Validating Calculation of Transport Cost for Freight Carriers on the Last Mile

Conducting a Cluster Analysis on a Case Study in the Municipality of Delft to Enhance the Usability of the Last-Mile Scan Calculation Model

G.A. SCHONEWILLE

ABSTRACT

Growing population, urbanization and rising consumption per capita have resulted in a growing amount of urban freight distribution. Urban freight distribution is not efficient in both economical (due to lack of cooperation and low loading ratios) and environmental terms (heavy and old vehicles driving on conventional fuels) and is a heavy burden on citizens (social burden). Little is known on performance of freight carriers on individual level on economic and environmental aspects. Data availability on urban freight distribution of freight carriers is scarce and research on individual freight carrier's performance is therefore difficult to conduct. ‘Maatwerk Distributie’ has developed a model, called the Last-Mile Scan (LMS) that gathers specific urban freight distribution information summarized in a data sample. A cluster analysis methodology is conducted that statistically generates clusters of similar individual freight carriers on self-chosen key performance indicators. The result of the analysis provides information to government bodies on the urban freight distribution system in their region and to know how (individual and groups of) freight carriers active in their region are performing on economical and environmental aspects.

Keywords: urban freight distribution; transport cost, last mile; multi-actor context; cluster analysis; freight carriers
1. Introduction on urban freight distribution

In scientific literature various definitions on urban freight distribution (UFD) can be found. According to DaBlanc UFD can be defined as: 'a large number of different types of freight flows that cross an urban environment'. Flows can consist out of consumer goods, waste products, building materials, postal mail and others. The term other is not further defined [1]. Muñuzuri et al. define urban logistics as: 'movements of goods that are affected by particularities associated to urban traffic and morphology' [2] [3]. Allen et al. define urban freight distribution as: 'The delivery of consumer goods (not only by retail, but also by other sectors such as manufacturing) in city and suburban areas, including the reverse flow of used goods in terms of clean waste' [4]. In this scientific article the definition of Allen et al. will be used. The reverse flow is out of scope since the model that is analyzed to collect data and calculate transport cost for freight carriers is delineated on urban freight distribution flows only.

Trends that drive UFD

Urban freight distribution demand is growing rapidly the last decades. Trends that drive and push urban freight distribution are growing population, urbanization, global e-commerce growth and dynamic inventory management concepts. The first trend, worldwide population growth, is increasing gradually the last fifty years. In 1961 the world population was around 3 billion. Two years ago, in the year of 2013 the world population grew over 7 billion persons in total [5]. Secondly, last half a century there is a worldwide trend of urbanization happening. From 1960 till 2013 upper middle-income range countries become almost fully urbanized. Now between 40 to 90% of total worldwide populations live in cities [6][5]. The third and more recent trend is the emerging e-commerce sector. The global e-commerce growth is a fact that cannot be denied [7]. The aggregate level of e-commerce is a research field that is not extensively developed, yet. Previous research is mainly focused on individual, firm or country level. Little is known about the aggregate level of e-commerce. Potential drivers of e-commerce are: GDP per capita, geographic and demographic characteristics, urbanization, information infrastructure, cost to shop on the internet, adequacy of economic and financial resources, cosmopolitism, education and human capital [7]. Overall, global e-commerce growth leads to higher demand of freight, mainly in high-density regions, which are generally urban areas. Information technology and other dynamic inventory management concepts have contributed to higher perform abilities of freight carriers. These higher standards have been translated, due to fierce competition among distributors towards higher standards demanded by consumers. The final trend that affects demand of urban freight distribution is the concept of 'just in time' (JIT). This has led to lean manufacturing and a continuously increase in complexity and decrease in stock amounts. Lower stock amounts result in a cost advantage for operators [8].

Problems regarding UFD

Due to an increase in urban freight distribution demand a number of problems in cities occur. In general, UFD problems can be characterized as economic, environmental or social of nature. Urban freight distribution flows occupy around one fourth of the street traffic of an average city (facts derived from surveys made by the Laboratoire d'Economie des Transports (Lyon, France) with the support of the French national research program on Goods in Cities). The amount of urban space that is demanded related to urban freight distribution is even greater. On top of these flows, loading and unloading spots, storage, conditioning and packing
capacity is needed [1]. The amount of greenhouse gas emissions is a problem in urban areas. Most vehicles operating in urban areas are conventional vehicles that still drive on petrol. Livability of an urban area is negatively affected by pollution. Environmental standards are too low to ensure maximum allowed emission polluted within urban boundaries [2]. Economic inefficiency can occur since freight carriers are operating with low loading rates due to lack of critical mass or JIT concepts. Consolidating freight from both supply and demand perspective can increase economic efficiency. The not-at-home problem can be resolved partially as well. As mentioned, lack of space and low loading rates decreases freight carriers operating efficiencies. This leads to a higher amount of vehicles on the road than necessary. Moreover, historic city centers are small and not designed for large amounts of traffic. The existing infrastructure has limited capacity. These factors all together result in congested [2]. Congested roads of course lead to higher inefficiencies. Inefficient bundling from a city perspective is mainly caused due to a lack of cooperation and consultation between receiver and carrier [9]. Transportation flows could be enlarged when actors improve cooperation. Consolidation increases loading rates and lowers the amount of vehicles operating in an urban area. Safety and noise issues regarding old vehicles operating in urban areas have a negative impact on livability of cities. Smaller clean and quiet electric vehicles can reduce the amount of noise and accidents. Another problem that exist mainly in rural areas is the lack of critical mass. These areas do not generate freight flows that are large enough to arrange efficient freight distribution. Distribution in these areas is expensive and inefficient since carriers operate with low loading rates [10].

Sustainable UFD
Sustainable urban freight distribution is a concept that has developed strongly the past years. In a literature review conducted by [11] a structured overview is provided on the origin of sustainable development. In their review the key concepts and principles of sustainable development are introduced and a set of indicators that describes SUFD is developed. Key concepts of sustainable development are two folded. Sustainable development has to meet needs of present generation without abiding to have the ability of future generations to meet their needs [12][11][13]. This definition of sustainable development can be divided in three principles, which include economic growth, social equity and environmental protection [12] (Wolff, 2004). Social equity will meet needs of present generations while the dimension of environmental protection will safeguard to meet needs of present generations and ensures the ability of future generations to meet their needs. In regard to this article the definition and concept of sustainable urban freight distribution is key.

Solutions regarding UFD
Solutions can be compared and analyzed from various approaches. These are physical versus behavioral solutions, private versus public versus private-public partnership, carrier perspective versus receiver perspective (demand versus supply driven) and centralized versus decentralized [4]. An example of a physical orientated solution is the traditional urban consolidation centre (UCC). A UCC is a distribution center specially built to consolidate freight on the border of an urban area. Another solution, which is behavioral oriented and focused on retailers, carriers or both, is consolidation from a retailer/carrier perspective. An example from the retailers’ perspective is that retailers in certain region do not cooperate with each other. Every retailer places an order individually at the moment they...
receive an order from their consumers or when restock of their stores is needed. This demand is driven by the needs of the consumer. Changes on their behavior by increasing cooperation among retailers to synchronize orders can contribute to optimized freight flows [9]. The organizational structure of different initiatives can generally be divided into private owned/operated and public-private partnerships. An example of a PPP is a traditional UCC subsidized by a government body. The UCC is generally operated by a private party but cannot exist without procurement.

2. The Last-Mile

From a supply chain perspective the last section of the supply chain is in literature known as the last mile. This section of the supply chain is for most freight carriers a section of the supply chain that can be considered problematic. The last mile challenge has risen in the last ten years since an explosion in consumer deliveries such as delivery in grocery, office supply, packages and pharmaceutical products [14]. In general, the last mile is regarded as more expensive, less efficient and more polluting compared to the entire logistic chain. A frequently occurring issue of the last mile delivery is the not-at-home problem. This issue results in high delivery failure and an increase in empty trip rates. This has negative impacts on cost efficiency and environmental performance [10].

The Last-Mile Scan

In 2009 Connekt initiated a workgroup with various stakeholders from the transportation sector. In the workgroup different interest groups such as ‘Transport en Logistiek Nederland’ (TLN) and ‘Ondernemersorganisatie voor logistiek en transport’ (EVO) and various municipalities as Delft, Arnhem, Nijmegen, Utrecht, Amsterdam and Breda participated. Derived from this working group the initiative ‘Maatwerk Distributie’ (MD) has been established. MD is procured and owned by the municipality of Delft. Dhr. B. Coremans, employed by the Municipality of Delft, is project manager of MD and developed the LMS to a great extent in participation with some other parties like Buck Consultants International (BCI), Panteia and municipality Delft. The model is eventually called the ‘Last Mile Scan’.

The LMS uses formulas, assumptions, indicators and tools to conduct calculations. Outcome of the model consist of dimensions’ distance, time and cost. These three dimensions are used throughout the entire model. The general result of the LMS is twofold. Firstly, the LMS estimates transport cost of freight carries on the last mile based on their characteristics, expressed in euros. Secondly, the LMS calculates the probability of saving. This probability is based on the outcome of the LMS (estimation of transport cost) compared to an alternative situation using a central urban freight distributor with competitive tariffs. Information on tariffs is received from PostNL.

In order to give freight carriers’ insight in the last mile cost structure, knowledge and characteristics of the shipment is needed. Variables such as vehicle type, origin, destination, number of stops, number of small packages, number of large packages, place of activity (‘in’ and/or ‘outside city center’), active during peak hours and number of days active per week is needed and used as input variables. Besides, a various number of assumptions are made to calculate the cost of the last mile and to make assumptions on the probability of saving money.

Purpose of development of LMS

The lack of information on the cost structure of the last mile of the supply chain can be identified as a knowledge gap. The desired situation is a full and accurate understanding of the cost structure. From the perspective of the municipality of Delft this line of reasoning can be drawn a bit further.
They do not only like to know more about transport cost but would also like to have an indication on which carriers are operating in a sustainable and society responsible matter. This information and knowledge could help in bridging differences of involved stakeholders and could potentially benefit in finding common ground. Collaboration and information sharing among stakeholders is an important factor that could contribute to improve proper functioning of the UFD system and to achieve higher sustainability, more economic gains and other social and environmental benefits.

**Case: Municipality of Delft**
In case of the municipality of Delft several of foregoing problems occur. Delft is an old town with a historic city centre that is rather small of size. Lack of space and congestion are definitely problems that apply since distribution with large trucks is difficult and therefor have a larger impact on citizens due to their relatively high disturbance compared to urban areas with more space. Greenhouse emissions are definitely a problem since old conventional vehicles have a negative impact on the liveability of the municipality. Also in Delft the economic efficiency is not fully utilized due to a large number of independent operators who do not communicate or collaborate to create win-win opportunities. Due to this lack of cooperation and communication between receiver and carrier, receiver and receiver and carrier and carrier potential efficient bundling is not yet achieved. Furthermore, the truck fleet that is active in the urban area of Haaglanden could be improved. In 2015 there are still many old vehicles active that have a large negative impact on the environment and cause nuisance and safety issues [15]. Moving towards smaller clean and quiet vehicles should be further emphasized.

3. **Methodology cluster analysis**
The identification of groups of individuals that are similar to each other but different from other individuals can be of added value to companies to increase profits or to gather knowledge for scientific research purposes. Clustering individual freight carriers in categories can increase understanding in what drives them. A cluster is a group of relatively homogenous cases or observations [16]. The dependence and segregation of clusters is realized on the basis of different attributes and characteristics. Parameters of the LMS model can serve as a possibility on how to cluster individual carriers. Furthermore, these clusters can be labeled and weighted on important criteria or key performance indicators (KPI's) to measure performance of urban freight distribution on economical and environmental aspects and even social impacts on a region.

The result of the cluster analysis gives insight to government bodies in general, or in scope of the case study, to the municipality of Delft on freight carriers active in their region. This information could give them knowledge on what freight carriers contribute in sustaining the UFD system and how freight carriers perform from an environmental and societal point of view. This knowledge gives the municipality of Delft a starting point to develop custom made policy suitable for the region or on a detailed carrier level.

**Data sample**
All data collected by MD is used in the analysis. A total of 267 scans are collected over six regions. All types of freight carriers are present and thereby all types of shipments. This data is used as input for the cluster analysis. Distinction of freight carriers can be based on attributes or KPIs. The attributes that can be used to distinguish clusters are given below. Most of them are already present in the LMS model. Some KPIs are not yet
present and can be calculated by conducting a few simple mutations. KPIs that are not yet present in the LMS model are cost parameters: cost per unit/minute/km, input parameter total load (package plus pallet) and footprints. These KPIs are used as criteria to see how well a cluster (group of freight carriers that are similar) performs on different KPIs.

**Software package SPSS**

Software package SPSS can statistically generate clusters on the base of selected attributes. After these clusters are formed, they provide a value on selected evaluation fields. Evaluation fields or criteria are in this case KPIs. This line of reasoning can also be reversed since clusters can be created on the base of KPIs. Therefore, attributes can also be evaluation fields or KPIs. A cluster can be developed on the base of key performance indicator cost per unit/minute/kilometer (economical) and/or footprints (environmental). In general, the possibilities of creating clusters and analyzing data are endless.

**Research question**

This scientific article tries to answer the following research question:

*How can the results and data collected by the LMS model be used to gain more information and knowledge on clusters of freight carriers with similar characteristics?*

Figure 1 captures the framework and design of a cluster methodology analysis. This framework is applied to specific data collected by the LMS model and uses available parameters and variables if possible.

---

**Choose attributes/inputs**

Choose evaluation fields

Apply cluster analysis

- Distinction on the base of cost per trip;
- Distinction on the base of time per trip;
- Distinction on the base of distance per trip;
- Distinction on the base of load (package);
- Distinction on the base of load (pallets);
- Distinction on the base of total load (package plus pallets);
- Distinction on the base of number of stops;
- Distinction on the base of vehicle type;
- Distinction on the base of cost per unit;
- Distinction on the base of cost per minute;
- Distinction on the base of cost per kilometre;
- Distinction on the base of footprints;
- Statistical distinction on the base of chosen attributes/inputs conducted by SPSS

**Figure 1 - Design of cluster analysis methodology**
Results

Formulated attributes and evaluation fields need to correspond with parameters present in the LMS model. Parameters can be used directly when they are present as a variable or can be created by means of a simple calculation. All evaluation fields (key performance indicators) are calculated by a simple mutation. Table 1 gives an overview of which parameters of the LMS match which distinction base and what mutations is conducted.

A freight carrier can be classified and evaluated on parameters represented in Table 1. A category is based on one or more input variables (attributes). Interesting to explore is which freight carriers perform well on economical and environmental aspects. Economical aspects correspond with KPIs cost per unit, cost per minute and cost per kilometre. These KPIs can be calculated by conducting a few simple mutations as mentioned in Table 1.

Environmental aspects can be added to the LMS model by calculating footprints of trips with aid of an emission table. Footprints can be calculated by using output variable distance multiplied by the value derived from the emission table (based on type of vehicle and activity in peak hour).

Cluster analysis ( economical aspects)

A first cluster analysis is conducted focused on economical aspects only. The cluster analysis is based on inputs and evaluation fields presented in Figure 2.

Table 1 - Attributes of freight carriers and corresponding parameters LMS model

<table>
<thead>
<tr>
<th>Distinction base</th>
<th>Parameters LMS model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attribute/ input variables</td>
<td></td>
</tr>
<tr>
<td>Distinction on the base of cost per trip</td>
<td>Output variable cost</td>
</tr>
<tr>
<td>Distinction on the base of time per trip</td>
<td>Output variable time</td>
</tr>
<tr>
<td>Distinction on the base of distance per trip</td>
<td>Output variable distance</td>
</tr>
<tr>
<td>Distinction on the base of load</td>
<td># of pallets/packages</td>
</tr>
<tr>
<td>Distinction on the base of number of stops</td>
<td># of stops</td>
</tr>
<tr>
<td>Distinction on the base of vehicle type</td>
<td>Type of vehicle</td>
</tr>
<tr>
<td>Key Performance Indicator</td>
<td></td>
</tr>
<tr>
<td>Distinction on the base of cost per unit</td>
<td>Output variable cost divided by # of pallets/packages</td>
</tr>
<tr>
<td>Distinction on the base of cost per minute</td>
<td>Output variable cost divided by output variable time</td>
</tr>
<tr>
<td>Distinction on the base of cost per km</td>
<td>Output variable cost divided by output variable distance</td>
</tr>
<tr>
<td>Distinction on the base of footprints</td>
<td>Footprints calculated by output variable distance; type of vehicle; emission table</td>
</tr>
</tbody>
</table>

Figure 2 - Cluster analysis economic aspects, chosen inputs and evaluation fields

As setting a minimum number of 4 clusters is chosen. Each cluster can be labeled based on the evaluation fields (performance) that are economical of nature. The clusters are generated on the basis of shipping load (number of packages and pallets) and number of stops.

Attributes/inputs:
- Number of package
- Number of pallets
- Number of stops

Evaluation fields:
- Cost per unit
- Cost per minute
- Cost per kilometer
The model summary of the cluster analysis is presented in Figure 3. A TwoStep algorithm has been used. In total three inputs created a total of four clusters. Furthermore, the cluster quality is around 0.6, which is rather good.

Cluster one consists of a relative small shipping load (1,23 packages and 2,27 pallets) and a few stops (2,21) and is the largest cluster (73.5% of the dataset, 197 cases). Cluster three has a medium size shipping load (12,46 packages and 1,46 pallets) and a medium amount of stops (6,38). In total 37 cases are included in this cluster. The second cluster has a medium size shipping load (3,04 packages and 20,42 pallets) and a high amount of stops (11,33). In total 24 cases are included in this cluster. The fourth and smallest cluster has a relatively large shipping load (43,60 packages and 3,40 pallets) and consist of many stops 18,30). In total 10 cases are included in this cluster.

Figure 3 - Model summary of cluster analysis on economic aspects

Clustering

Clusters

<table>
<thead>
<tr>
<th>Cluster</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Inefficient</td>
<td>Medium efficient</td>
<td>Efficient</td>
<td>Efficient</td>
</tr>
<tr>
<td>Description</td>
<td>Small load Little stops</td>
<td>Medium load * little pallets Medium stops</td>
<td>Medium load * many pallets Many stops</td>
<td>Large load Many stops</td>
</tr>
<tr>
<td>Size</td>
<td>73.5% (197)</td>
<td>13.8% (37)</td>
<td>9.0% (24)</td>
<td>3.7% (10)</td>
</tr>
<tr>
<td>Inputs</td>
<td>Package 3.23</td>
<td>Package 12.46</td>
<td>Package 3.09</td>
<td>Package 43.60</td>
</tr>
<tr>
<td></td>
<td>Pallets 2.27</td>
<td>Pallets 1.46</td>
<td>Pallets 20.42</td>
<td>Pallets 3.40</td>
</tr>
<tr>
<td></td>
<td>Stops 2.21</td>
<td>Stops 6.38</td>
<td>Stops 11.33</td>
<td>Stops 18.30</td>
</tr>
<tr>
<td>Evaluation fields</td>
<td>CostPerUnit 12.17</td>
<td>CostPerUnit 3.10</td>
<td>CostPerUnit 4.25</td>
<td>CostPerUnit 1.45</td>
</tr>
<tr>
<td></td>
<td>CostPerKm 8.05</td>
<td>CostPerKm 9.11</td>
<td>CostPerKm 12.31</td>
<td>CostPerKm 0.34</td>
</tr>
<tr>
<td></td>
<td>CostPerMin 0.57</td>
<td>CostPerMin 0.38</td>
<td>CostPerMin 0.21</td>
<td>CostPerMin 0.36</td>
</tr>
</tbody>
</table>

Figure 4 - Cluster analysis results on economic aspects
The first cluster can be labeled as inefficient because it scores relatively bad on evaluation fields cost per minute and cost per unit. Variables related to time are most important variables because they have the greatest influence on cost. Comparing or valuating clusters on evaluation fields cost per unit is difficult because shipment consist of pallets and/or packages that differ in cost, time and effort. This is confirmed by looking at clusters 3 and 4 (both large amount of packages). These clusters have a relatively low score on evaluation field cost per unit due to a shipping load with relatively many packages.

Furthermore, it can be concluded that the first cluster is the largest. This means that from the entire dataset, which are actually real participants active in the urban area of Haaglanden, the majority is conducting activities relatively inefficient. For the entire cluster (or freight carriers) it is beneficial to consider a switch to a central distributor. A first classification framework on economical aspects can be made. Classes are distinguished on the basis of key performance indicators cost per unit/minute/kilometer. A total of four classes are created. The classification typology is described according the following labels:

- Inefficient (relatively small load and relatively little stops);
- Medium efficient (relatively medium load, relatively many stops);
- Efficient (relatively medium load, relatively medium stops);
- Efficient (relatively large load, relatively many stops);

Cluster analysis (environmental aspects)
A second cluster analysis is conducted focused on environmental aspects only. The cluster analysis is based on inputs and evaluation fields presented in Figure 6.

Figure 6 – Cluster analysis environmental aspects, chosen inputs and evaluation fields

As setting a minimum number of 4 clusters is chosen. Each cluster can be labeled based on the evaluation fields (performance) that are environmental of nature. The clusters are generated on the basis of cost of a trip, number of pallets, number of packages and number of stops.

Attributes/inputs:
- Cost of a trip
- Number of package
- Number of pallets
- Number of stops

Evaluation fields:
- Cost per unit
- Cost per minute
- Cost per kilometer

<table>
<thead>
<tr>
<th>Label</th>
<th>Dimension/input values</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inefficient</td>
<td>Small load, little stops</td>
<td>Single day fly</td>
</tr>
<tr>
<td>Medium efficient</td>
<td>Medium load, many stops</td>
<td>HoReCa</td>
</tr>
<tr>
<td>Efficient</td>
<td>Medium load, medium stops</td>
<td>Supply of supermarket</td>
</tr>
<tr>
<td>Efficient</td>
<td>Large load, many stops</td>
<td>E-commerce</td>
</tr>
</tbody>
</table>
The cluster analysis has used a TwoStep algorithm. In total four inputs created a total of four clusters. Furthermore, the cluster quality is around 0.55, which is rather good.

Cluster number two can be labelled as a relatively medium pollutant. This group has a small shipping load (2.34 packages and 1.47 pallets), only a few stops (2,12) and is relatively cheap (18.89). Cluster number one can be labelled as a large pollutant. This group has a medium-shipping load (2.54 packages and 4.00 pallets), a medium amount of stops (4,57) and is relatively expensive (58.02). Cluster number three can be labelled as a small pollutant. This group has a large load (28.22 packages and 4.83 pallets), many stops (16.00) and is average costly (51.67). Cluster number four is a small pollutant. This group has a large shipping load (0.17 packages and 30 pallets), a medium amount of stops (4.92) and is relatively expensive (76.63).
4. Conclusion

In this section the research question of this article is answered. The research question was:

*How can the results and data collected by the LMS model be used to gain more information and knowledge on clusters of freight carriers with similar characteristics?*

Distributors can be classified on the basis of characteristics or on performance levels. In the cluster analysis characteristics are translated into attributes/inputs and performance levels into key performance indicators. To be able to filter the data sample these characteristics need to match parameters or results of the LMS model. Distinction can be based on parameters present in the LMS model.

A cluster analysis can give valuable knowledge on how freight carriers are performing on key performance indicators. Depending on the objective and needs for specific information the right settings can be chosen. Settings can vary on economic and environmental aspects or can even be combined.

The practical usability of the cluster analysis is three-folded. First, the analysis gathers detailed knowledge on already available data derived by the LMS model for both participants (freight carriers) and government bodies. Secondly, this knowledge can be used to develop custom policy for specific freight carriers that are performing undersize. Thirdly, the results can contribute in mediating cooperation between involved freight carriers because cluster analysis provides them with specific knowledge on their performance on economical and environmental indicators.

The conclusion drawn from the cluster analysis is that relatively smaller shipments perform less good in both economical and environmental terms and can better outsource their activities and use an alternative shipping method.
5. Bibliography


