). Balancing the criticisms: Validating multi-agent

- models of social systems, *Simulation Practice and Theory*, 16, 242-256.
- Mathew, T. V., & Krishna Rao, K. V. (2007). Introduction to Transportation Engineering, *NPTEL*, 7.1-7.5.
 - Mladenovic, M. N. & Trifunović, A. (2014). The Shortcomings of the Conventional Four Step Travel Demand Forecasting Process, 5-11.
 - Modi, K., Zala, L., Umrigar, F., & Desai T. (2011). Transportation Planning Models: A Review. *National Conference on Recent Trends in Engineering & Technology*.
 - Nabil, A., Safwat, K. & Hasan, M.K. (2004). Predicting International Freight Flows for Trade: Simultaneous Multimodal, Multi-Commodity, Network Equilibrium Model. *Transportation Research Record*, 129-139.
 - National Academies of Sciences, Engineering, and Medicine. (2012). *Freight Trip Generation and Land Use*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/23437>.
 - NC State University. (2011). *Supply Chain Resource Cooperative - Economic order quantity (EOQ) Model: Inventory Management Models: A Tutorial*, Online <https://scm.ncsu.edu/scm-articles/article/economic-order-quantity-eoq-model-inventory-management-models-a-tutorial> consulted on 24/05/2018
 - Ortúzar, S. J. D., & Willumsen, L. G. (1990). *Modelling transport*. (4th ed) Chichester, West Sussex, England: John Wiley & Sons.
 - Oum, T. H. (1989). *Alternative demand models and their elasticity estimates*. *Journal of transport economics and policy*, 23 (2), 163-188.
 - Pavon J., Arroyo M., Hassan S. et Sansores C. (2008). Agent-based modelling and simulation for the analysis of social patterns, *Pattern Recognition Letters*, 29, 1039-1048.
 - Piotte, J. (2013). *Modélisation des flux routiers de marchandises, de la génération de la demande par une approche multi-agents à l'affectation temporelle*. (Doctoral Thesis) Université Catholique de Louvain, Louvain-la-Neuve. Online <https://dial.uclouvain.be/pr/boreal/fr/object/boreal%3A126150>
 - Pourabdollahi, Z. (2011) A Behavioral Freight Transportation Modeling System. In Samimi, A., Mohammadian, A. & Kawamura, K. (Eds.) University of Illinois, USA: Chicago
 - Regan, A. & Garrido, R. (2002). Modeling freight demand and shipper behavior: state of the art, future directions. In Hensher, D., (ed) *The Leading Edge of Travel*

Behaviour Research, 185-216. Oxford (UK): Pergamon Press

- Safwat, K. & T. Magnanti. (1988) A Combined Trip Generation, Trip Distribution, Modal Split and Traffic Assignment Model. *Transportation Science*, 22 (1), 14–30.
- Saqalli M., Gérard B., Biielders D. & Defourny P.(2010). Testing the impact of social forces on the evolution of Sahelian farming systems: A combined agent-based modeling and anthropological approach, *Ecological modelling*, 221, 2714-2727.
- Sarkar, P. (2017). *Transportation Planning: principles, practices and policies* (2nd ed) India: New Delhi
- Statista. (2016). *Number of vehicles in use worldwide 2006-2015*. Online <https://www.statista.com/statistics/281134/number-of-vehicles-in-use-worldwide/> Consulted on 5/03/2018
- Suryani, E., Chou, S., Hartono, R., & Chen, C. (2010). Demand scenario analysis and planned capacity expansion: A system dynamics framework. *Simulation Modelling Practice and Theory*, 18(6), 732-751. doi:10.1016/j.simpat.2010.01.013
- Szimba, E., Kraft, M., Schimke, A., Schnell, O., Kawabata, Y., Newton, S., Breemers, T., Versteegh, R., Van Meijeren, J., Jin-Xue, H., De Stasio, C., Fermi, F., Ihrig Eckhard Szimba, J. (2013). *ETISplus Database: Content and Methodology*. Netherlands: Zoetermeer
- Thurston Regional Planning Council. (2016). *Level of Service Standard and Measurements*. Online <https://www.trpc.org/DocumentCenter/View/2798/Appendix-O--Level-of-Service-Standard-and-Measurement> Consulted on 7/05/2018
- U.S. Census Bureau. (2002). *North American Industry Classification System (NAICS)* Online <https://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart=2002> Consulted on 15/05/2018
- University of Idaho. (n.d.). *Travel Demand Forecasting – Theory and Concepts: Gravity Model*. Online <http://www.princeton.edu/~alaink/Orf467F08/The%20Gravity%20Model.pdf> Consulted on 3/05/2018
- Valbuena D., Verburg P.H., Veldkamp A., Bregt A.K. & Ligtenberg A. (2010). Effects of farmers' decisions on the landscape structure of a Dutch rural region: An agent-based approach, *Landscape and Urban Planning*, 97, 98-110.
- Vandresse, M. (2012). Perspectives de l'évolution de la demande de transport en

Belgique à l'horizon 2030. In Gusbin, D., Hertveldt, B. & Hoornaert, B. (Eds.)
Bureau fédéral du Plan, Belgium: Brussels

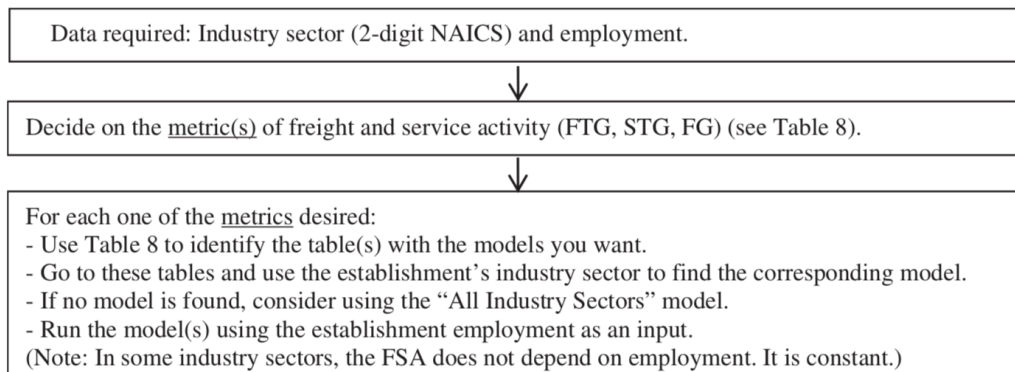
- Winston, C. (1985). Conceptual developments in economics of transportation: An interpretative survey. *Journal of economic literature*, 23 (3), 57-94.

7 Appendices

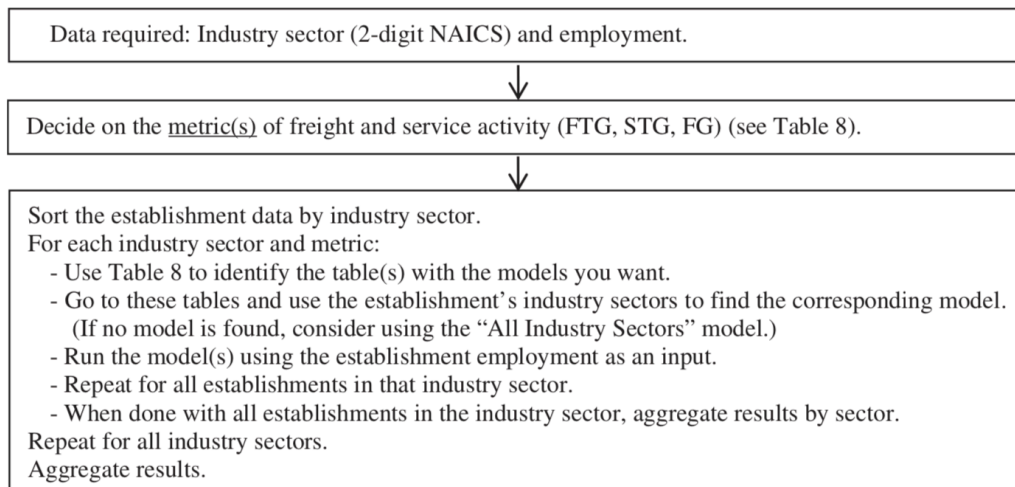
Appendix 1

The next flow chart summarizes the procedure to quantify the FSA according to Holguín-Veras, J. et al. The scenarios are the ones explained above in Table 5.

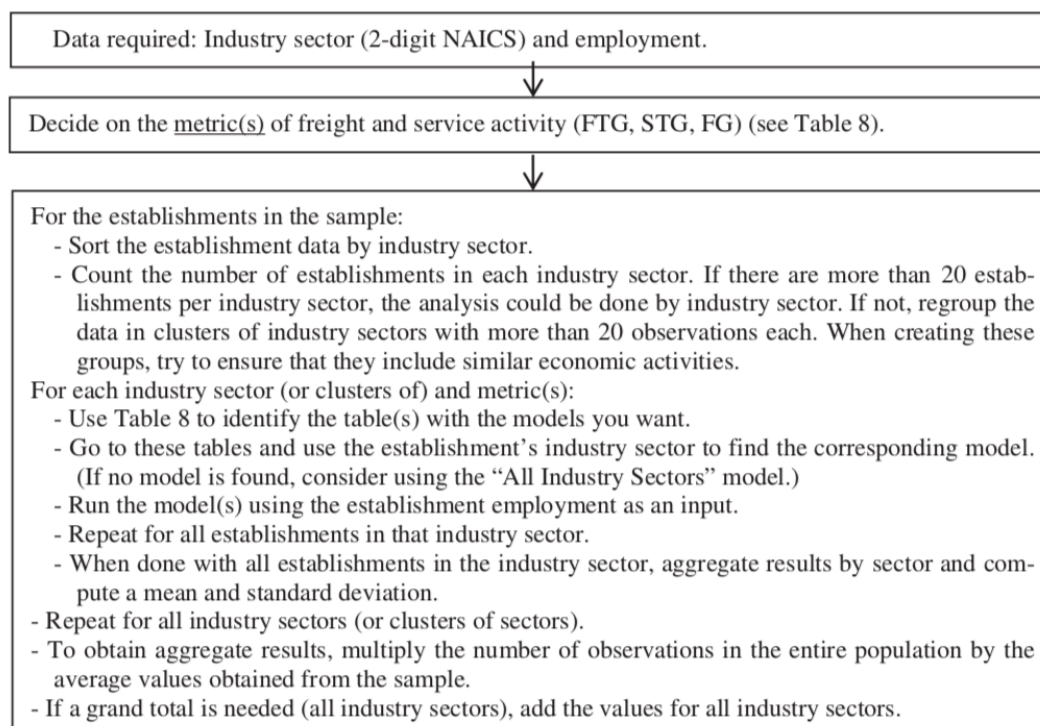
Scenario 1: Single Establishment (Establishment-Level Data Available, Linear/Non-Linear Models)



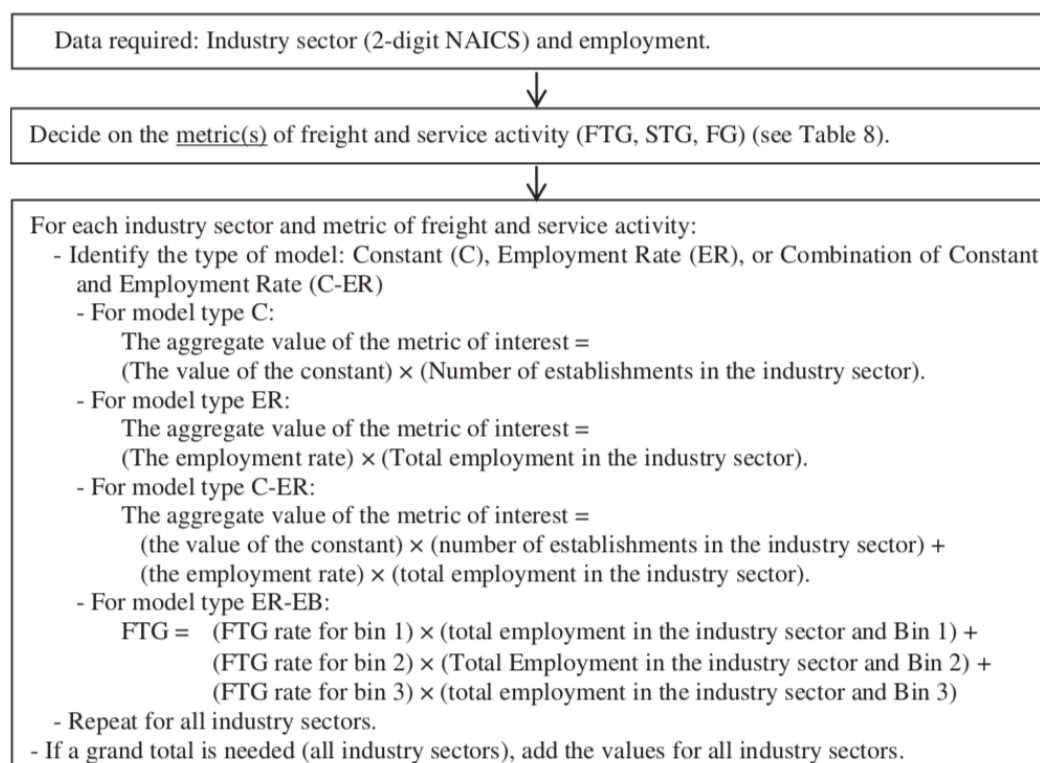
Scenario 2: Complete Enumeration (Multiple Establishments; Establishment-Level Data Available for All; Linear/Non-Linear Models)



Scenario 3: Sample Estimation (Multiple Establishments; Establishment-Level Data Available for a Sample; Linear/Non-Linear Models)



Scenario 4: Only Aggregate Data (Multiple Establishments; Only Aggregate Data Available; Linear Models)



(Holguín-Veras et al., 2016)

The models suggested separated the generation of demand from the generation of the FSA traffic, which allows them to be more flexible, realistic and implementable to a wider range of situations.

Even though the demand generation is the expression of economic fluctuations, the generation of vehicle trips depends on logistics and distribution. Depending on their locations, two identical establishments could generate different numbers of trips (if the establishment is close to other facilities, suppliers could group deliveries and have them more frequent, for instance).

This decomposition requires the usual measurements of demand forecasts (deliveries/day, shipments out/day, service calls/day, or pounds/day) to be transformed into vehicle trips. Appendix 2 summarizes the conversion of these measurements (Holguín-Veras et al., 2016).

Holguín-Veras, J. et al (2016) have displayed five limits of the models they have gathered:

- All their data comes from the United States, hence the relevance to any substantially different sector might result in inaccuracies.
- Only basic variables common to all databases were used in order to ensure continuity between models.
- They were not able to gather data on Service Trip Production and Freight Attraction
- Most of the data collected concerns small businesses. Data from larger establishments (over 100 employees) is hard to obtain.
- The assumption of linking employment to the demand forecasts hinders the explanatory power

Appendix 2

Table 7. Summary of output units and conversion factors.

Metric of Freight and Service Activity (FSA)	Output unit	Conversion Factor to Vehicle-Trips (multiply by)
Freight Trip Generation (FTG) (HARTGEN)	shipments/day	$(\gamma_A \phi_A + \gamma_P \phi_P)$
Freight Trip Attraction (FTA)	deliveries/day	$\phi_A = 1.00$
Freight Trip Production (FTP)	shipments/day	$\phi_P = 0.5$
Service Trip Generation (STG)	-	-
Service Trip Attraction (STA)	service calls/day	$\phi_A = 1.00$
Service Trip Production (STP)	-	-
Freight Generation (FG) (RPI)	-	-
Freight Attraction (FA)	pounds/day	Table 20, 21
Freight Production (FP)	pounds/day	Table 22, 23

ϕ_A and ϕ_P are the conversion factors for attractions and productions of the various metrics of FSA.
 γ_A and γ_P are the ratios of attractions and productions with respect to the total generation.

(Holguín-Veras et al., 2016)

Appendix 3

[The following table] displays the various Mode Choice Factors that affect freight mode choice. Displayed also are the stages in the freight decision making process when each factor comes into play. The stages are:

- Immediate
- Mid-term
- Final

Each decision factor is also given a number between one and nine to further break out when the factor comes into play during the freight mode choice decision (CUTR, n.d., p.54).

Table 4.1 Timing of Mode Choice Factors

Factor 1	Factor 2	Immediate			Mid-Term		Final			
		1	2	3	4	5	6	7	8	9
Total logistics cost	Transportation charges									
	Capital carrying cost in transit									
	Service reliability costs									
Modal characteristics	Trip time and reliability									
Physical attributes of goods	Shipment size									
	Package characteristics									
	Shipment shelf life									
	Shipment value									
	Shipment density									
Flow & spatial distrib. of shpmnts	Distance of shipment									
Firm characteristics	Shippers and receivers situated on rail line									
	Shippers near highway									
	Firms own small trucks									
Flow & spatial distrib. of shpmnts	Shipment frequency									
Modal characteristics	Capacity									
	Equipment availability									
	Handling quality - damage loss reputation									
Total logistics cost	Order and handling costs									
	Loss and damage costs									
	Inventory carrying cost at destination									
	Unavailability of equipment costs									
	Intangible service costs									
Modal characteristics	Customer service									
Firm characteristics	Firm size	Background throughout process								

(CUTR, n.d.)