Environmental considerations in supply chain decisions: Proposing a new methodology for locating cross-docking terminals considering CO_2e emissions*

Etiënne Albers (s2618028), e.h.m.albers@student.utwente.nl

Faculty of Behavioural, Management and Social Sciences, University of Twente
Supervisors: dr. P.B. Rogetzer & dr. ir. E.A. Lalla
Company supervisor: N. Prins

Business Administration, Digital Business and Analytics
Master's Thesis

July 12th, 2022

** From the beginning, the main goal of this research was to propose and apply a new methodology for locating cross-docking terminals in a supply network considering the emissions of CO_2e. In the first phase, the results of the methodology seemed promising, with a reduction of CO_2e as high as 23.79% compared to the current supply network in use. During the completion of the research, as a result of a careful examination of the CO_2e figures provided by the organization, it turned out that the CO_2e figures provided describing the current supply network were calculated in a different way as initially given. Consequently, the previously mentioned reduction of 23.79% was based on a different calculation method as the one used for CO_2e calculations for the proposed methodology. After re-calculating the emissions for the current supply network with the same method, emissions in the supply network proposed by the methodology turned out to be 50.13% higher as compared to the current situation. This influenced the contribution of the research. However, the research still contains many meaningful insights regarding the use of clustering and the centre of gravity and how these methods can be used when considering CO_2e emissions in supply network decisions.*

Environmental considerations in supply chain decisions: Proposing a new methodology for locating cross-docking terminals considering CO_2e emissions.

Etiënne Albers (s2618028), e.h.m.albers@student.utwente.nl

Faculty of Behavioural, Management and Social Sciences, University of Twente
Supervisors: dr. P.B. Rogetzer & dr. ir. E.A. Lalla

KEYWORDS

Environmental considerations
Supply chain decisions
Cross-docking terminals
Facility location
Hub location
Centre of gravity
K-means clustering

ABSTRACT

Goal: The goal of this research is to propose and test a new methodology that considers the emission of CO_2e during transportation in determining the locations of cross docking terminals in a supply network. Simultaneously, the research aims to show how clustering and the centre of gravity can be used when considering CO_2e emissions in supply chain decisions. **Design/Methodology/Approach:** K-Means clustering is used to cluster suppliers based on geographical location. By applying the centre of gravity twice, the weight of supply and demand points are included. The amount of CO_2e emitted is estimated using Proxio, following an activity-based approach. **Findings:** The methodology proposed suggests using four cross-docking terminals, being the lowest number of terminals with the highest effect. The CO_2e emissions in the proposed network are 50.13% higher compared to the current situation where 8 cross-docking terminals are used. The result shows that the proposed method is not successful in the case of Scania. Consequently, Scania should not change their supply network accordingly. Instead, the organization should focus on the use of sustainable transportation in their current network and consider a more advanced method for analysing their entire supply network. **Practical Implications:** The research is limited by the fact that only the organizations own shipments are considered, resulting in sub-optimization. In addition, a continuous approach is used which may cause unrealistic locations for cross docking terminals. **Originality/value**: This paper is especially valuable for practitioners in the field of supply network design. Besides, it contributes to the literature by proposing and testing an alternative methodology to the hub location problem.

1. Introduction

Organizing a supply network involves decisions on strategic, tactical and operational level (Ivanov, 2010). Strategic decisions tend to be made over a longer time horizon, while tactical and operational decisions occur at a daily basis. However, the operational and tactical decisions in a supply network are influenced by strategic decisions regarding the overall layout of the network. So, strategic decisions affect the day-to-day operation.

One of the consequences of the daily operations that is increasingly drawing attention is the emission of greenhouse gases during transport. According to Cristea (2013), 33% of world-wide trade-related related emissions are attributable to international transport. By efficiently organizing transportation, an organization can reduce these emissions.

One way of increasing transport efficiency is by using a cross-docking terminal (Vogt, 2010). This efficiency, however, is influenced by the location of the cross-docking terminal within the supply network and determining a location for one or multiple cross-docking terminals is a subject vividly discussed in scientific literature.

Despite the wide range of publications on the subject, there does not seem to be a universal method that applies to every situation. Proposed methodologies within the literature are tailored to specific situations, only consider costs or are mathematically complex. In addition, methods currently available regularly consider conditions that are irrelevant in the case of carbon emissions or too specific for the situation, such as opening hours of the cross-docking terminal or the amount of docks available (Campbell, 2005).

In this research, a new approach for determining a location for a cross-docking terminal is proposed. The proposed method is applied at an international manufacturer in the automotive industry. The method considers CO_2e emissions and combines k-means clustering and the centre of gravity in a novel way to tackle the hub location problem. Simultaneously, the research aims to show how historical data can be used when considering CO_2e emissions in supply chain decisions.

Opposite to many of the methods currently available in scientific literature regarding hub location problems, the method proposed in this research aims to be accessible for practitioners in the field of supply network design employed at the organization where the research is conducted. The intermediate steps in the method provide meaningful insight in how suppliers and drop-off locations influence the locations of cross-docking terminals and therefore impact the design of a supply network.

This paper is structured as follows. The rest of the introduction is used to describe Scania, the organization where the proposed methodology is applied. Section 2 is dedicated to the literature, in which facility location problems, CO_2e and environmental considerations in supply chain decisions are addressed in detail. Section 3 describes the proposed methodology and provides more details on the data used. Results are discussed in Section 4. The final section, Section 5, contains the conclusion, as well as practical and theoretical contributions and suggestions for further research.

1.2 Company description

The methodology proposed in this paper is applied at Scania Logistics Netherlands (SLN). SLN is responsible for the entire supply network of inbound supplies for Scania, a Swedish manufacturer of trucks, buses, and power solutions (engines). Scania holds a market share of 15.5% for trucks in Europe and sold 85,930 trucks worldwide in 2021, making it one of the biggest companies active in the truck market (Scania, 2021b).

Scania has multiple production units worldwide, namely in Södertälje (Sweden), Zwolle (Netherlands), Angers (France), Słupsk (Poland) and São Bernardo (Brazil). Besides the production units for final products, Scania has several facilities for the production of components, such as cabins (in Oskarshamn, Sweden) and frame components (in Luleå, Sweden). These components are thereafter transported to the aforementioned production units. At Scania, the processes within a production unit mainly involve the assembly of parts and components. In addition to the production units, there are warehouses run by Scania for distribution and storage of parts and components, such as Scania's logistics centre in Hasselt (Netherlands).

Scania distinguishes between parts, such as bolts, tires, and gas tanks originating from suppliers, and components, such as engines and cabins, originating from Scania's own production units. However, to produce those components, parts from suppliers are still required. The distinction

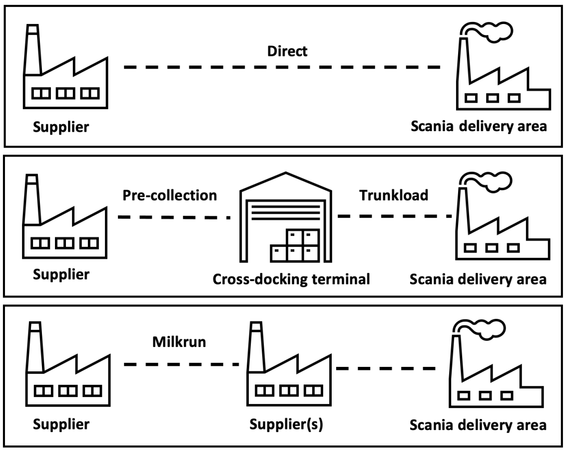

Figure 1. Three different supply methods used by SLN for transportation of parts between suppliers and Scania delivery areas.

between parts and components is important because the organization has a lot of influence on the production at their own facilities as compared to the production of parts and suppliers, where their influence is limited. Providing all production units with parts and components in an efficient way on this scale is challenging and therefore requires an advanced supply network.

1.3 Scania's Supply Network

Scania's supply network consists of around 1,000 direct suppliers (Scania, 2021b). Parts from these suppliers are transported to Scania's production

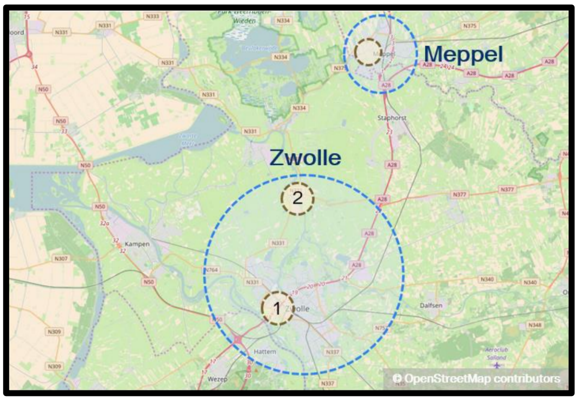

Figure 2. Delivery area Zwolle with two delivery clusters, Meppel and Zwolle (dotted circles). Within the Zwolle cluster, two unloading areas are distinguished, Zwolle and Hasselt. (Source: Scania (2021a)).

units in three different ways: via a direct run (1), which is a transport from the supplier to a production unit without any stopover, via pre-collection (2), which entails transport from a supplier to a cross-docking terminal where parts with the same destination from different suppliers will be combined and transported towards a production unit, and via a milkrun (3), where a truck visits two or more suppliers to collect smaller batches of goods and eventually depart towards a production unit with a full trailer. Figure 1 contains an overview of the different types of transportation.

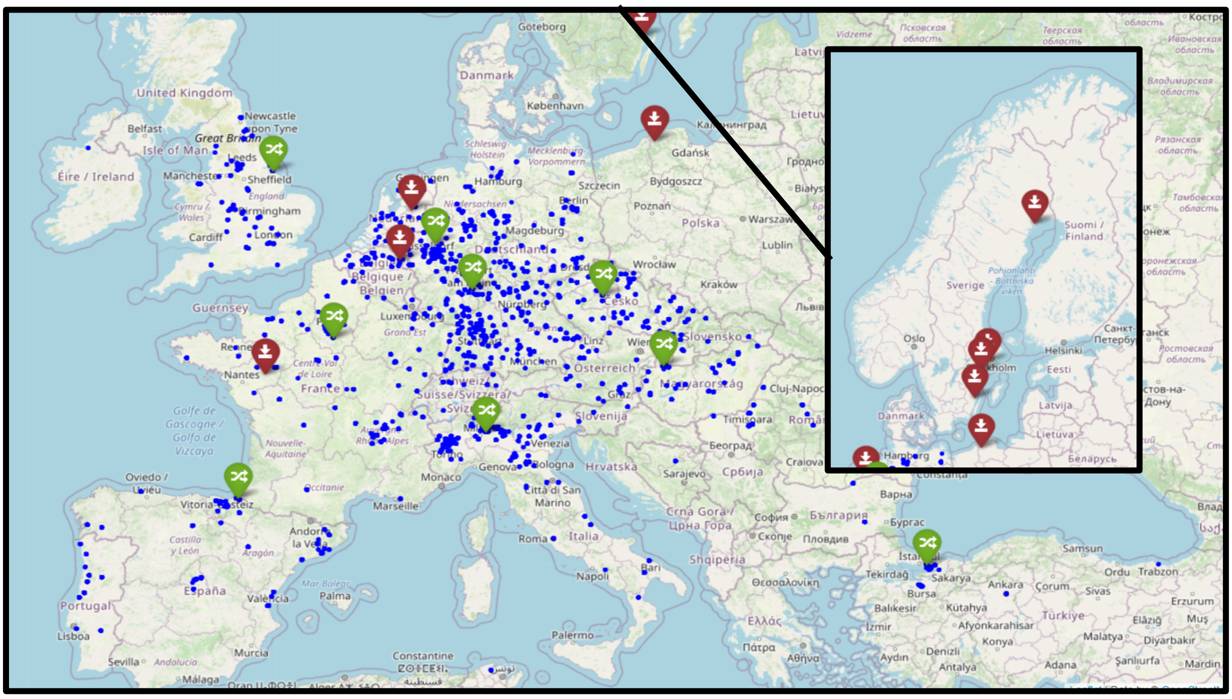

Figure 3. Current situation in Central Europe. Suppliers (blue, dots), cross docking terminals (green marker with shuffle-arrows) and delivery areas (red marker with arrow pointing down) that are part of Scania's supply network. Map: ©OpenStreetMap-contributors.

The production units in Scania's supply network are part of a hierarchical delivery structure that Scania created to manage the incoming supplies. Delivery areas are the biggest granulation in this hierarchical structure. Within a delivery area, multiple delivery clusters can occur. Similarly, a delivery cluster can contain multiple (un)loading areas, with multiple (un)loading places in one area. Figure 2 provides an overview of this hierarchy. As an example, the delivery area Zwolle contains two clusters, Zwolle and Meppel. Within the Zwolle cluster, there are two unloading areas, e.g., the production unit in Zwolle and the logistics centre in Hasselt. Finally, there are multiple unloading places within an unloading area. Separate docks dedicated to the unloading of cabins and gearboxes are an example of this.

The delivery areas together with the cross-docking terminals and suppliers make the supply network that SLN is responsible for. The case study is limited to Scania's suppliers located in central Europe. Suppliers located in Scandinavia, the Baltic states and Poland are therefore not included in the case study. Suppliers in central Europe producing parts for delivery areas in these areas are included, hence the delivery areas in this region. See Figure 3 for the geographical area that SLN is responsible for, including the suppliers, delivery areas and cross-docking terminals.

1.4 Problem description at Scania

The transports taking place within Scania's supply network are outsourced to a variety of different carriers. SLN concludes a contract with a carrier for every existing route in its supply network for a predetermined period, commonly several years. This allows the organization to be flexible in the long term while creating stability in the short term. After the contract period expired, SLN launches a tender for the corresponding route, which enables the organization to select a new carrier that best meets the current requirements. Thus, tenders allow SLN to change and therefore re-design Scania's supply network.

As is the case with the transports, the cross-docking terminals together with the material handling processes within the terminal are also outsourced. SLN concludes a contract with an operator of a cross-docking terminal for a predetermined period. During this period, Scania can send shipments via the contracted cross-docking terminal, consolidate(s) them, and send(s) them forward to their delivery areas. The main advantage of this construction is the ability to choose a cross-docking terminal that is strategically located in Scania's supply network, which is constantly changing.

The main factor causing change in Scania's supply network is the selection of suppliers. Scania is aware of the trend that suppliers increasingly come from Eastern Europe. As a result, parts originate more frequently and in higher volume from suppliers located further away from Scania's delivery areas. Consequently, the location of a cross-docking terminal can become geographically inefficient when the assigned suppliers are located elsewhere.

Scania aims to avoid transport of small batches of parts over long distances. Instead, consolidating at a cross-docking terminal and sending a full truck load towards a delivery area is the preferred approach. Since some of the delivery areas are producing the same product, parts from a supplier can be sent towards different delivery areas. This increases the importance of a cross docking terminal being present in the supply network, and simultaneously increases the complexity of the network. Suppliers are located in Eastern Europe irregularly, and one at a time. Also, transports between suppliers and cross-docking terminals and between cross-docking terminals and delivery areas are contracted one at a time. As a result, a cross-docking terminal can have different, and possibly less, suppliers assigned that are located further away. What once was an efficient or optimal location, can become inefficient or obsolete over time.

Within SLN the question has raised to what extent the current cross-docking terminals are efficient in terms of their location within the supply network. This question is reinforced by the increasing attention paid by SLN to the emission of greenhouse gases during transportation. Scania wants to have not only a supply network that is cost efficient, but also aims at reducing their direct and indirect emissions as much as possible while simultaneously meeting quality standards e.g., on time delivery.

In 2021, Scania emitted 100,080,000,000 Kg of CO_2e (Sum of scope 1, 2 and 3 emissions) (Scania, 2021b). More information on the emission of CO_2e

is provided in Section 2. Emissions in Scania's supply network during inbound transportation total 68,010,280 Kg of CO_2e in 2021. 18,064,995 Kg CO_2e is emitted during transportations where a cross-docking terminal is involved. This 18,064,995 Kg of CO_2e is the sum of 10,619,304 Kg CO_2e that is emitted during pre-collection and 7,445,691 Kg of CO_2e that is emitted during transportation from cross-docking terminals towards Scania delivery areas.

Currently, when a contract with a cross-docking terminal expires, a new tender is presented considering the suppliers that are currently using the associated cross-docking terminal. This causes sub-optimization, because suppliers cannot be assigned to a cross-docking terminal other than the one destined for the whole group of suppliers. Re-considering every supplier-cross-docking terminal combination every time that a supplier is located elsewhere would result in an unmanageable amount of work during the tender period and is therefore not executed.

In summary, SLN is constantly evaluating Scania's supply network, aiming at a higher efficiency in terms of costs, quality, and emissions. The cross-docking terminals that Scania is currently using are an important factor regarding the efficiency of the supply network, but the organization lacks insights in how the locations of the cross-docking terminals influence the performance in terms of CO_2e emissions. More specifically, the organization wants to know what the most optimal locations for cross-docking terminals are in their supply network and how their historical transport data can be used for such strategic questions.

2. Literature Review

2.1 Cross-docking

Various definitions of cross-docking are present within scientific literature. According to Vogt (2010), there is no generally accepted definition of cross-docking within a supply. In fact, Vogt (2010) states that there is a lack of formal taxonomy for all the different types of cross-docking in a supply network. Akkerman et al. (2022) state that cross-docking is typically defined as "the process of consolidating less-than-truckload (LTL) shipments with the same destination to full truckloads (FTL), with the additional trait that products are stored

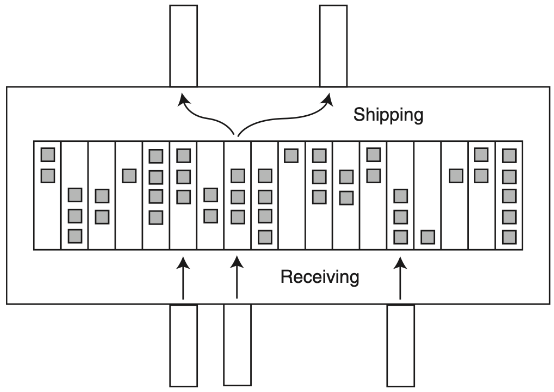

Figure 4. The layout of a single-stage cross-docking terminal. Source: (Gue & Kang, 2001).

up to a maximum of 24 hours (Boysen & Fliedner, 2010)" (Akkerman et al., 2022, p. 71).

Within the literature, there is a distinction between the concept, often referred to as *cross-docking* and the physical building itself, most referred to as *cross-docking terminal* (Stephan & Boysen, 2011) or *cross- docking facility* (Magableh et al., 2005).

Inside a cross-docking terminal, two different layouts can be distinguished: single-stage and two-stage (Gue & Kang, 2001). The input and output of these different layouts are the same, but the material handling in the cross-dock is different. Terminals with a single-stage layout unload a trailer after arrival and sort the incoming goods just before loading the goods into the outbound trailer, as shown in Figure 4.

Terminals with a two-stage layout unload a truck and sort the incoming goods in a dedicated area, after which the sorted goods are loaded into the outbound trailers, as shown in Figure 5.

Scientific literature contains a rich number of publications regarding the operational problems inside a cross-docking terminal, mostly focussing on efficiency improvement. Comprehensively discussing the internal efficiency of cross-docking terminals is out of the scope of this paper. For an overview of the available literature on this topic, see the works of Agustina et al. (2010) and Ladier and Alpan (2016).

Using a cross-docking terminal in a supply network has multiple benefits. Foremostly, transportation costs can be reduced because goods can move through the distribution network in Full Truck Loads (FTL) more often (Apte & Viswanathan,

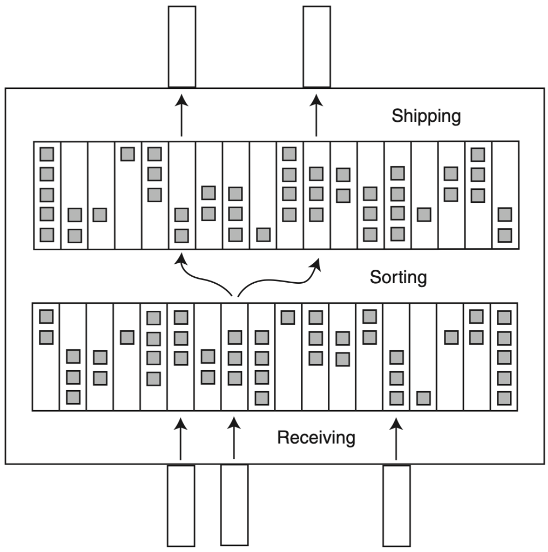

Figure 5. The layout of cross-docking terminal applying a two-stage design. Source: (Gue & Kang, 2001).

2000), since Less Than Truck Loads (LTL) are consolidated (Agustina et al., 2010). Other benefits of using a cross-docking terminal, as compared with traditional warehouses, are (1) reduced delivery lead-times, (2) reduced stock and (3) a reduced risk for loss and damage of stored goods (Van Belle et al., 2012).

The use of cross-docking is preferred in situations where demand is fairly constant (Apte & Viswanathan, 2000). Successful applications of cross-docking mentioned are those of grocery store Walmart (Van Belle et al., 2012), car-manufacturer Toyota and mailing company UPS (Larioui et al., 2017). According to Kreng (2008), the use of a cross-docking distribution strategy in a supply network can result in 'tremendous' savings compared to a supply network using standard distribution centres.

Three tactical decisions required for establishing an efficient supply network using cross-docking terminals are *location, transportation* and *consolidation,* as formulated by Gümüs and Bookbinder (2004). The former of these three will be discussed in more detail, the latter two are of equal importance but are out of scope of this research. Besides the location of a cross-docking terminal, the number of terminals is also a factor of influence. Determining the location of a cross-docking terminal is relatively simple when a single terminal has to be located, but complexity arises when multiple terminals are concerned (Stephan & Boysen, 2011).

2.2 Locating facilities in distribution- or supply networks

Strategically locating a facility in a distribution- or supply network is a familiar challenge within operational research. This challenge results in an extensive body of publications addressing the challenge, each with slightly different approaches or assumptions. Klose and Drexl (2005) comprehensively cover a variety of different types of models available. A detailed description of all available models regarding the facility location problem falls outside the scope of this paper. Similar to Klose and Drexl, Arabani and Farahani (2012), provide a broad overview of the different models and classify them accordingly.

One of the classifications by Arabani and Farahani (2012) is whether facility location problems (FLPs) are continuous or discrete. In continuous models a facility can be located everywhere inside the planning area, while facilities can only be placed in predetermined (fixed) points when a discrete model is concerned. Using a continuous approach can result in proposed locations that are in unrealistic areas, such as an ocean or nature reserve. However, this does not mean that continuous approaches are useless. As mentioned by Drezner (1994), it could be that the most optimal location is not within the set of predetermined (fixed) points. Besides, discrete and continuous approaches can be combined to find an optimal location, as mentioned by Adeleke and Olukanni (2020). Facilities close to the proposed points of the continuous approach can be used as input for an analysis with a discrete approach.

Another classification that Arabani and Farahani (2012) distinguish within all the facility location problems are *hub-location problems.* Hub location problems (HLPs) focus on the location of hubs. Such hubs are best defined by Alumur and Kara (2008) as " special facilities that serve as switching, transshipment and sorting points in many-to-many distribution systems" (P. 1). Such a hub can be considered a cross docking terminal, hence the definition of Kinnear (1997) as provided in Section 2.

Publications on hub location problems are manifold. Over a 100 hub-location studies have been conducted up until 2006 (Alumur & Kara, 2008). Hub-location problems occur in a variety of disciplines, such as air-transportation and telecommunication (Alumur & Kara, 2008). Publications focusing on the location of a hub, such as Campbell (2005) are using mathematical single-objective optimization models. Most of these models are focused on finding an optimal solution given a set of constraints. These constraints consider a variety of different variables, such as opening hours of the hub and a hubs capacity.

In a broader sense, this is a trend that also appears in other publications regarding the facility location problem: models are mathematically complex, or only consider costs. This is recognized in the work of Treitl and Jammernegg (2014), who state that models used to solve the facility location problem are formulated as (non)-linear problems or mixed integer problems solving for costs.

2.3 Environmental considerations

There are a few exceptions to the above. Chen and Zhang (2018) developed a model that considered carbon emissions in finding the optimal location for a cross-docking terminal. Here too, the authors note that little attention is paid to emissions during transport within the existing literature. Treitl and Jammernegg (2014) also developed a model that considers carbon emissions in facility location problems, but the researchers recon that the majority of the publications on this topic leave environmental consequences out of their scope.

McKinnon and Piecyk (2010), distinguish two different approaches for the estimation of CO_2 during freight transportation, the *energy-based approach* and the *activity-based approach.* The energy-based approach is based on the exact amount of energy consumed, multiplied by an emission factor that converts the energy into CO_2. This is only possible if the exact emissions are known, and as pointed out by McKinnon and Piecyk (2010), this is often not the case because transport operations are likely to be outsourced. As an alternative, the activity-based approach can be used. The activity-based approach is a rough estimation of CO_2, which is equal to the transported weight (in tonnes) multiplied by the distance travelled and by a CO_2 emission factor.

A common methodology used to keep track of emissions is by calculating the CO_2e, or *Carbon Dioxide Equivalent.* CO_2e describes all the different greenhouse gases in one common unit that allows for comparison. CO_2e expresses in CO_2 the amount of CO_2 which would have the equivalent amount of global warming impact, regardless of the type and quantity of the greenhouse gas (Brander & Davis, 2012).

The amount of CO_2e depends on the type of GHG, which all have a different global warming potential (GWP). The GWP indicates how much a GHG contributes to the global warming over a certain period. Therefore, GWP is an index and CO_2 is in the index with a value of 1. i.e., nitrous oxide has a GWP of 298, meaning that nitrous oxide causes 298 times more warmth over a given period as compared with CO_2 (Brander & Davis, 2012).

A major benefit of using CO_2e, as mentioned by Brander and Davis (2012), is that it allows for easy comparison of different GHGs, since they are expressed in a single number. The downside of this, however, is that one should be cautious when comparing CO_2e figures, because one should know if the GHGs included in the CO_2e figures are equal for all figures.

CO_2e figures for transport emissions are often expressed as *WTW, TTW* or *WTT. WTW* stands for *Well to Wheel* and includes the emissions for a fuel during production, as well as during the use in a powertrain. In contrast, *TTW* which stands for *Tank to Wheel,* only considers the emissions that occur during consumption of the fuel. Finally, figures containing *WTT* only capture the emissions that occur during the production of the fuel, hence the definition *Well to Tank* (Villante et al., 2018).

2.4 Alternative approaches for locating facilities

As mentioned in Section 2.2, proposed solutions to the facility location problem mostly comprise computationally heavy and highly complex mathematical models. Consequently, these propositions are theory based, which questions their applications in practice. As a result, scholars attempt to develop and apply more accessible methods, which settles a part of the accuracy for practicality.

One of the most common methods for determining the location for a single facility is the centre of gravity approach. The application is described by, among others, Krajewski et al.

(2013) and Zi-xia and Wei (2010) and applied by Irwanto and Hasibuan (2018) and Xueying (2014). In these applications, weights, such as transported volume, are attached to each demand point characterizing the relative importance of every node in the network. The centre of gravity takes these weights into account when calculating the average point, resulting in the most optimal location for a facility. Mathematically, the centre of gravity can be expressed as:

$$(1)\ X = \frac{\sum_i l_i x_i}{\sum_i l_i}$$

$$(2)\ Y = \frac{\sum_i l_i y_i}{\sum_i l_i}$$

where X equals the x-coordinate of the centre of gravity, which is the sum of the multiplication of the x-coordinates of the nodes (x_i) by their corresponding load (l_i), divided by the sum of all the loads, and Y is the y-coordinate of the centre of gravity, which equals the sum of the multiplication of the y-coordinates of the nodes (y_i) by their corresponding load (l_i), divided by the sum of all the loads. Equations (1) and (2) are derived from Krajewski et al. (2013).

The centre of gravity approach comes with some caveats, as described by Ballou (1973). Foremostly, when the weights of the demand- and supply points differ greatly, the influence of points with a low magnitude is neglected. This in combination with a highly asymmetrical shape of demand or supply points can result in a centre of gravity that is close to one point, and far away from all others. Secondly, the average percentage of errors decreases when the number of points in the centre of gravity-calculation increases. Ballou (1973) recommends using the *mode method* when the number of points concerned in the calculation is about seven or less. The mode method suggests the point with the highest weight as the optimal location for a facility (Ballou, 1973), instead of the centre-of gravity. When many points are concerned, the centre of gravity method provides a near-optimum location: when trucks are the dominant mode of transportation, the error can be as low as between 0.16% and 0.39% (Ballou, 1973). When rail is the dominant mode of transport, the expected error is 1.6% with a maximum of 3.9% (Ballou, 1973). The difference in error for rail and truck is due to the way transport rates are included in Ballou's (1973) experiment.

As the name implies, the centre of gravity method is only able to calculate one optimal location, the centre. While multiple facilities can increase the efficiency in a supply network, the centre of gravity method is unusable for determining their locations. There are, however, workarounds proposed to this limitation that first divide all demand- and supply points in different groups, and thereafter solve the problem for a single facility, as described by Brimberg and Drezner (2019). The authors propose to divide the supply- and demand points into K clusters, to then determine one facility for every cluster. This way, the multiple facility problem is divided in multiple smaller independent problems. A similar approach is adopted used by Cai et al. (2020), who determine the clusters by applying a clustering algorithm called k-means.

The application of k-means clustering is not limited to the allocation of facilities, but has successfully been used in a variety of fields. Kanungo et al. (2002) describe k-means clustering as follows:

> Given a set of *n* data points in real *d-dimensional* space, $\mathbb{R}^d$, and an integer *k*, the problem is to determine a set of *k* points in $\mathbb{R}^d$, called centers, so as to minimize the mean squared distance from each data point to its nearest center. (Kanungo et al., 2002, p. 881)

The k-means clustering algorithm assigns all datapoints to a predefined (k) number of clusters, while minimizing the total mean squared distance. A key step in the application of k-means clustering is determining k, the number of clusters to be assigned. Kodinariya and Makwana (2013) discuss six different methods for determining k, but do not prefer one method over the others.

2.5 Conclusion based on literature review

Based on the reviewed literature one can conclude that number of publications within scientific literature regarding the topic of FLPs, or more specifically HLPs, is very extensive. Simultaneously, the publications available regarding this topic are mathematically complex or consider very specific constraints. Besides, publications seldom consider environmental impact regarding the location decision.

Some of the alternative approaches for solving FLPs are already less mathematically complex and do consider the environmental consequences of a location decision. Also, methods like the centre of gravity or K-Means clustering allow for a more transparent analysis as compared with the mathematical models. However, publications using these methods concern the location of facilities instead of hubs.

This leaves the question if a combination of accessible methods, such as the centre of gravity and K-Means clustering, can be used to locate cross-docking terminals (hubs) considering the emissions of CO_2e.

3. Methodology

3.1 Methodology Description

The methodology proposed in this paper aims to determine the number of cross-docking terminals and their locations in such a way that CO_2e emissions during transportation are lowest. This is done by applying K-Means clustering using suppliers' coordinates. Next, the centre of gravity is calculated for each cluster, which results in one centroid for each cluster. Next, another centre of gravity is applied to determine the location for a cross-docking terminal. This second centre of gravity, calculated for each cluster, uses the cluster centroid (representing the weight of every supplier) and the delivery areas (using the total delivered weight) to arrive and the location for a cross-docking terminal. After this, the CO_2e emissions can be calculated for the pre-collections and trunkloads between the suppliers, cross-docking terminals and delivery areas. The methodology is described step-by-step in the remainder of this section.

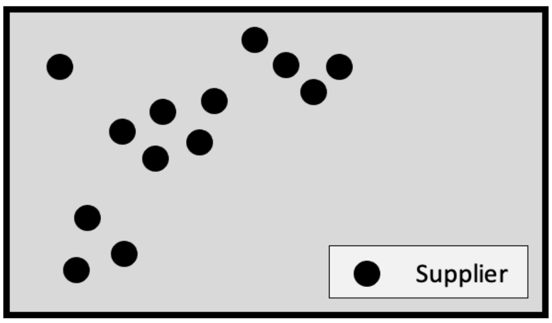

Figure 6: Situation prior to methodology Step 1.

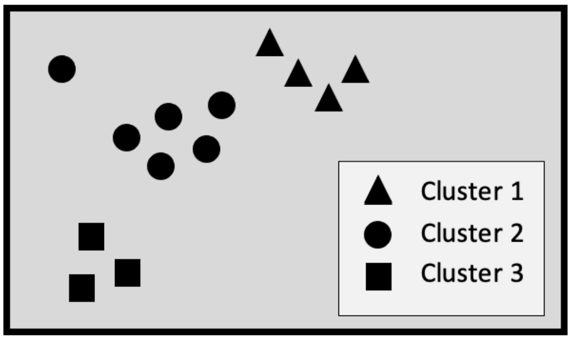

Figure 7: Situation after methodology Step 1. Every supplier is assigned to a cluster.

Step 1: Clustering suppliers using K-Means

Step 1 requires the following data: (1) the coordinates for every supplier in the supply network and (2) the total weight of the supplies produced by each supplier. The situation prior to Step 1 is shown in Figure 6.

K-Means clustering is used to cluster the suppliers into different groups (clusters) based on their longitude and latitude. The clustering is applied using K-Means from scikit learn, a package in Python. Additionally, the *sample weight* parameter is set to the total gross weight of the supplies originating from the supplier. K-Means would cluster the suppliers solely on their coordinates when the gross weight that originates from each supplier is neglected. When all suppliers produce an equal weight of supplies, the sample weight parameter would be redundant. Besides, using the sample weight parameter has advantages in step 2. K, the number of clusters, is chosen using the elbow method, the oldest method available for determining K (Kodinariya & Makwana, 2013). Furthermore, the method for initialization is set to *k-means++* and the number of iterations is 10, the default parameter. See *sklearn.cluster.kmeans* (n.d.) for more details on the different parameters.

After applying K-Means clustering, every supplier is assigned to a cluster, as shown in Figure 7.

Step 2: Determining the centre of gravity (COG) for each cluster.

The output of Step 1, as shown in Figure 7, is the starting point of Step 2. Step 2 does not require any additional data but makes use of the weights and coordinates that were required in Step 1.

After all suppliers are clustered, the centre of gravity (COG) can be calculated for each cluster, as

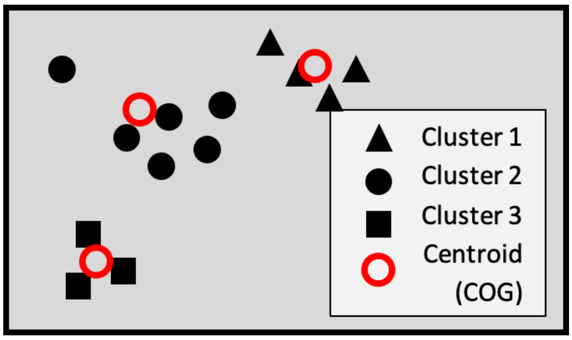

Figure 8. Situation after methodology Step 2.

shown in Figure 8. The easiest way to retrieve the COG in this context is calling the *cluster_centers_* function from K-Means. This function provides the exact coordinates of the centroids for each cluster. And since the *sample weight* parameter is already set for the clusters in Step 1, *cluster_centers_* automatically considers the gross weight of all suppliers in a cluster.

For clarification, *centroids* is a term commonly used during clustering to refer to the middle point of each cluster. The centre of gravity (COG) is the centre point of a selection of points considering their location and weight. In this specific case, suppliers are clustered based on coordinates and weight by using the *sample weight* parameter, as mentioned during Step 1. And since the weights are already considered during clustering, the centroids are equal to the centre of gravity. When clustering is applied without the *sample weight* parameter, one can calculate the COG using equations 1 and 2 from Section 2.4.

This COG would normally be used as the definitive location for a facility, as is done in Cai et al. (2020) and Esnaf and Küçükdeniz (2009). In this case, where a cross-docking terminal is considered, one should not only consider the weight of the suppliers, but also the weight of the final destinations. The weight of the final destinations is included in Step 3.

Step 3: Calculating the COG considering delivery areas.

The output of Step 2, suppliers assigned to a cluster and a COG for every cluster, as shown in Figure 8, is the input for Step 3. Step 3 requires additional data consisting of (1) the total weight produced by suppliers in each cluster, (2) the total weight transported to each delivery area from each cluster and (3) the coordinates of the delivery areas.

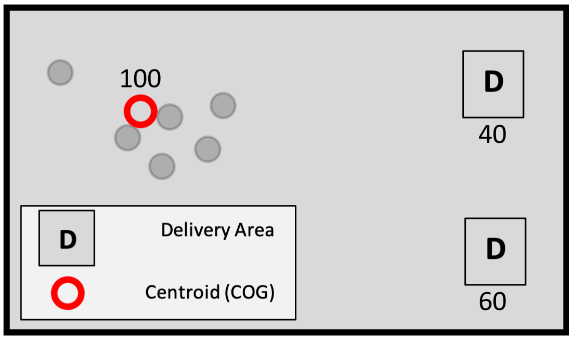

Figure 9. Situation prior to methodology Step 3. 100 is the sum of all the weight produced by the suppliers in the cluster. 40 and 60 are the weight of the supplies from the suppliers in the cluster to the Delivery Area. The sum of the Delivery Areas (40 + 60) is equal to the weight of the cluster centroid (100).

At the end of Step 3, the locations for the cross-docking terminals are calculated. During Step 3, the total weight transported to a Delivery Area form each cluster is used in a second centre of gravity calculation, together with the COG from Step 2. Because a cross-docking terminal does not hold any storage, the weight of the incoming goods is equal to the weight of the outgoing goods over a given period. Thus, the weight of all delivery areas together is equal to that of the COG of the suppliers. This is shown in Figure 9.

After calculating the total weights for the delivery areas and the suppliers in the cluster, the second centre of gravity calculation can be applied. This calculation uses the weight of the delivery areas and the cluster centroid, as shown in Figure 10. This second centre of gravity is where the second cross-docking terminal should be located. It is important to note that in the example, the calculations are conducted for one cluster. In practice, step 3 should be repeated for each cluster, resulting in one cross-docking terminal per cluster.

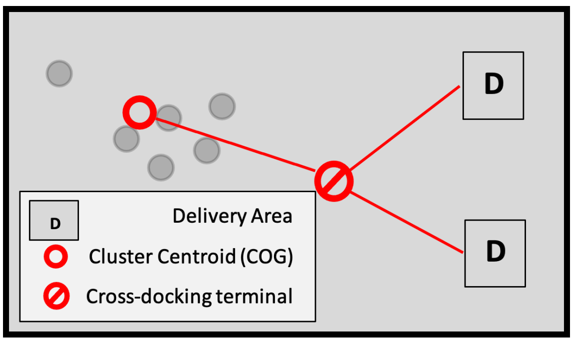

Figure 10. Situation after methodology Step 3. The output of the second centre of gravity calculation is the location for the cross-docking terminal.

Step 4: Finding addresses for calculated COG (optional).

The inputs for Step 4 are the coordinates of the centre of gravity calculations in Step 3. These coordinates represent the locations for cross-docking terminals. The final two steps of the methodology, Step 4 and Step 5, are optional in cases where one is not interested in the emissions in the supply network. Besides, a different method for CO_2e calculation is used, Step 4 and Step 5 can be considered redundant.

The COGs calculated in Step 3 provide latitude and longitude coordinates. However, for the CO_2e calculations in Step 5, addresses are required as input. Therefore, the coordinates of the COGs obtained in Step 3 are transferred to addresses. This can be done manually, by using an online web mapping platform, or automatically using an API. Since a continuous approach is applied, manually retrieving addresses is preferred because the COG can be in forests or rivers, requiring small adjustments to obtain an address that is suitable for Step 5.

Step 5: Calculating CO_2e emissions using Proxio

Calculating the CO_2e for the supply network with the new cross-docking terminal locations is done via Proxio. For more information on Proxio, see the Section *3.2.* As input, Proxio requires postal codes, cities, and countries, as well as the weights that are shipped. Table 1 contains an overview of the required input, as well as an example of a record.

Table 1

Required input and example record for Proxio input.

Input	Example of a record
Transport date	09-05-2021
Origin postal code	061099
Origin city	Bucharest
Origin country code	RO
Destination postal code	63-630
Destination city	Stogniewice
Destination country code	PL
Case weight (Kg)	19329
Consignment id	-
Shipping id	-

Note. – are optional entries.

In Table 1, *Origin city* can be a supplier or a cross docking terminal since this is where transports depart from. *Destination city* can be either a cross-docking terminal, in case of a pre-collection ride, or a delivery area, in case of a trunkload. This way, Proxio can calculate the CO_2e emissions for all transports between suppliers and the new cross-docking terminal, and between the cross-docking terminal and the delivery area.

The weight of transports between the new cross-docking terminal location and the delivery areas cannot be based on historical data because the composition of the trunkload has changed because suppliers are assigned to cross-docking terminals differently. Therefore, weights between cross-docking terminals and delivery areas are based on the average trunkload weight in the current situation. The same applies for the number of transports between dross-docking terminals and delivery areas.

3.2 Data collection and -pre-processing

The proposed methodology is applied at data that originates from different systems used in Scania's daily operation. This section provides details on the data used in the case study, such as a description of the source and preparations made to the data prior to application of the method. For a detailed description of the preparation of the data, see Appendix A.

The data used originates from three different systems, *iNet TA* (Transport Analytics), for historical transport data, *Proxio*, for CO_2e calculations, and *Vista,* containing supplier details, such as addresses and coordinates. Figure 11 shows how the data from the different systems is used prior to the analysis. The data covers all inbound (incoming transport from Scania's perspective) transports between April 1st, 2021, and June 30th, 2021. According to the organization, this period is representative for Scania's operations, capturing the variations in supply and demand that typically occur in the supply network. 22,368 inbound transportations originating from 830 unique suppliers are covered by the data used.

iNet TA

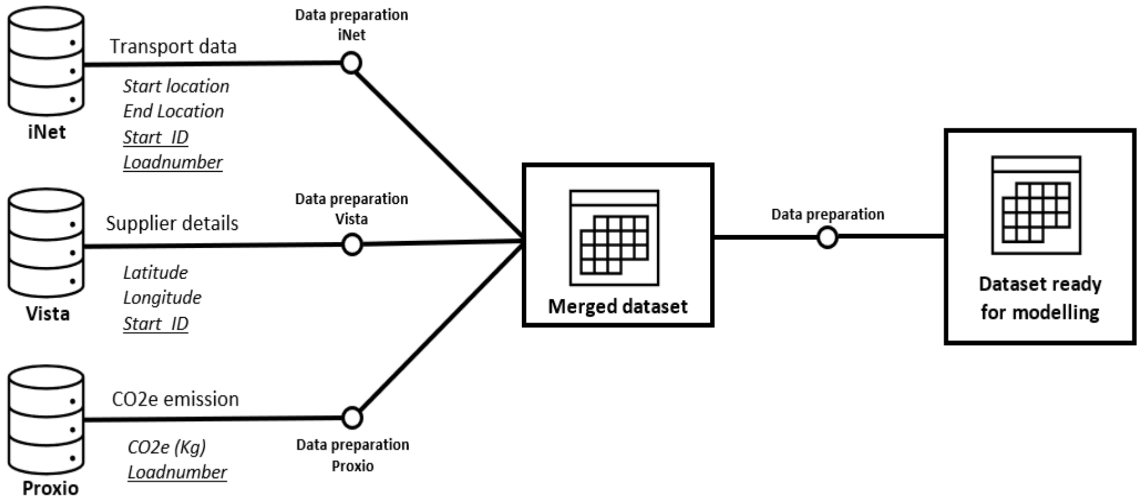

Figure 11: Systematic overview of the relationship between the different datasets used.

iNet TA (Transport Analytics) is the system used to analyse historic transport data. iNet TA contains attributes such as pick-up and drop-off locations, transport distances and volumes transported for every transportation leg in Scania's supply network. The data in iNet TA originates from iNet, SLN's transport management system. When a transport takes place within SLN's supply network, it is created in iNet. iNet TA is used to access the transport data in iNet and its main purpose is to satisfy the organization's need for information that can be used in analysis.

Proxio

Proxio is a third-party system used by SLN to calculate the amount of CO_2e emitted during inbound transportations in Scania's supply network. For every inbound transport, Proxio calculates the amount of CO_2e emitted based on predefined parameters that characterize the route from pick-up to drop-off. For the calculations, Proxio considers parameters such as fuel consumption (including, but not limited to, engine type, road type and road gradient), weight of the load, the distance of transportation and an emission factor. Reported CO_2e emissions in this research are all WTW (Well to wheel). And by default, Proxio assumes that a trailer is filled for 60% of the maximum weight. So, weights entered lower than 60% of maximum capacity will be included as if 60% of the truck is filled (for weight).

Within Proxio, different modes of transport can be combined systematically to create so-called *emission profiles.* Proxio uses these profiles to calculate the CO_2e emissions occurring on specific routes. For example, supplies transported to delivery areas in Sweden from a cross-docking terminal in Germany might involve a Ferry. For these transports an emission profile is created that takes into account the emissions for every leg of transportation. Proxio uses ZIP codes for start- and end locations to assign the corresponding emission profile.

Vista

The final data required for the research originates from Vista, another third-party system used by SLN. Vista is a deck-coverage calculator used by SLN to organize the packaged parts inside a trailer or container. All packaging dimensions are known by Vista, and in combination with the quantities provided by the supplier it calculates how the supplies fit inside a trailer or container. Besides, Vista contains important supplier details, including names, addresses, and coordinates.

It should be mentioned that preparing the data from the above-mentioned sources is an intensive, careful, and time-consuming job. This is mainly due to the fact that the data originates from practice, resulting in deficiencies during automatic analyses. Examples of this are the lack of coordinates in Vista, spelling mistakes or other incorrect notations, such as an O instead of a 0 in zip codes. It should be noted that every adjustment is registered, so that the transparency of the analysis is not lost.

4. Results

As indicated on the title page, the results of the method changed due to different calculation method used for the proposed solutions and the comparison value of the current situation. All results stated in this section are based on the same calculations, allowing for a fair comparison. Previous results, that suggested a successful method, are included in Appendix B. This concerns Tables 4 and 5.

Applying the proposed methodology

830 unique suppliers are currently using cross-docking terminals as consolidation points in SLN's supply network. Applying K-Means clustering using suppliers' coordinates resulted in k = 4, as shown in Figure 12. In the case of Scania, k = 4 means that having more than four cross-docking terminals in the supply network barely reduces the sum of squared distance between the suppliers and their cluster centroid.

More specifically, k = 4 does not necessarily mean that having more than four cross-docking terminals has no effect regarding CO_2e emissions. As Figure 12 shows, the sum of squared errors still decreases beyond k = 4, indicating that also the total distance travelled between suppliers and

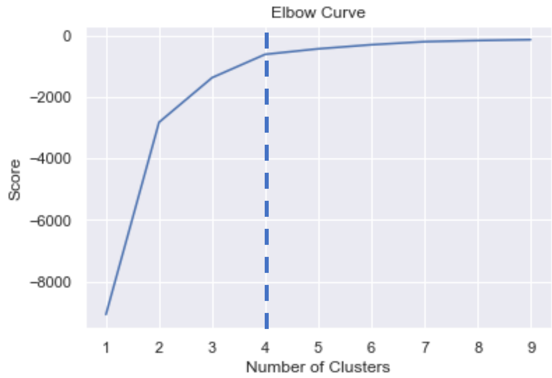

Figure 12. Elbow-curve of K-Means clustering applied at Scania's suppliers. K>4 rarely improves the score (Sum of Squared Error based on distance between suppliers).

cross-docking terminals will decrease when more cross-docking terminals are added.

In the case of Scania, changing the amount of cross-docking terminals in use is relatively simple compared to organizations that have their own cross-docking terminal. Since Scania uses cross-docking terminals from third parties, the organization does not have to consider long-term investments in the facilities itself. For organizations where this is not the case, and the cross-docking terminals are operated by themselves, increasing the number of terminals from 4 to 5 is likely to increase the costs significantly, while only having limited benefits regarding the total distance travelled between suppliers and cross-docking terminals. For

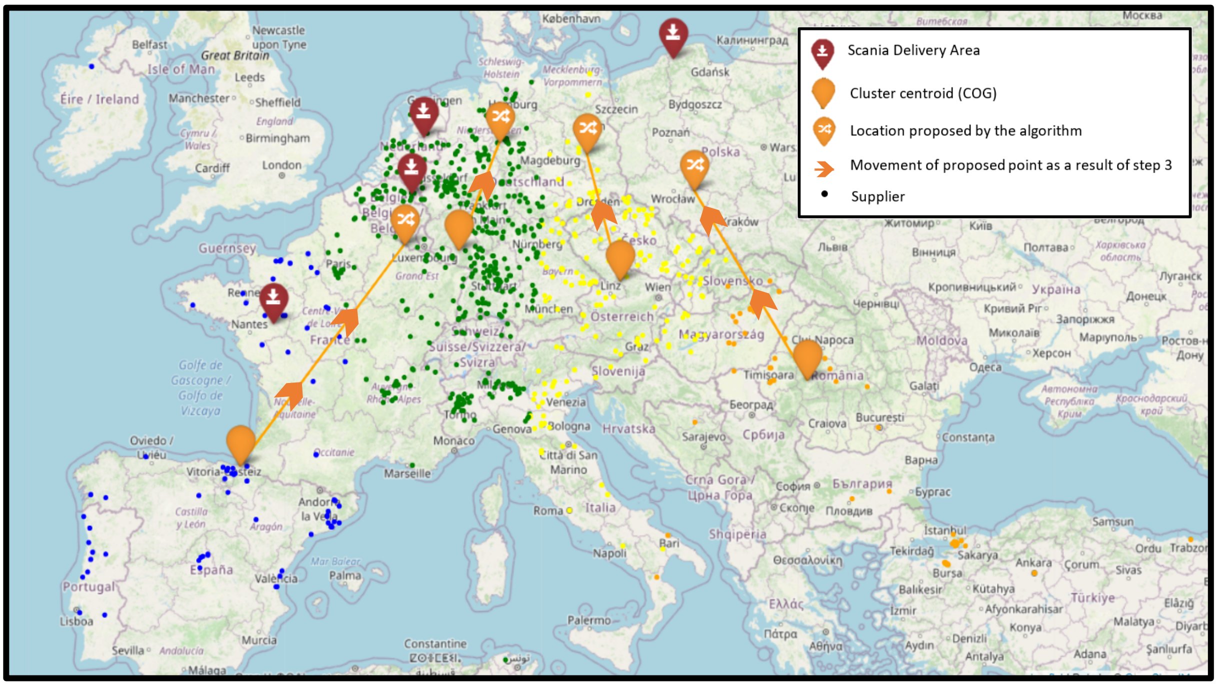

Figure 13. Proposed network based on the 4 clusters. Filled orange markers are set at suppliers' centre of gravity. Orange markers with the mixed arrows are set at the centre of gravity of the delivery areas and suppliers centre of gravity. Orange lines mark how the centres of gravity move when delivery areas are included. ©OpenStreetMap-contributors.

Table 2

Details for Figure 13: Cluster labels and details for clusters' COG (City, ZIP Code and Country Code).

Cluster Label	Latitude	Longitude	City	ZIP Code	Country Code
0 (yellow)	48.734636	14.716168	Černé Údolí	382 41	CZ
1 (green)	49.486116	7.587073	Kollweiler	66879	DE
2 (blue)	43.889325	-2.068653	San Sebastian	20011	SP
3 (orange)	45.902591	23.034441	Deva	330182	RO

Table 3

Details for Figure 13: Cluster labels and details for clusters' COG considering the weight of delivery areas (markers with arrows). (City, ZIP Code and Country Code).

Cluster Label	Latitude	Longitude	City	ZIP Code	Country Code
0 (yellow)	52.080	13.231	Nuthe-Urstromal	14947	DE
1 (green)	52.369	9.396	Hohnhorst	31559	DE
2 (blue)	49.668	5.211	Florenville	6820	BE
3 (orange)	51.112	17.951	Stogniewice	63-630	PL

Table 4

CO_2e emissions and TonneKMs in the current situation, with weighted centroids and with unweighted centroids.

	Current Situation		Weighted Centroids		Unweighted Centroids	
Situation	Kg CO_2e	TonneKM	Kg CO_2e	TonneKM	Kg CO_2e	TonneKM
Pre-collection	831,668	14,812,767	2,051,456	35,545,190	1,297,736	22,531,350
Trunk loads	2,088,672	54,539,560	2,332,988	40,492,209	3,011,134	59,607,993
Total	2,920,340	69,352,327	4,384,444	76,037,399	4,308,870	82,139,343
% Difference	-	-	+ 50.13	+ 9.64	+ 47.55	+ 18.44

clarification, the methodology considers one cross-docking terminal for every cluster, located at the clusters' centres of gravity.

The exact coordinates of the four clusters' centres of gravity are included in Table 2 and plotted as filled markers in Figure 13. Calculating the centre of gravity for each of the four clusters considering Scania's delivery areas results in the points marked as markers (twisted arrows) in Figure 10. These markers are located at the locations proposed by the methodology for the cross-docking terminals. Details on these markers are included in Table 3.

Addresses corresponding with the calculated COGs are used as origin and destination points for CO_2e calculations in Proxio. The results of the Proxio calculations are provided in Table 4. Results show that the total amount of CO_2e emitted during the current situation is 2,920,340 Kg, and 4,384,444 Kg for the situation proposed by the algorithm. This means that CO_2e emissions during transportation in the situation proposed by the method are 50.13% higher than in the current situation. In Table 4, this value is included as *% Difference* in the column *Weighted Centroids.*

Weighted Centroids is the term used to describe the calculations that consider the weight of the delivery areas in the COG calculation, i.e., the result of step 3 in Section 3. The weighted centroids are marked as a marker (twisted arrows) in Figure 13. *Unweighted Centroids* are the COG's that do not consider the weight of the delivery areas, i.e., the output of step 2 in Section 3. The unweighted centroids are marked in Figure 13 with the filled marker.

As a comparison, CO_2e emissions are calculated for Weighted Centroids as well as Unweighted Centroids. Including the unweighted centroids provides an interesting comparison value. As shown in Table 4, the total amount of CO_2e for the weighted centroids is only slightly more than the total amount of CO_2e for the unweighted centroids. However, there are differences between CO_2e emissions during pre-collection and trunk loads.

During pre-collection rides in the weighted centroids situation, larger distances are travelled with lower weights. In the unweighted centroids scenario, where the cross-docking terminal is located closer to the suppliers, the amount of TonneKM is lower than the equal rides in the weighted centroid scenario. Thus, the second COG

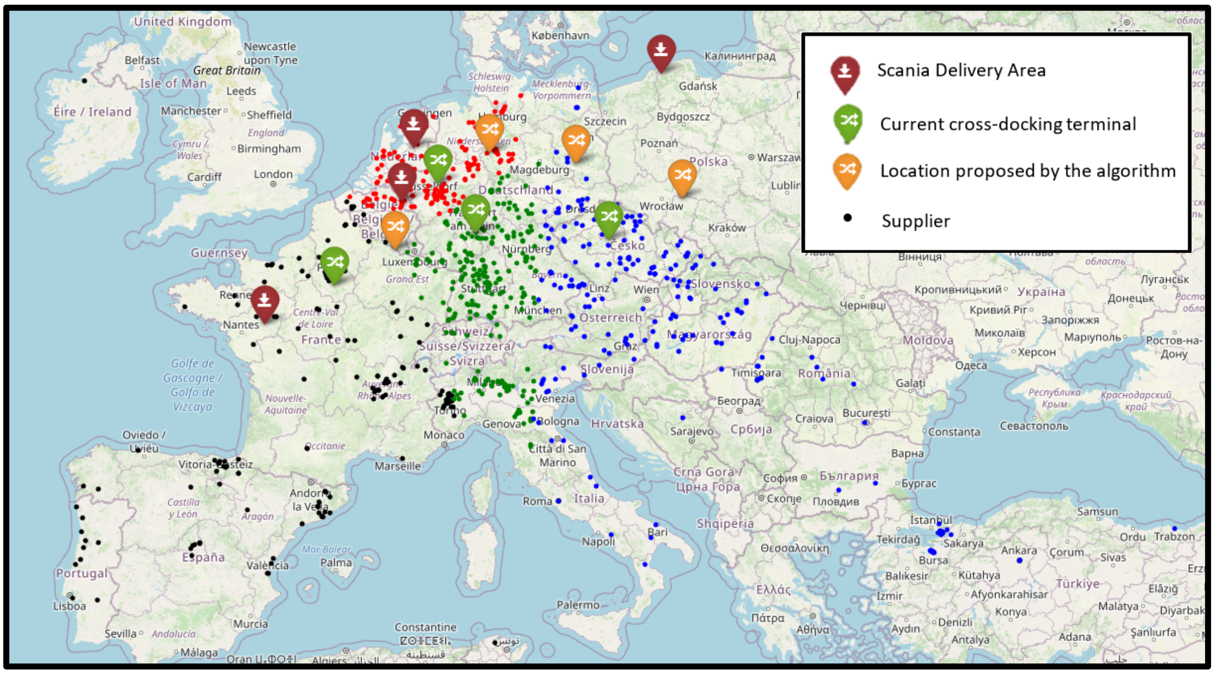

Figure 14. Part of Scania's supply network showing the four current cross-docking terminals that are closest to the locations proposed by the algorithm. ©OpenStreetMap-contributors.

Table 5

CO_2e emissions and TonneKMs using four of the cross-docking terminals currently in use that are closest to the locations proposed by the algorithm.

	Current Situation		**Weighted Centroids**		**Four current Cross-docking terminals**	
Situation	Kg CO_2e	TonneKM	Kg CO_2e	TonneKM	Kg CO_2e	TonneKM
Pre-collection	831,668	14,812,767	2,051,456	35,545,190	1,217,386	21,684,424
Trunk loads	2,088,672	54,539,560	2,332,988	40,492,209	2,089,475	50,724,594
Total	2,920,340	69,352,327	4,384,444	76,037,399	3,306,861	72,409,018
% Difference	-	-	+ 50.13	+ 9.64	+ 13.26	+ 4.41

calculation negatively influences the distances travelled and the CO_2e emissions during pre-collection.

Besides, the value for TonneKM in the proposed scenario (40,492,209) is lower than the TonneKM in the current situation (54,539,560). As shown in Figure 13, the cross-docking terminals in this situation are much closer to the delivery areas as compared with the current network (shown in Figure 2). And since greater distances are travelled during pre-collection in the proposed situation, travelling shorter distances with trunk loads is well-reflected in the emission results. Travelling greater distances with pre-collection rides does not result in lower CO_2e emissions.

Additional analysis: using four cross-docking terminals currently in use

In addition to the previous results, it is valuable to see how the four locations for cross-docking terminals proposed by the algorithm compare with a situation where only four of the current cross-docking terminals are used. To conduct this analysis, the four cross-docking terminals currently used in Scania's supply network that are closest to the locations proposed by the algorithm are selected. Thereafter, suppliers are assigned to the closest of the four cross-docking terminals. Subsequently, CO_2e emissions are calculated for pre-collection and trunkload transports. Results are included in Table 5. Figure 14 shows the locations proposed by the method and the nearest four current cross-docking terminals.

Results show that using the four locations for cross-docking terminals proposed by the algorithm does not result in a reduction of CO_2e emissions compared to a situation where four of the current cross-docking terminals are used. CO_2e emissions are higher for both pre-collection and trunk load transportation in the *weighted* *centroids* situation, but the difference is mainly due to pre-collection rides.

In the situation proposed by the algorithm, suppliers are not always assigned to the nearest cross-docking terminal, resulting in transports that are unrealistic or counterintuitive from a practical standpoint. For example, as shown in Figure 10, pre-collection rides originating from Spain must drive all the way to the Belgium border for a cross-docking terminal. Simultaneously, the suppliers close to the Belgium border are not using cross-docking terminal in Belgium but must ride towards Northern Germany for their consolidation point. This is reflected in the CO_2e emissions for pre-collection shown in Table 5.

Additional analysis: emissions during pre-collection for different values of k

In addition to the two previous analyses, a final analysis is conducted to analyse how CO_2e emissions change during pre-collection for different levels of k. In the case of Scania, k is equal to the number of cross-docking terminals used in the supply network. For context, applying k-means clustering for $k = 1$ and subsequently calculating the CO_2e emissions resembles a supply network that uses one cross-docking terminal located at the centre of gravity of all suppliers. For $k = 10$, CO_2e emissions during pre-collection are calculated for a supply network using ten cross-docking terminals.

This additional analysis is interesting for several reasons. Foremostly, the majority of CO_2e in the current situation is emitted during pre-collection transports, as shown in Table 5. This justifies focussing more on this part of the supply network. Besides, all pre-collection transports in the current situation are executed using diesel, which is not the preferred option from a CO_2e perspective. In addition, grouping supply points to solve the multiple facility location problem has been discussed in Section 2. However, an important caveat here is that the influence of delivery areas in the facility location is disregarded in this analysis.

K-Means clustering is applied to the coordinates of Scania's suppliers for $k = 1$ until $k = 10$. For the output of every value of k, CO_2e emissions are calculated using Proxio. The results are plotted in

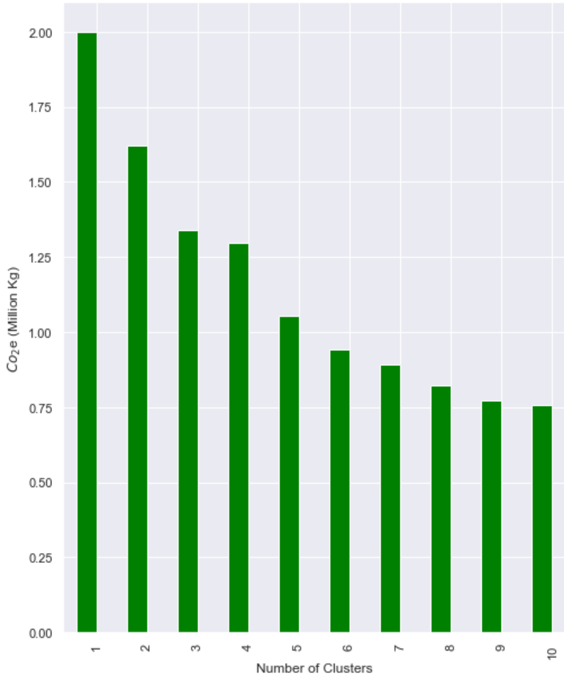

Figure 15. CO2e emissions for different values of k.

Figure 15. The bars show the amount of CO_2e emitted for every value of k between 1 and 10. Figures containing the clustered suppliers are included in Appendix C.

The distance to a cross-docking terminal decreases when more terminals are present in the network, which results in lower CO_2e emissions. This, however, only holds until a certain amount of cross-docking terminals. One of the factors causing the use of cross-docking terminals to be efficient is the fact that supplies can consolidated. When there are many cross-docking terminals present in a supply network, and only a few suppliers are assigned to each cross-docking terminal, an organization would not be able to combine shipments at a high frequency. In such a scenario, supplies have to be stored until a full trunkload of supplies can be forwarded to a delivery area. This has consequences regarding, storage and lead-times and is therefore not desirable.

Plotting the cluster centroids for $k = 8$ combined with the cross-docking terminals currently in use provides an interesting insight in Scania's supply network, see Figure 16. Emissions for pre-collection in the current situation (831,668) are only slightly higher than the scenario for $k = 8$ (823,590). The Figure 16 shows that for every cluster centroid (triangle marker) calculated by K-Means clustering for $k = 8$ there is a cross-

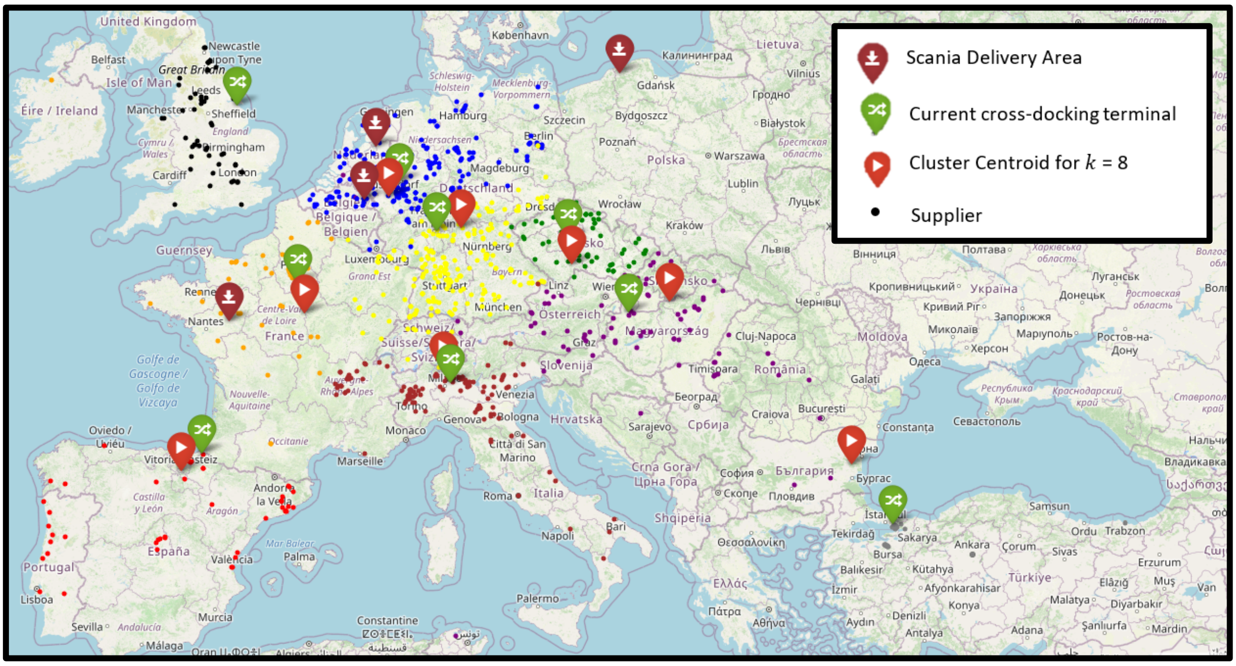

Figure 16. Scania's supply network including current cross-docking terminals and cluster centroids for k =8. Suppliers are coloured by current cross-docking terminal. (Suppliers in Great Britain are included in the image for completeness of network overview but are not part of clustering for k =8.) ©OpenStreetMap-contributors.

docking terminal that is currently being used by Scania. So, solely focussing on pre-collection transports there is no reason to assume that Scania's supply network is not efficiently organized.

5. Conclusion, limitations, contributions, and future research.

In this research a new methodology for locating multiple cross-docking terminals considering the emissions of CO_2e in a supply network with multiple supply and demand points was proposed and applied at a large manufacturer in the automotive industry: Scania. The methodology follows a continuous approach and combines k-means clustering with the centre of gravity. The methodology aimed to determine the amount of cross-docking terminals to be used as well as their position in the supply network. Simultaneously, by using historical data, the research aimed to provide insight in how historical transport data can be used in supply chain decisions considering the emission of CO_2e.

The supply network design as proposed by the methodology, i.e., the number of cross-docking terminals and their location, does not result in a reduction of CO_2e emissions as compared to the current situation where 8 cross-docking terminals are used. The CO_2e emissions calculated for the proposed situation are 50.13% higher compared with the current situation. In fact, the supply network design proposed by the methodology results in transports that are unrealistic or counterintuitive from a practical standpoint.

The lack of success of the methodology in the case of Scania can be attributed to several factors. First of all, the methodology does not assign suppliers to the nearest cross-docking terminal. An additional analysis showed that a in a situation, with four cross-docking terminals currently in use, where each supplier is allocated to the closest cross-docking terminal, the CO_2e emissions are lower than in the situation proposed by the method.

Secondly, during Step 3 in the Method, where the weights of the delivery areas are considered, the proposed location of a cross-docking terminal moves towards the north for every cluster. This is due to the fact that the vast majority of the suppliers are located south with respect to the delivery areas. Perhaps, in a situation where delivery areas are located more randomly within the supply network, this movement is less extreme, resulting in proposed location within or close to the cluster.

A final factor that limits the success of the method proposed in the case of Scania is the fact that there are currently eight cross-docking terminals being used by the organization. And while the method proposes four, having eight cross-docking terminals was proven to reduce the total distances between suppliers and cross-docking terminals. Since Scania does not operate the cross-docking terminals themselves, and as a result, therefore, does not have to consider any investments, the minor improvements in distance from four to eight terminals are worth the extra effort.

Finally, an additional analysis solely focussing on pre-collection rides (from suppliers to cross-docking terminals) whereby suppliers are clustered based on location and supplied weight results in locations for cross-docking terminals that are close the cross-docking terminals already in use by Scania. So, solely focussing on pre-collection transports there is no reason to assume that Scania's supply network is not efficiently organized when it comes to CO_2e emissions.

Resume, locations for cross-docking terminals proposed by the methodology do not result in the reduction of CO_2e emissions. However, by applying K-Means clustering and the centre of gravity, historical transport data is proven to be useful in supply chain decisions considering the emission of CO_2e.

This research is **limited** by several factors. First, due to the use of a continuous approach, the locations suggested by the methodology as well as in the additional analysis are not equipped with a cross-docking terminal, meaning that the amount of CO_2e calculated by the model will deviate from reality when a terminal nearby is used.

Another important limitation to this research is the fact that only supplies from Scania are considered in the analysis. This led to the assumption that trucks were only filled with Scania supplies or were empty otherwise. In practice, however, carriers are free to combine shipments from different customers, or pick up supplies from different customers at the same supplier. Assuming that carriers are also aiming to optimize their shipments and therefore seek to combine them, a change to the supply network that decreases CO_2e for one, might in fact increase the overall CO_2e-emissions.

Additionally, the application of K-Means clustering also comes with several limitations, especially in the case of geographical data. Firstly, the distances concerned within a K-Means cluster are Euclidian distances. In practise, transportation is limited to the available infrastructure, resulting in a travelled distance between suppliers and cross-docking terminals that is higher than the Euclidian distance between them. Secondly, K-Means clustering neglects the presence of water, mountains, or other insurmountable objects within a cluster. This can result in suppliers being assigned to a cross-docking terminal that is not the best option in practise.

This research has several **practical contributions.** Foremostly, it provides meaningful insights in the dynamics of the supply network of the organization where the methodology is applied: Scania. The methodology proposed in this research aimed to be more accessible and more transparent for practitioners in supply network engineering as compared with other publications on this topic. Despite the lack of success of the methodology itself, the research shows that, by applying K-Means clustering and the centre of gravity, historical data can be used in in supply chain decisions considering the emission of CO_2e.

This research **contributes to the literature** by proposing and applying a new methodology for locating multiple cross-docking terminals considering the emissions of CO_2e in a supply network with multiple supply and demand points. The methodology is an alternative approach to the multi-facility location problem, or hub-locations problem more specifically. By combining two accessible and transparent methods, K-Means clustering and the centre of gravity, the research aimed to fill a gap in the literature with an approach that is less mathematically complex. However, due to the lack of success of the method, the research failed to fill this gap with a successful method. Additionally, this research is another contribution in the field of facility location problems that emphasizes the need for sustainable considerations in supply chain decisions.

Future research at Scania can be manifold. Considering the limited mathematical knowledge within the organization regarding traditional approaches on this topic, applying a more sophisticated approach, such as (non)-linear models, can provide additional insights in Scania's supply network and the use of cross-docking terminals within it. Besides CO_2e emissions, transportation costs could be considered in future research to contribute more concretely to a potential business case.

On the short term, the organization should focus on using different types of transportation or fuel during pre-collection transports that are more environmentally friendly. Especially during pre-collection transports, where the majority of CO_2e is emitted and only diesel is used, improvements should be considered.

In general, researchers should continue to develop models that support strategic decision making considering the emissions of greenhouse gases. Simultaneously, these models should become more accessible for a wider audience and practitioners in the work field of supply network design.

References

Adeleke, O. J., & Olukanni, D. O. (2020). Facility location problems: models, techniques, and applications in waste management. *Recycling*, *5*(2), 10. https://doi.org/10.3390/recycling5020010

Agustina, D., Lee, C., & Piplani, R. (2010). A review: Mathematical modles for cross docking planning. *International Journal of Engineering Business Management*, *2*(2), 47-54. https://hrcak.srce.hr/65649

Akkerman, F., Lalla-Ruiz, E., Mes, M., & Spitters, T. (2022). Cross-Docking: Current Research Versus Industry Practice and Industry 4.0 Adoption. In T. Bondarouk & M. R. Olivas-Luján (Eds.), *Smart Industry – Better Management* (Vol. 28, pp. 69-104). Emerald Publishing Limited. https://doi.org/10.1108/S1877-636120220000028007

Alumur, S., & Kara, B. Y. (2008). Network hub location problems: The state of the art. *European Journal of Operational Research*, *190*(1), 1-21. https://doi.org/10.1016/j.ejor.2007.06.008

Apte, U. M., & Viswanathan, S. (2000). Effective Cross Docking for Improving Distribution Efficiencies. *International Journal of Logistics Research and Applications*, *3*(3), 291-302. https://doi.org/10.1080/713682769

Arabani, A. B., & Farahani, R. Z. (2012). Facility location dynamics: An overview of classifications and applications. *Computers & Industrial Engineering*, *62*(1), 408-420. https://doi.org/https://doi.org/10.1016/j.cie.2011.09.018

Ballou, R. H. (1973). Potential error in the center of gravity approach to facility location. *Transportation Journal*, *13*(2), 44-50. https://www.jstor.org/stable/20712268

Brander, M., & Davis, G. (2012). Greenhouse gases, CO_2, CO_2e, and carbon: What do all these terms mean. *Econometrica, White Papers*. https://ecometrica.com/greenhouse-gases-co2-co2e-and-carbon-what-do-all-these-terms-mean/

Brimberg, J., & Drezner, Z. (2019). Solving multiple facilities location problems with separated clusters. Operations Research Letters, 47(5), 386-390. https://doi.org/10.1016/j.orl.2019.07.007

Cai, C., Luo, Y., Cui, Y., & Chen, F. (2020). Solving Multiple Distribution Center Location Allocation Problem Using K-Means Algorithm and Center of Gravity Method Take Jinjiang District of Chengdu as an example. *IOP Conference Series: Earth and Environmental Science*, *587*(1), 012120. https://doi.org/10.1088/1755-1315/587/1/012120

Campbell, J. F. (2005). Strategic network design for motor carriers. *Logistics systems: Design and optimization*, 245-278. https://doi.org/https://doi.org/10.1007/0-387-24977-X_8

Chen, D., & Zhang, Y. (2018). Location Optimization of Cross Dock Network Considering Carbon Emissions. In *CICTP 2017: Transportation Reform and Change—Equity, Inclusiveness, Sharing, and Innovation* (pp. 1519-1527). American Society of Civil Engineers Reston, VA. https://doi.org/doi:10.1061/9780784480915.161

Cristea, A. (2013). Trade and the greenhouse gas emissions from international freight transport. *Journal of Environmental Economics and Management*, *65*(1), 153. https://doi.org/https://doi.org/10.1016/j.jeem.2012.06.002

Drezner, T. (1994). Optimal continuous location of a retail facility, facility attractiveness, and market share: an interactive model. *Journal of Retailing*, *70*(1), 49-64. https://doi.org/10.1016/0022-4359(94)90028-0

Esnaf, Ş., & Küçükdeniz, T. (2009). A fuzzy clustering-based hybrid method for a multi-facility location problem. *Journal of Intelligent Manufacturing*, *20*(2), 259-265. https://doi.org/https://doi.org/10.1007/s10845-008-0233-y

Gue, K. R., & Kang, K. (2001). Staging queues in material handling and transportation systems. Proceeding of the 2001 Winter Simulation Conference (Cat. No. 01CH37304), 1104-1108. IEEE. http://doi.org/10.1109.WSC.2001.977421

Gümüs, M., & Bookbinder, J. (2004). Crossdocking and its implications in locationdistribution systems. *Journal of Business Logistics*, *25*(2), 199. https://doi.org/10.1002/j.2158-1592.2004.tb00187.x

Irwanto, B., & Hasibuan, S. (2018). Determination of Pharmaceutical Industrial Distribution Center Location Using Center of Gravity Method: Case Study At Pt Jkt. *Jurnal Operations Excellence: Journal of Applied Industrial Engineering*, *10*(3), 228-239. https://doi.org/10.22441/oe.v10.3.2018.003

Ivanov, D. (2010). An adaptive framework for aligning (re) planning decisions on supply chain strategy, design, tactics, and operations. *International Journal of Production Research*, *48*(13), 3999. https://doi.org/10.1080/00207540902893417

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. *IEEE transactions on pattern analysis and machine intelligence*, *24*(7), 881-892. https://doi.org/10.1109/TPAMI.2002.1017616

Klose, A., & Drexl, A. (2005). Facility location models for distribution system design. *European Journal of Operational Research*, *162*(1), 4-29. https://doi.org/https://doi.org/10.1016/j.ejor.2003.10.031

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies.*, *1*(6), 90-95.

Krajewski, L. J., Ritzman, L. P., & Malhotra, M. K. (2013). *Operations Management* (10 ed.). Pearson Education, Inc.

Kreng, V. B. (2008). The benefits of a cross-docking delivery strategy: a supply chain collaboration approach. *Production Planning and Control*, *19*(3), 229. https://doi.org/https://doi.org/10.1080/09537280801916157

Ladier, A.-L., & Alpan, G. (2016). Cross-docking operations: Current research versus industry practice. *Omega*, *62*, 145-162. https://doi.org/https://doi.org/10.1016/j.omega.2015.09.006

Larioui, S., Reghioui, M., Elfallahi, A., & Elkadiri, K. E. (2017). From a classical Cross dock to a PI cross dock: Latest Advances and New Challenges. *Revue Marocaine de Management, Logistique et Transport*(2).

Magableh, G. M., Rosetti, M. D., & Mason, S. (2005). Modeling and Analysis of a Generic Cross-Docking Facility. *Proceedings of the Winter Simulation Conference, 2005.*, 1613. https://doi.org/10.1109/wsc.2005.1574430

McKinnon, A. C., & Piecyk, M. (2010). Measuring and managing CO2 emissions in European chemical transport. http://www.cefic.org/Documents/IndustrySupport/Transport-and-Logistics/Sustainable%20Logistics/McKinnon%20Report%20Transport%20GHG%20emissions%2024.01.11.pdf

Scania. (2021a). *Delivery structure for inbound transports* [Internal document [Description]].

Scania. (2021b). *The Scania Report 2021*. https://www.scania.com/content/dam/group/investor-relations/annual-review/download-full-report/scania-annual-and-sustainability-report-2021-final.pdf *sklearn.cluster.kmeans*. ((n.d.). Sci-kit Learn. Retrieved June 9 from https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html

Stephan, K., & Boysen, N. (2011). Cross-docking. *Journal of Management Control*, *22*(1), 129-137. https://doi.org/10.1007/s00187-011-0124-9

Treitl, S., & Jammernegg, W. (2014). Incorporating environmental aspects in an inventory routing problem. A case study from the petrochemical industry. *Flexible Services and Manufacturing Journal*, *26*(1), 143-169. https://doi.org//10.1007/s10696-012-9158-z

Van Belle, J., Valckenaers, p., & Cattrysse, D. (2012). Cross-docking: State of the art. *Omega*, *40*(6), 827-846. https://doi.org/10.1016/j.omega.2012.01.005

Villante, C., Anatone, M., & De Vita, A. (2018). On WTW and TTW Specific Energy Consumption and CO_2 Emissions of Conventional, Series Hybrid and Fully Electric Buses. *SAE International Journal of Alternative Powertrains*, *7*(1), 5-26. https://www.jstor.org/stable/26559565

Vogt, J. J. (2010). The successful cross-dock based supply chain. *Journal of Business Logistics*, *31*(1), 99-119. https://doi.org/10.1002/j.2158-1592.2010.tb00130.x

Xueying, Z. (2014, 2014/05). Based on Gravity Method of Logistics Distribution Center Location Strategy Research. Proceedings of the International Conference on Logistics, Engineering, Management and Computer Science, 586-589. https://doi.org/10.2991/lemcs-14.2014.134

Zi-xia, C., & Wei, H. (2010). Study and application of Center-of-Gravity on the location selection of distribution center. 2010 International Conference on Logistics Systems and Intelligent Management (ICLSIM), 981-984. https://doi.org/10.1109/ICLSIM.2010.5461101

Appendix A Data pre-processing

Pre-processing – iNet TA dataset (Inbound).

The data exported from iNet TA used in this research covers all inbound transports taking place between 00:00 April 1st, 2021, and 23:59 June 30th, 2021. An inbound transport is a transport between a supplier and a delivery area, whether or not via a cross-docking terminal. The period is selected because, according to the organization, it well represents the characteristics of daily operations in the supply chain, making it a representative sample. During the selected period, operations did not or only to a very limited extend suffered due to circumstances caused by the covid-19 pandemic.

The selected period contains 32,023 unique records (rows), which means that 32,023 transports took place in the selected period. A transport in this case can best be defined as the movement of goods between the start location (supplier, cross-docking terminal, warehouse) and its destination (Delivery area). Every record in the dataset contains a unique loadnumber, which is used to distinguish between the different transports that took place. Besides the loadnumber, the dataset contains fifteen other attributes that describe a transport. A detailed description of all attributes is included in Table A1, the attributes most relevant for the analysis are elaborated in the next paragraph. The attributes not discussed here mainly cover additional information concerning suppliers and destinations, such as countries, postal codes, and cities.

Table A1

Attribute names, attribute definition and new attribute names.

Attribute name	Definition	Renamed to:
005TMLD Start Request Date	Start date of transport	Start_Request_Date
005TMLD End Request Date	End date of transport	End_Request_Date
005TMLD Business Case	Type of transport. All entries have Inbound as business case.	-
005TMLD Load Number	Unique number to distinguish a transport	Loadnumber
005TMLD Start ID	Unique string to identify a starting point (supplier, cross-docking terminal or delivery area).	Start_ID
005TMLD Start City	City of departure for transport.	Start_City
005TMLD Start ZIPCode	ZIP Code of departure location for transport.	Start_ZIP
005TMLD Start Country	Country of departure for transport.	Start_Country
005TMLD End ID	Unique string to identify end point (cross-docking terminal or delivery area).	End_ID
005TMLD End City	City of arrival for transport.	End_City
005TMLD End ZIPCode	ZIP Code of transport destination	END_ZIP
005TMLD End Country	Destination country	End_Country
005TMLD End CountryCode	Country code for destination country	End_CC
005TMLD Gross Weight in kg	Gross weight of the goods transported. Gross weight is the weight of the parts and the packaging	Gross_Weight
005TMLD Volume in m^3	Volume of the transported parts, including packaging.	Volume
005TMLD Calculated Costs in	Costs calculated for the transport.	Costs_EUR

Note. – implies that the attribute is deleted from the dataset.

The attribute *Start ID* is a unique code that describes where the transport departures. All suppliers and cross-docking terminals have a unique *Start ID,* but because multiple transports can originate from one supplier *Start ID's* can occur more than once in the data. Similar to *Start ID* is the attribute *End ID*, a unique code that describes the destination of a transport. It is important to note that a *Start_ID* as well as an *End_ID* can be a cross-docking terminal, since supplies can arrive at and departure from it. For both attributes, cross-docking terminals have the same unique ID. e.g., HUB_IT_MIL is the *Start_ID* and *End_ID* for the cross-docking terminal in Milan, Italy.

Other attributes in the iNet data are *Gross Weight*, describing the gross weight of the transported supplies and packaging in kilos, *Volume in* m^3*,* describing the transported volume in cubic meters. Table A2 contains an example of a record in the iNet TA dataset, a direct load from a supplier to Oskarshamn (MO1).

Table A2

Example of a record in the iNet TA dataset.

Loadnumber	Start_ID	End_ID	Gross Weight	Volume in m^3	…
1548392	6935-0A	MO1	4,377	*10.69*	…

Note. The three dots (…) represent the additional attributes present in the data.

As mentioned in Section 1, Scania created a hierarchical structure to manage the unloading areas in its supply Network. The methodology includes delivery areas, but this data is not present in the dataset. The delivery area can, however, be derived from the *End_ID* attribute. A new attribute is therefore added to the iNet TA dataset, called *Delivery_area* and is by default filled with the value 'Undefined' for all records. Thereafter, for all records with an *End_ID* that is within the Zwolle delivery area, the value 'Undefined' is replaced with 'Zwolle' in the *Delivery_area* attribute. This is done for all delivery areas. Table A3 contains an overview of how all *End_IDs* are assigned to a specific delivery area.

Table A3

Overview of which values for the End_ID attribute are part of which delivery area

Delivery Area	End_IDss in the iNet TA dataset that belong to the corresponding delivery area.
Angers	ANGERS, 7285-01, SOFICA
Södertälje	SODERTALJE_EN, SODERTALJE, SODERTALJE_AH, SODETALTJE_ALM, SODERTALJE_ME, EN6, 3683-01, HUB_SE_NYK
Oskarshamn	OSKARSHAMN, MO1
Zwolle	MEPPEL, HASSELT, MM4, ZWOLLE_RW, ZWOLLE_LOGISTICS, 58290-0B, PNL806-01, HUB_NL_HAS, 58290-0A, 7168-01
Oudsbergen	OPGLABBEEK

Lulea	LULEA
Slupsk	SLUPSK, 18011-02
Nyköping	OR3
X-Dock	HUB_CZ_KUN, HUB_DE_DUI, HUB_DE_MAI, HUB_ES_IRU, HUB_FR_PAR, HUB_HU_LEB, HUB_IT_MIL, HUB_TR_GB

After assigning the *End_IDs* to the corresponding delivery areas, 416 out of the original 32,023 records still hold 'Undefined' as a value. These records are manually checked and excluded from the dataset because of their irrelevance for the analysis. Some of these instances considered the transportation of empty packaging, others were special deliveries to locations that are not part of the daily operations. After this clean-up, 31,607 records are left in the dataset.

Another attribute is added to the iNet TA dataset called *Departure_From*. This attribute describes where the transport departures from. The default value for this attribute is 'Supplier' and gets replaced by 'X-Dock' or 'Delivery_Area' based on the value present in the *Start_ID* attribute. Table A4 contains an overview of the *Start_IDs* belonging to the corresponding *Departure_From* value.

Table A4

Assignment of the different values for Start_ID to the attribute 'Departure From'.

Departure From	Start_ID in original dataset
Delivery_Area	ZWOLLE_RW, ZWOLLE_LOGISTICS, ANGERS, PNL806-01, 7233-0C, 58290-0A, HASSELT, OPGLABBEEK, 58290-0C, 58290-0B, 7168-01, HUB_NL_HAS, HUB_SE_NYK
X-Dock	HUB_CZ_KUN, HUB_DE_DUI, HUB_DE_MAI, HUB_ES_IRU, HUB_FR_PAR, HUB_HU_LEB, HUB_IT_MIL, HUB_TR_GB

Now that the attributes *Delivery_Area* and *Departure_From* are present in the dataset, another filter can be applied. Transports starting at a supplier and ending at a delivery area have to be excluded from the dataset since these do not involve a cross-docking terminal. After filtering on these criteria, 24,149 records are left in the iNet TA dataset, meaning that it contained 7,458 direct transports.

A similar filter will be applied to the transports starting at a delivery area and ending at a delivery area. In practice this occurs when transport take place from two delivery zones within one delivery area. For example: from Hasselt to Meppel, both in delivery area Zwolle). After filtering on this criteria, 22,769 records are left.

Other records that are included in the dataset but not relevant for the analysis are the ones starting at a delivery area and ending at a cross-docking terminal. These instances also originate from practise and are likely to concern return-transports. After excluding these, 22,683 records are left in the dataset.

The iNet TA dataset now contains all data necessary for analysis. 22,683 records with 16 attributes. The final iNet TA dataset does not contain missing values. Table A5 provides an overview of the records deleted and other adjustments.

Table A5

Effect of the adjustments and applied filtering to the iNet TA dataset.

Adjustment	# Records after adjustment
Original dataset	32,023
Adding attributes Delivery_Area and Departure_From	32,023
Excluding 'Undefined' delivery areas	31,607
Excluding direct transports from suppliers to delivery areas	24,149
Excluding transports from delivery areas to delivery areas	22,769
Excluding transports from delivery area to cross-docking terminal	22,683

Pre-processing – Proxio dataset.

The data exported from Proxio contains the calculated amount of CO_2e emitted during all inbound transports taking place between 00:00 March 28th, 2021 and 00:00 July 15th, 2021. The amount of CO_2e emitted is aggregated by *shipping_id*, which is unique for every record. A *shipping_id* in the Proxio dataset is equal to a *Loadnumber* in the iNet TA dataset. The Proxio dataset contains 67,768 records, which implies that the CO_2e of 67,768 inbound transportations are calculated in the selected timeframe. Besides the *shipping_id*, eight other attributes are present in the Proxio dataset. A detailed description of all these attributes is included in Table A6. The most relevant attributes are explained in detail in the next paragraph.

Table A6

Attribute names, attribute definition and new attribute names in the Proxio dataset.

Attribute name	Definition	Renamed to:
Transport_date	Start date of transport.	-
Business_case	Type of transport. All entries have Inbound as business case.	-
shipping_id	Unique number to identify the transports in the Proxio dataset.	Loadnumber

distance	Combined distance for several shipments in the shipping_id.	-
Distance per load	The distance the load has travelled. This is the distance between the *Start_ID* and *End_ID* in the iNet TA dataset.	Not renamed
Sum of avg distance km leg	Average distance per leg in km's. A transport is divided in multiple legs in Proxio*	-
wtw_co2e	The amount of CO_2e equivalent emitted during transportation.	Not renamed

Note. – implies that the attribute is removed from the dataset. * A calculation in Proxio can be based on several legs. A leg is a part of the transportation and contains a specific type of vehicle or fuel. When moving goods from A to B, it can occur that somewhere during the ride the shipment is loaded on to a train. In these cases, not Diesel but electricity is considered. This is registered using different legs.

The main attribute of the Proxio dataset is *wtw_co2e. wtw_co2e* contains the amount of CO_2e emitted from well to wheel in Kg. The other attribute present in the dataset that is relevant for the research is *Distance per load*, which captures the distance in km from pick-up to drop-off. Table A7 contains an example of a record in the Proxio dataset after cleaning. Attributes present in the dataset that are not relevant for the analysis are removed. The attributes left in the dataset are *shipping_id, wtw_co2e* and *Distance per load*. Next, the attribute *shipping_id* is renamed as *Loadnumber* and *Distance per load* as *Distance_CO2e.*

Table A7

Example of a record in the Proxio dataset.

Loadnumber	Wtw_co2e	Distance per load (km)
1503629	137	1282

Pre-processing – Vista dataset.

The data exported from Vista contains names, addresses and coordinates (latitude and longitude) for every supplier in the system. The supplier-data is aggregated by *Key*, which is equal to the *Start_ID* from the iNet TA dataset. The entire dataset has 9559 records, but that does not equal the number of unique suppliers. An unknown number of suppliers has a double entry due to separate *Start_IDs* for the delivery of packaging, which are out of the scope of the research

The main attribute of the Vista dataset are the *Latitude* and *Longitude* coordinates. The other attributes in the dataset that are relevant for the research are the *Name,* which will be anonymized later, and the address, which can be constructed via *Street, ZIP Code,* and *City.* A detailed description of all these attributes is included in Table A8.

Table A8

Attribute names, attribute definition and new attribute names in the Vista dataset.

Attribute name	Definition	Renamed to:
Key	Unique string to identify a supplier in the dataset.	Start_ID

Name	Supplier Name	Name
City	Supplier’s city of residence.	-
ZIP Code	Supplier’s ZIP Code	ZIP_Code
Latitude	Supplier’s latitude coordinate	Not Renamed
Longitude	Supplier’s longitude coordinate	Not Renamed
Street	Supplier’s street name	Not Renamed

Note. – implies that the attribute is deleted from the dataset.

Pre-processing – Combined dataset.

The combined dataset is a combination of the dataset from iNet TA, Proxio and Vista. The CO_2e data from Proxio is merged with the iNet TA dataset based on the *Loadnumber*. Suppler details from Vista are merged with the iNet TA dataset using the Start_*ID*. The merged dataset contains 25 attributes and a total of 22,987 records. A variety of attributes are added for analytical purposes.

The attribute *lat_long* is added to the dataset to enable the researcher to exclude double coordinates. Hence, latitudes and longitudes can be equal, but the combination must be unique.

Some records are adjusted manually after error showed up in different section of the modelling. This was done to the following *ZIP Codes*: 514O1 (51401), ST. A (LL17 0JB) and DLI 4PW (DL1 4PW).

In total, 131 suppliers missed a *longitude* or *latitude* coordinate.

An extra attribute named *Address1* is added which consists of the street, ZIP Code, City and Country Code of the supplier. Similar addition is done named *Address2 but* does not include the street name and only *ZIP Code, City and Start Country. Address1* and *Address2* enable automated web-scraping. *Address2* is used for the records that still have missing coordinates after *Address1* was applied but was not able to retrieve coordinates. Geolocator in Python is used to automatically retrieve the missing addresses.

The addresses that are missing after applying Geolocator to *Address1* and *Adress2* are retrieved manually from the internet and are included via a separated DataFrame.

During the modelling phase, some errors in coordinates or other supplier details appeared. These are adjusted to:

Coordinates from Start_ID = 61039-01 are changed to 52.28087 (Latitude) and 4.76159 (Longitude).

ZIP Code from Start_ID = 82539-01 is changed to 565 2AE.

ZIP Code from Start_ID = 58588-01 is changed to 51401.

ZIP Code from Start_ID = 71872-01 is changed to 515400.

Start_IDs that are equal to MANUAL are excluded from the dataset. The same applies to *Start_IDs* that did not have any information attached that enabled retrieval (Such as missing addresses and name). This applies to four Start_IDs: 99-99-613521, 99-99-77840-02, 99-99-77840-02, 99-99-59088-01, 99-99-7784702.

Only Central Europe is included in the Analysis, Great Britain is excluded.

Appendix B – Results prior to discovering differences in CO_2e calculations

Table 4

CO_2e emissions and TonneKMs in the current situation, with weighted centroids and with unweighted centroids.

	Current Situation		**Weighted Centroids**		**Unweighted Centroids**	
Situation	Kg CO_2e	TonneKM	Kg CO_2e	TonneKM	Kg CO_2e	TonneKM
Pre-collection	3,231,962	67,986,932	2,051,456	35,545,190	1,297,736	22,531,350
Trunk loads	2,515,703	58,191,886	2,332,988	40,492,209	3,011,134	59,607,993
Total	5,747,665	126,178,818	4,384,444	76,037,399	4,308,870	82,139,343
% Difference	-	-	- 23.79	-39.74	-25.11	-34.90

Table 5

CO_2e emissions and TonneKMs using four of the cross-docking terminals currently in use that are closest to the locations proposed by the algorithm.

	Current Situation		**Weighted Centroids**		**Four current Cross-docking terminals**	
Situation	Kg CO_2e	TonneKM	Kg CO_2e	TonneKM	Kg CO_2e	TonneKM
Pre-collection	3,231,962	67,986,932	2,051,456	35,545,190	1,217,386	21,684,424
Trunk loads	2,515,703	58,191,886	2,332,988	40,492,209	2,089,475	50,724,594
Total	5,747,665	126,178,818	4,384,444	76,037,399	3,306,861	72,409,018
% Difference	-	-	- 23,79	- 39.74	- 25,11	- 42.61

Appendix C – Outcome of clustering for k=1 until k=10

Markers indicate the weighted cluster centroid (centre of gravity). This would be the proposed location for a cross-docking terminal when only pre-collection transport is considered.

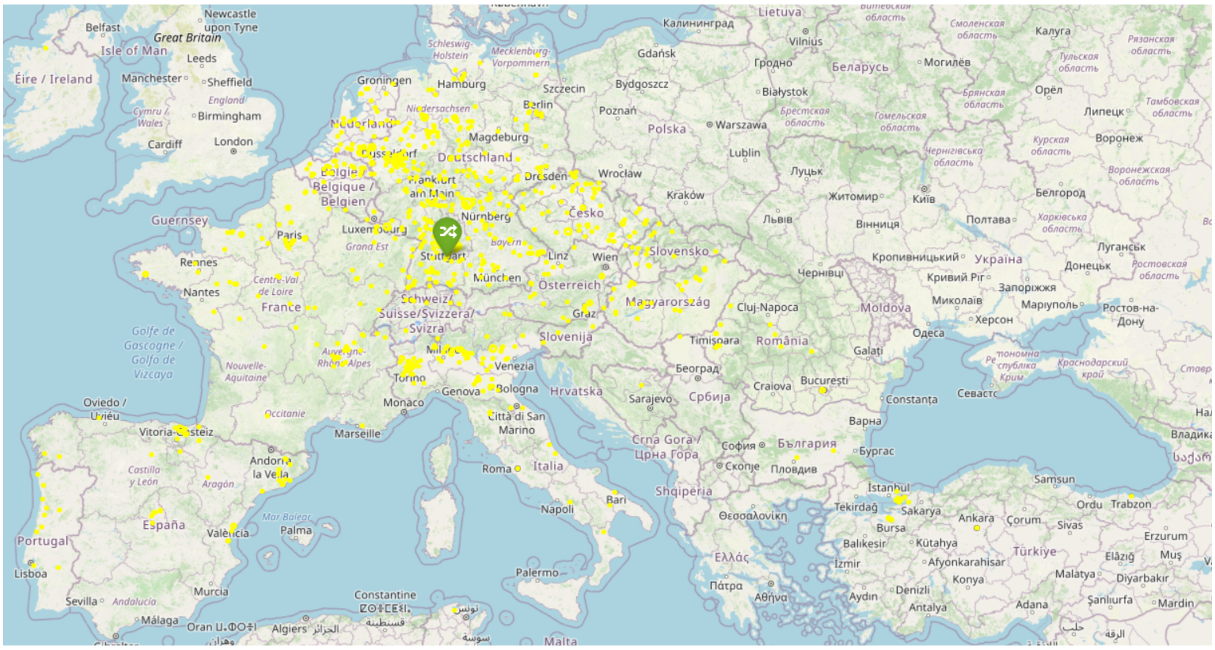

Figure 17: Weighted cluster centroid for k=1.

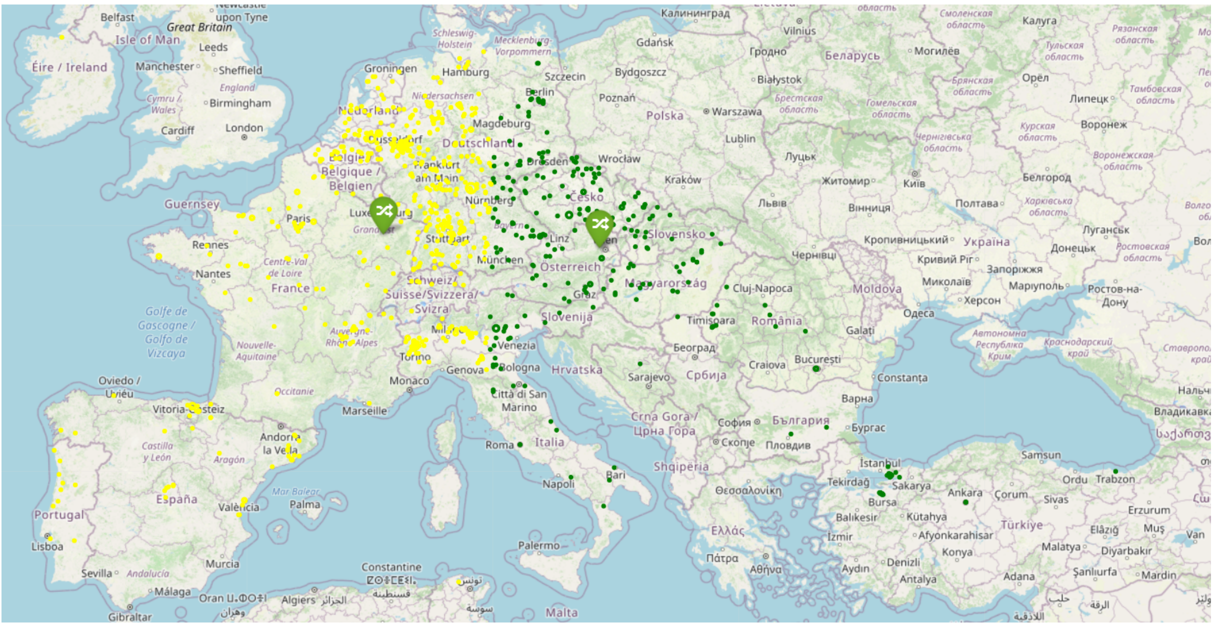

Figure 18: Weighted cluster centroids for k=2.

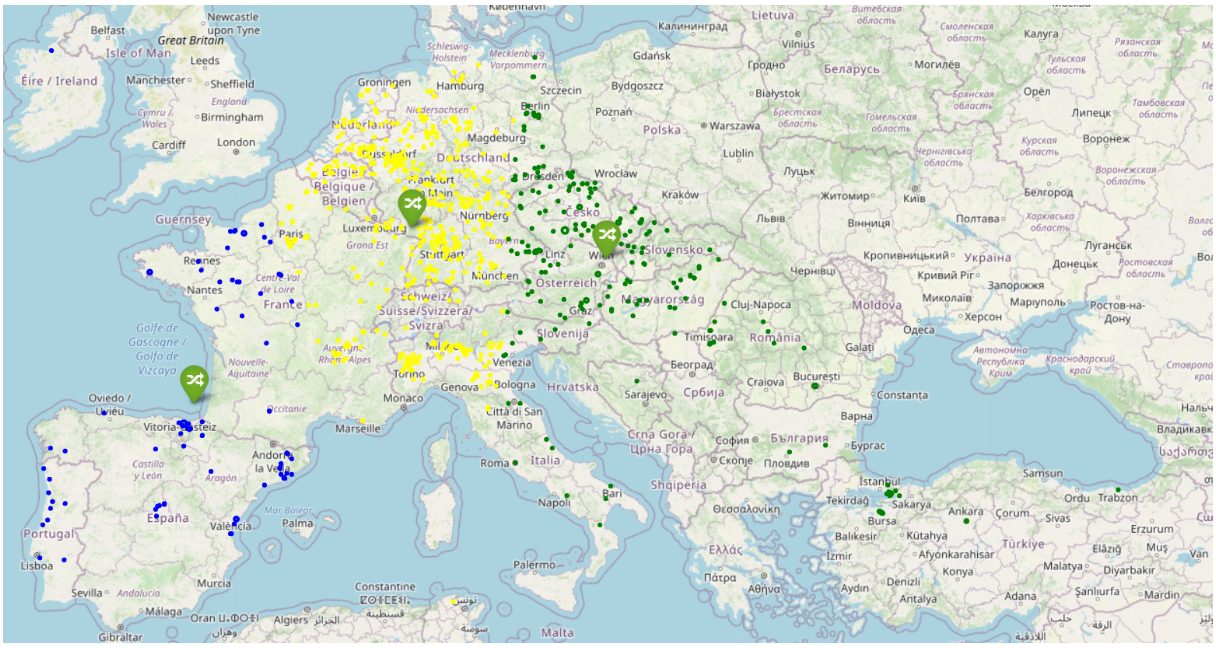

Figure 19: Weighted cluster centroids for k=3.

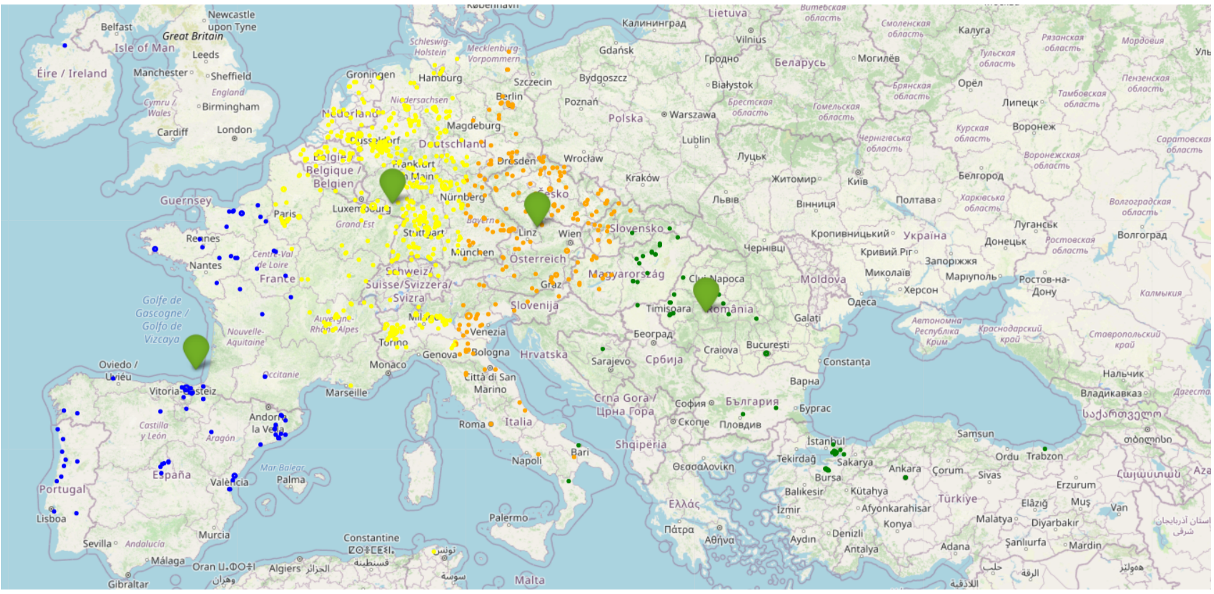

Figure 20: Weighted cluster centroids for k=4.

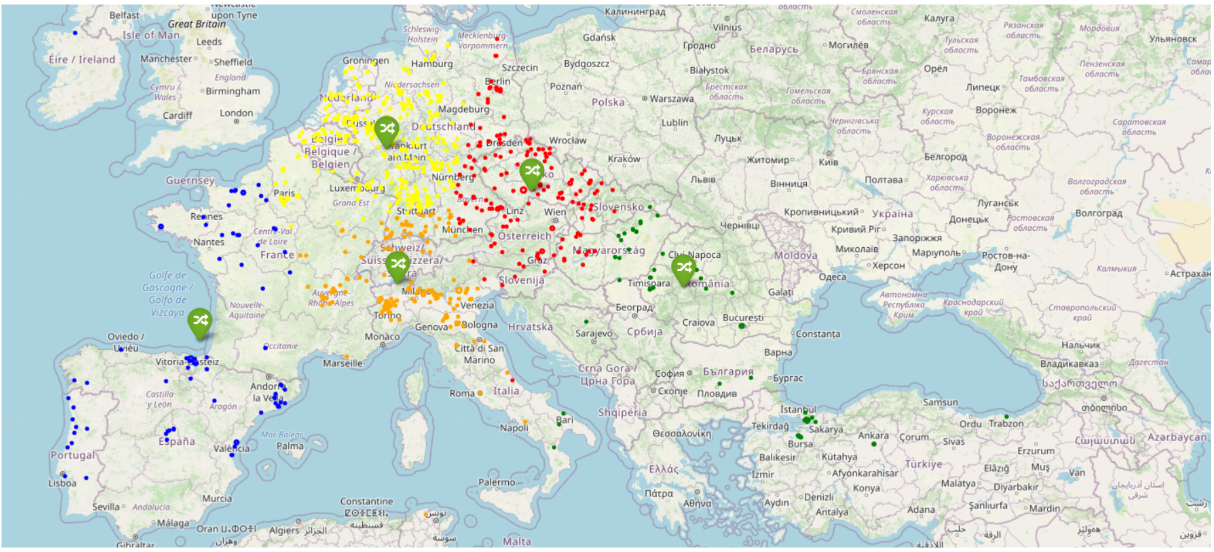

Figure 21: Weighted cluster centroids for k=5.

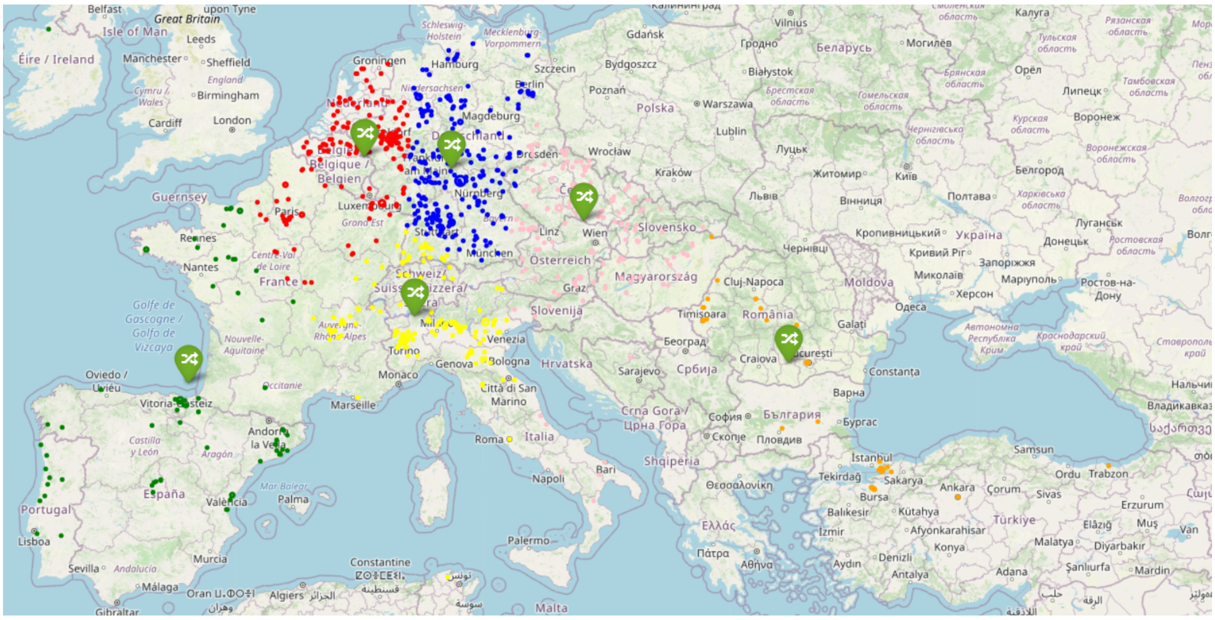

Figure 22: Weighted cluster centroids for k=6.

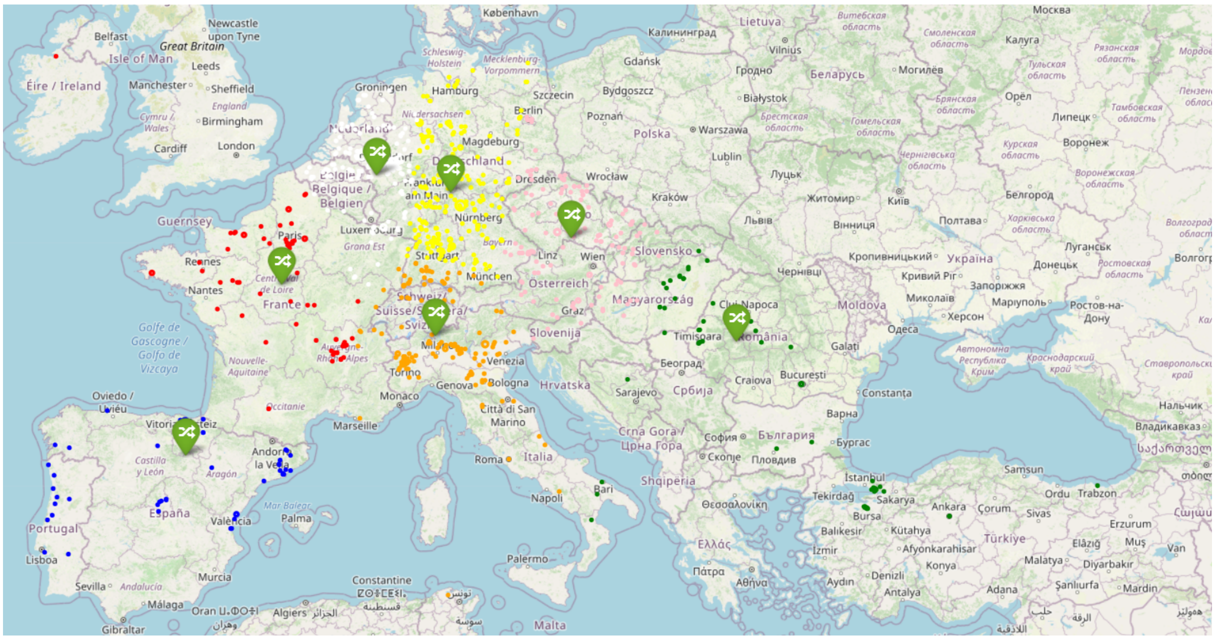

Figure 23: Weighted cluster centroids for k=7.

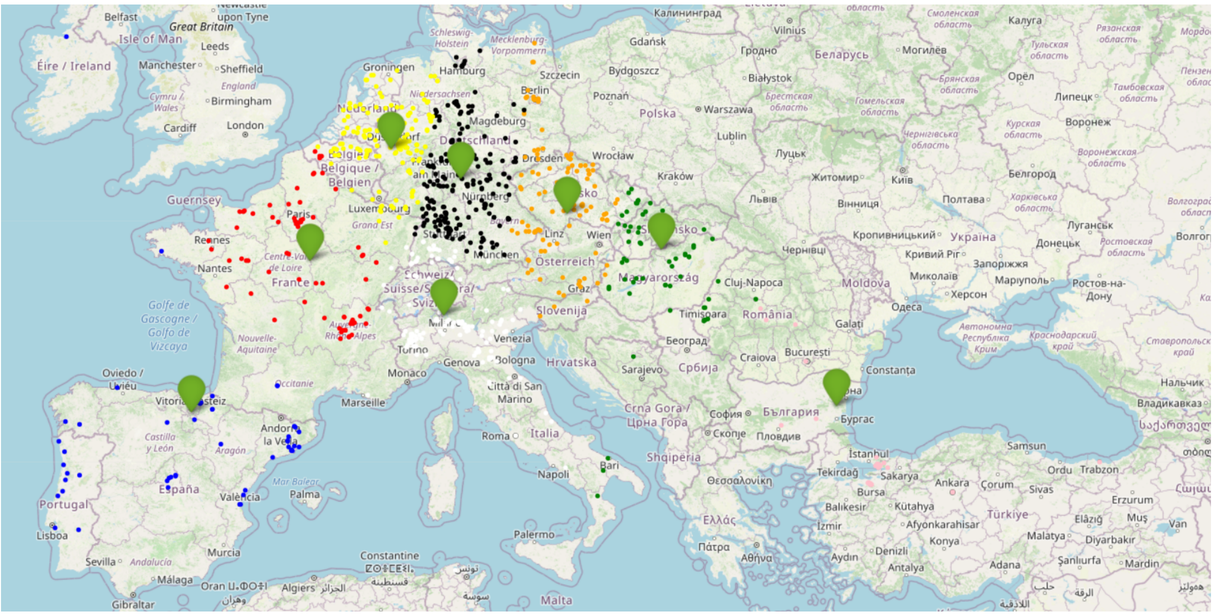

Figure 24: Weighted cluster centroids for k=8.

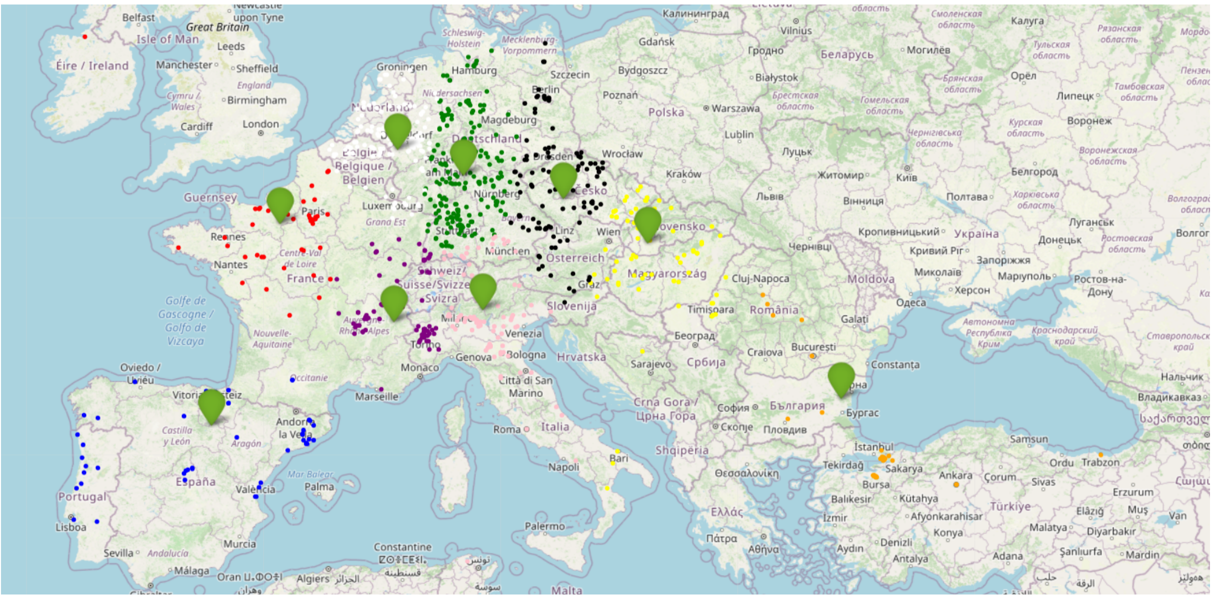

Figure 25: Weighted cluster centroids for k=9.

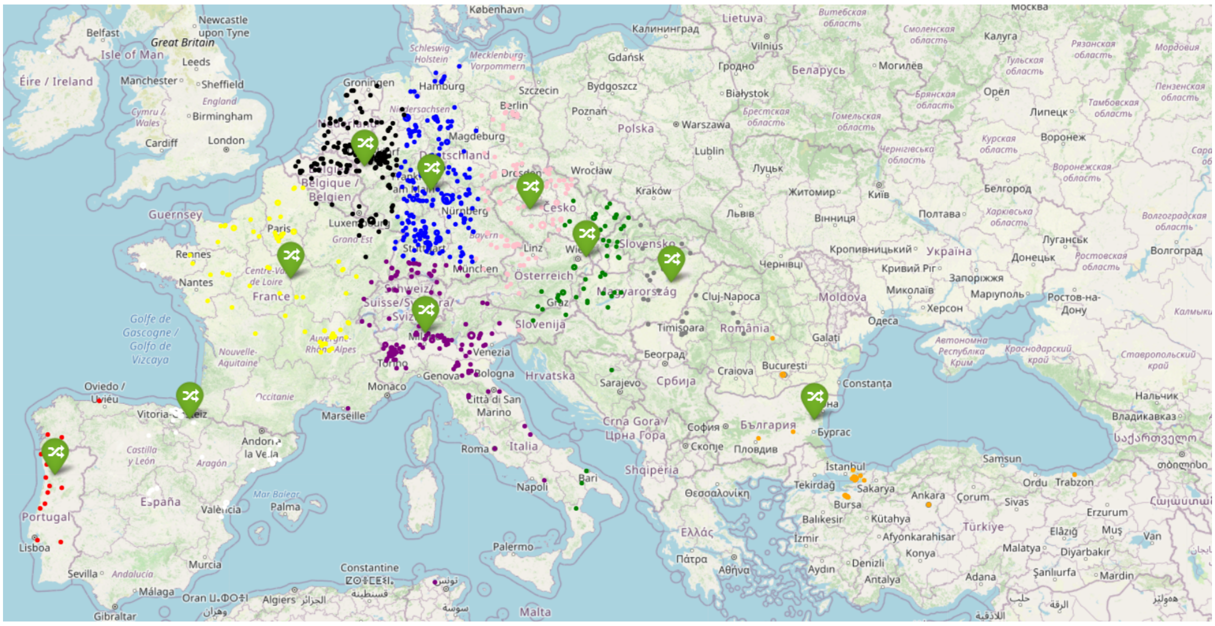

Figure 26: Weighted cluster centroids for k=10.