Master's degree in Supply Chain, Transport, and Mobility management

Operations research and other good practices for the design of a logistic network

Author: Anna Gabarró Sala
Supervisor: Maria Albareda Sambola
Tutor: Francesc Soriguera Marti
Call: September 2021

Escola Tècnica Superior
d'Enginyeria Industrial de Barcelona

ETSEIB

ABSTRACT

This project is intended to understand how to best use latest academic models and knowledge in the practical logistics world.

This is especially interesting in logistics, given the high level of objectivity of the problems' nature (compared to for example marketing, which is much more subjective). Reducing network inefficiencies is today increasingly important given that companies' success driver is currently swapping from demand to supply, and also considering the impact on planet's sustainability.

This thesis focuses on network design, a complex topic that today is optimized rather locally. The current general business trend here is to use simulation tools that provide best local alternatives assuming the rest of the network to be fixed. This approach fits in a stable environment where changes happen rather slowly and smoothly so that constantly reconsidering the full structure is not needed. Also, today's still existing lack of consistent and solid complete database triggers a necessity to manually obtain and summarize the needed input, hindering the extraction of the end-to-end picture and making those local scenario analyses the only feasible ones.

However, both trends are changing: proper digitalization is tackling the data bottleneck, while it is getting more and more clear that the VUCA (volatility, uncertainty, complexity and ambiguity) world we have experienced these last years is here to stay. Given those facts, we shall reconsider how to best pull data insights to provide powerful indications and help addressing management towards proper directions. In this regard, optimization models are proposed as a holistic complement of today's used local simulation ones.

The core of this project is the formulation of a linear integer program to optimally design a hub and spoke network to serve the Spanish customers from different EU origins (optimum balance between direct routing and shipment consolidation) and solve it on a small example data set. Given their high computational cost, exact integer programs are not suitable for real logistics problems of real scale. Therefore, the next steps would be to design an heuristic algorithm that, trading some controlled results precision for speed, could cope with the application of this model to real-life problems.

CONTENTS

1. GLOSSARY

- **Destination Zone**: the geographic area of large countries (i.e. Spain) is divided among different destination zones, based on destination zip codes.
- **EMEA**: Europe, Middle East, and Africa.
- **FTL (truck)**: Full TruckLoad (see Freightos, 2021). Type of trucking transport mode in which the journey is reserved for only one load/shipper. The shipper pays the full cost of the FTL trip [€]. *(See also LTL).*
- **FTL route**: route defined by unique origin and destination nodes that can be linked via an FTL truck.
- **LTL (truck)**: Less Than Truckload (see Freightos, 2021). Type of trucking transport mode in which the journey is split into many loads/shippers. The shipper pays the allocated cost of the LTL trip [€/kg]. *(See also FTL).*
- **OCP**: Origin Consolidation Point. Hub close to the factories that allows consolidating at origin.
- **PID** = **PackID**: indivisible shipment.
- **SC**: Supply Chain
- **VUCA**: Volatile, Uncertain, Complex, and Ambiguous.
- **Waybill**: group of shipments that travel together end-to-end (thus have same origin, destination, and routing).

2. INTRODUCTION

With the trend towards globalization experienced during the last decades, supply chains have become highly distributed with many different activities happening in a broad range of faraway geographies. Extraction of raw materials, manufacturing the components, assemblage, packaging, and final usage of the product by final customer, are a set of tasks that can be easily spread among different continents. Therefore, a proper strategy in logistics is very important for the overall company's competitiveness. This importance has been further exaggerated during 2020's Covid pandemic, which has empowered those companies with strong and efficient supply chains while nearly killed the ones with low—visibility or inefficient networks.

In order to improve the efficiency of a logistics network it can be very useful to have the right tools and techniques to deal with all data and extract from it insightful information that will help pointing company's management towards the right directions of the logistics strategy. This project is focused on identifying opportunities where academic models could help in this regard. After some interviews with people with different profiles and from diverse companies, it has been identified that today the logistic networks are mainly designed/operated via simulation models. In this regard, a potential area of opportunity could be to include also optimization models. Those optimization models can be either exact or heuristic. While exact models provide the exact optimum, they need a very high amount of computational effort, which disables them from real life's logistics applications. On the other side, heuristics provide only an approximated solution, but are much faster. The suggestion would therefore be to use heuristics models, but for their design and evaluation it is also needed to have first an exact one which will be used to frame the problem and also to tune the future heuristic program.

This project will further develop the diagnosis of today's company's network designs and their potential areas of improvement; and then focus on the development of an exact algorithm that optimizes the design of a hub & spoke network. This analysis will be based on subproblem of a multinational company, with several factories across Europe and delivery to the Spanish market (many different final destination customers, spread through Spain).

2.1. BACKGROUND AND MOTIVATION

After a solid period of academic studies, followed by several years of professional experience in logistics I have been surprised by how much potential but little leverage is there between the two worlds. Whilst any business will always fall into a very subjective environment, the same term "logistics" already indicates a discipline that after all derives from logics, very connected to algebra and the theoretical academic world. In this regard, it could be stated that this is an area where theoretical models may not have all answers, but for sure can strongly make use of data to provide proper directions, thus leading to better setups and reducing all type of costs: environmental, economic, time…

It is therefore a bittersweet situation to see the many advances happening at an academic level, which are adopted by private companies very slowly. Overwhelmed by many other core tasks, they do not find the time or resources to rethink their approaches and engage with each other, leverage efforts, and heavily improve their performances.

2.2. PURPOSE OF THE PROJECT

The aim of this thesis is, therefore, to work towards the improvement of such lack of mutual awareness and try to partially bridge the gap between the academic and the logistics business world. This includes a bidirectional approach: From one side it is needed that business take the necessary time and resources to invest towards finding new ways to work with data. From the other side it would also be helpful if the academia could also take some of their time to better understand general business requirements and adapt some of their research to address those needs. As shown in **Figure 1** both organizations could benefit from a larger alignment.

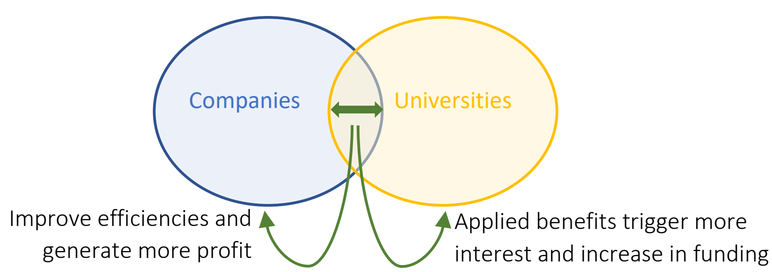

Figure 1. Benefits of leveraging logistics business with academia. Source: self-elaboration.

This project focuses on logistics network design and optimization. As detailed in the following sections, it has been observed that most business already make use of simulation tools, which can be very useful when evaluating a limited set of scenarios (i.e. different alternatives for a given distribution pattern, assuming the rest of the network is fixed). This approach seems to make full sense in a stable environment where the optimal network design rarely changes in time, thus allowing it to be defined by years of continuous improvements. However, the future seems to be all but stable, with ambiguity, uncertainty and rapid changes being the new normal.

Also, most companies still lack proper databases and instead have all their data scattered in many different places without consistency and end to end visibility. Input data is nowadays the main bottleneck, so model improvements would hardly bring any benefit. However, this data issue is currently being tackled with significant priority, hence the assumption is that this constraint will soon be solved and therefore open up the potential of investing towards more sophisticated data models that can provide better insights.

Figure 2 represents the different capabilities and requirements of simulation versus optimization tools.

	Simulation Tool *existing*	**Optimization Tool** **Would it be useful?**
Scope	Local / Isolated	Holistic
Evaluation of alternatives	One at a time	Combination
Conditions (volumes, rates, capacities…)	Stable	Fast and/or radically changing
Info (volumes, rates, capacities…)	Limited and/or unstructured	Digitalized
Computational power req.	Low	High

Figure 2. Simulation vs Optimization capabilities. Source: self-elaboration.

The first step right after setting up proper databases making all relevant data easily accessible is to develop new tools able to exploit these data and give quantitative support to decision making. This project is focused in this direction.

2.3. SCOPE

The major aim of this project is to explore optimization tools that can have real applications in logistics network design. In particular, in this project we have focused in the following activities:

1) Business framework: understand the current approach and identify potential upgrades. Simulation tools may not always be the best option, how about extending to optimization tools? What good would it do to know this optimal?

2) Designing of an exact model to optimize the hub and spoke of the logistics network of the considered company within Spain: balance between direct routing and consolidation benefits.

3) Providing guidelines on further developments to extrapolate the model to real life applications: heuristics, mix of continuous/integer programming, etc. This final section also includes a brief list of other existing initiatives that use academic models to improve logistics networks.

3. BUSSINESS FRAMEWORK - OPPORTUNITIES AND LIMITATIONS OF OPTIMIZATION MODELS

To understand a generic trend of companies' approach to logistics Network Design, I have conducted a series of interviews with logistics' team members of different companies (including multinationals, freight carriers, Supply Chain consulting, and even smaller startups with local delivery services). All of them follow a similar pattern and although data is used to support decisions, it falls far short from its potential. For confidentiality reasons, no more detail on the identities of the interviewees can be given.

In this project we focus in the particular case of a multinational with about 10 points of origin (factories and ports of entry from Asian factories); and thousands of customers all over EMEA (Europe, Middle East, and Africa).

3.1. Current approach and its limitations: sequentially solving subproblems yields suboptimal solutions

Although relying on computer tools, in practice, the logistics network *optimization* process is usually based on the human brain. There are a series of built-in *simulation* algorithms that are very useful when choosing among a few alternatives (you simulate each of them and keep the one that gives the best results according to your criteria). One can also extract information from data analysis and the simulation of "extreme" scenarios, which are useful to discard some options.

But a holistic optimization of the entire EMEA logistics network results in an exponential number of possible scenarios, impossible to handle from a human point of view.

Thus, one ends up deciding on a first "layer" based on common sense, experience, and existing infrastructure, and then improving small sections separately. The result is a network that a first glance looks reasonable, but with no optimality guarantees. That is, there might be other choices that could be more efficient. In addition, there is the conditioning of existing contracts and infrastructure that limit the number of alternatives and facilitate the decision. However, these trap decision makers in "temporary optima" (that keep changing as some contracts run out and new ones have to be set). Instead, it would be interesting to think about the global optimum, since it is well known that sequentially solving subproblems yields suboptimal solutions (see Darvish and Coelho, 2018): if there was a blank page, what would we do?

The complexity of devising such optimal solutions is heavily increased by the following aspects of the current business context:

- VUCA world (volatility, uncertainty, complexity, and ambiguity): inputs vary constantly, with radical changes becoming more and more frequent.
- Still very rigid systems (every small change needs an implementation time of at least 1-3 months).
- Currently, most companies also have a lack of a proper solid and consistent end-to-end databases. This is one of the most common and critical bottlenecks, however it is already being considered and treated with high priority. Hence, the assumption is that short term (quality) data should be a common asset.

All these issues present very interesting challenges within the field of mathematical optimization, both in terms of the dimension of the posed problems, and in terms of the complexity associated with decision problems under uncertainty. This has motivated a great bunch of research in the last years (see, for example, Lekić et al., 2021 or Pires, Parreira and Frazzon, 2021).

Clearly, the most common approaches to these problems are either the sequential optimization of different subproblems, which, as mentioned above, yields suboptimal solutions, and the use of heuristics and simulation (see, for instance, Tordecilla et al. 2021).

3.2. On the value of identifying optimal solutions

The main utility of knowing what is the optimal decision for a particular scenario is to have a basis for comparison. That is, an optimal solution can be used as a reference, to understand what is the gap between the decisions that are made under the current company policies and the best possible ones. Given the context of volatility together with the hassle and slowness of implementing changes, knowing this optimum is a long way from being implementable. But the ability to quickly determine what the best design would be given a particular set of conditions, even if it serves only as a reference, can be very powerful.

Knowledge about optimal solutions can also be useful to evaluate the benefits of possible new operating systems (dynamic optimization, autonomous decisions of digital twins, collaboration with other companies, ...).

3.3. Closing the gap

Based on those findings, it could be advisable for the companies to count on heuristic methods capable to deal with real instances (large sets of data). When developing them, it is convenient to first start with exact algorithms that help evaluating and calibrating the heuristic ones. Also, the initial construction of mathematical programs to model the situation helps to frame the problem and detail the questions that we would like to have answered. Considering the expected extension of this master project, we have stopped after the development of the mathematical program, and provided some guidelines on its limitations and potential further developments. So, the development of heuristic methods for the considered problem is left out of the scope of this project.

4. THE STUDY CASE

The particular optimization problem addressed in this project is a logistics network operation problem defined as follows: a list of shipments have to be distributed from a certain origin factory to a known set of destination customers given certain timing constraints. The goal is to do it using the best combination of the different shipment possibilities available in a hub and spoke network so that the total distribution cost is minimum. To simplify, in this project we have considered only the economic cost (€). However, it must be remarked that the cost can be any quantified parameter (Co2eq, €, km, or also even a subjective measure such as service quality to which different weights are assigned depending on business priorities).

For evaluating the computational burden associated with the obtained formulation, we will evaluate it on the particular case of a fraction of the above-mentioned company that comprises some factories located around Europe, and customers in Spain.

4.1. Study case description

The main elements required to define a particular problem instance are the list of shipments that need to be made and all the alternative possible ways to perform the shipments. More detail is given next.

The problem to solve is defined on an abstract network. The nodes of the network are divided into three subsets, one representing the set of factories where the goods are going to be shipped from, the second one representing the set of customers who have to receive the goods, and a third set of nodes representing possible hubs where shipments can be consolidated or reorganized.

The shipment orders and their corresponding characteristics are described in Section 4.1.1. The links of the above network, and the available transportation modes are defined in Section 4.1.2.

4.1.1. Input Data – Shipment orders

For a fixed network, instances are basically defined by a list of shipments, identified by their Pack IDs (PIDs). Each PID is characterized by its origin factory and a destination customer. On top of its origin and destination, each PID has assigned a weight and available shipment date (departure from the factory). The PIDs are also indivisible. Whenever there are several PIDs that are to be shipped from the same origin factory to the same destination customer on the same truck, those can be grouped in the so called "waybills", which physically relate to one or more pallets full of PIDS that will travel together end to end.

Also, as the list of customers is very large, they are grouped in different zones depending on their zip code. Hence, we have divided the area of Spain into 10 zones, which will determine different pricing rates.

4.1.2. Input Parameters - Available Network description

Next, we describe the costs and constraints of all potential route legs and nodes of the hub and spoke network considered in this study case. The proposed formulation could be easily extended to different network shapes, if the distribution constraints are kept. In this study case, on top of origins and destinations, we have considered four additional nodes:

One origin consolidation hub (OCP = O), close to the origin factories.

Three country destination hubs: in Barcelona (BCN), in Madrid (MAD), and in Zaragoza (ZAR).

The resulting network, including the available arcs, is depicted in Figure 3.

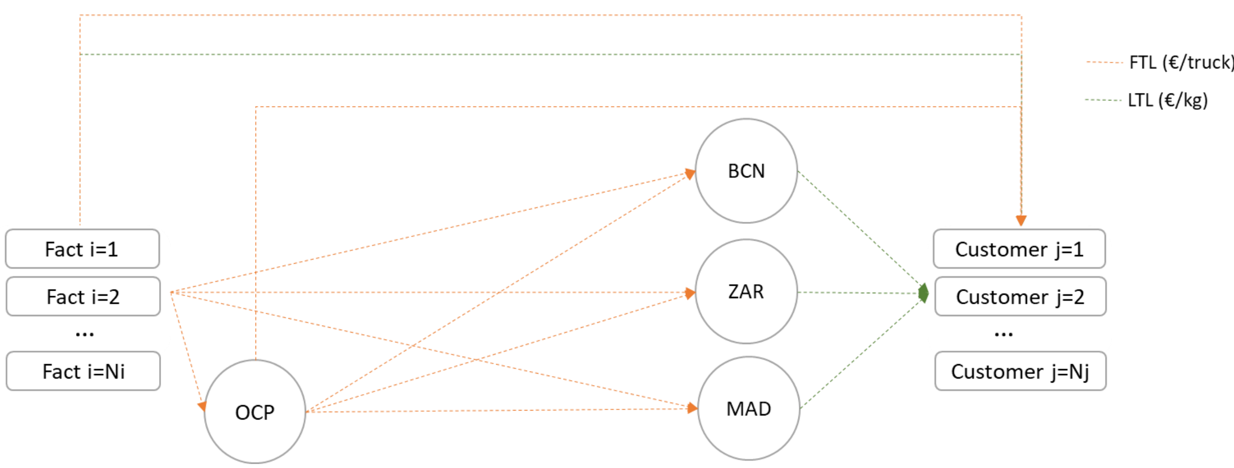

Figure 3. Graphic visualization of the modelled study case. Source: self-elaboration.

Note that, formally, the network depicted in Figure 3 is a multigraph. This is because there are two different transportation options and, in some cases, they are both available for a given pair of nodes of the network. As it is usual in the company that served as a basis for defining this problem, two types of shipments are possible: full-truckload (FTL) shipments, for which a fixed cost per shipment is paid, and less-than-truckload (LTL) shipments whose costs depend on the shipped weight.

We next explain in detail all the costs involved in our distribution problem.

- Handling costs (€/kg), only charged at the OCP: *OO*
- FTL trucks: we pay a fixed amount per truck (€/truck), i.e. the full truck capacity is utilized by the company's products.

  FTL route: defined by unique origin and destination nodes that can be linked with an FTL truck. The *total number of arcs corresponding to FTL routes is Ni*Nj + 4Ni + Nj + 3*, where *Ni* is the total number of origin factories and *Nj* the total number of destination customers.

  *Ni*Nj: FTL direct routes from each factory to each customer*

  +

  *4*Ni: from each factory to the 4 hubs (OCP, BCN, MAD, ZAR)*

  +

  Nj: direct from OCP to each customer

  +

  3: from OCP to the 3 destination hubs (B, M, Z)

  The corresponding cost, D_{ip}, depends on the origin *i* and destination zone, *p.*

- LTL trucks: we pay per kg (€/kg), i.e. the truck is split among other companies, so we pay only for a certain allocation. This €/kg rate is a function of the waybill weight, with decreasing cost per kg as we increase the waybill weight. Moreover, there is a minimum cost for waybills up to a minimum weight and a maximum cost per shipment. Figure 4 shows an example of this cost function, which is piecewise linear:

Min Cost	*10 €*
0-30kg	*1 €/kg*
30-200kg	*0.8 €/kg*
…	
Max Cost	*4000 €*

Figure 4. Common structure of LTL rates quotation. Source: self-elaboration.

To include this quoting in the model the function has been linearized, as explained in *section "4.2.1. Linearization of the LTL cost function"*.

<u>Constraints and other parameters to define the network operation:</u>

- Capacity of the trucks: The capacity of the trucks, *C*, indicates de maximum total weight a truck can hold, and is the same for all trucks.
- Maximum time that shipments can wait at factory or at the OCP: Again, this maximum time is independent of the node where the waiting time takes place. We denote it by *WF* and *WO*.
- Routing times from the factories to the OCP (to properly define continuity constraints on its inbound and outbound flow: O_i.

Note: we establish time constraints involving a maximum waiting allowance at the origin factories or at the hubs. Therefore, the routing times past those points are not needed.

4.1.3 Decision variables for the formulation

The intent is to find the routing of each PackID (i.e. the set of arcs and nodes that will bring each PID from its origin to its destination) so that the overall cost is minimum.

To do so, for each arc we have defined a set of binary routing variables that get activated (=1) if a certain PID (indiced by *k*) is using it with a certain truck (indiced by *l)*, otherwise they have 0 value, i.e.:

d_{kl} = 1 if packid *k* travels from its origin factory directly to its destination customer on the *l-th* truck

d_{kl} = 0 otherwise

Variables f_{kl}, od_{kj}, b_{kl}, z_{kl}, m_{kl}, o_{kl}, ob_{kl}, oz_{kl} and om_{kl} are defined analogously, to represent shipments through the other arcs of the network, as represented in Figure 5.

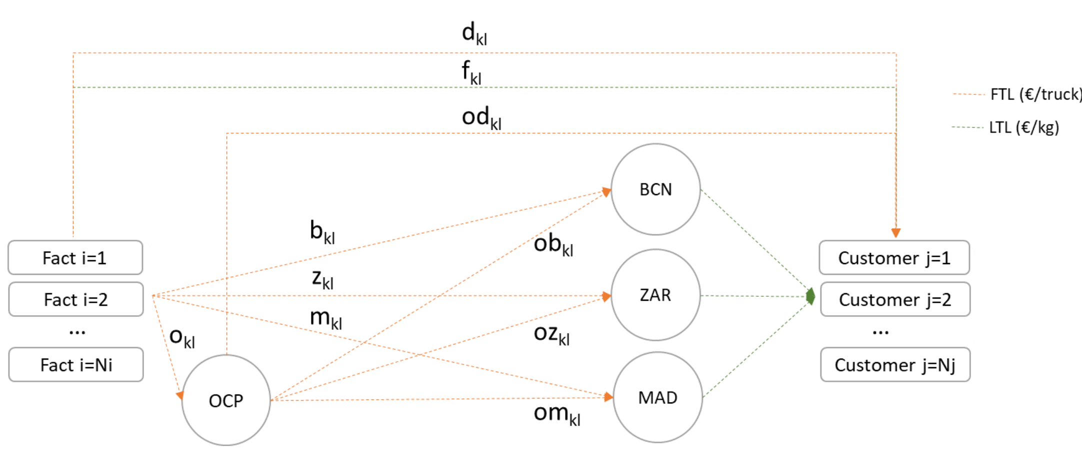

Figure 5. Graphic visualization of the network, including the routing variables (9 different options). Source: self-elaboration.

A set of constraints will ensure continuity (everything entering a hub needs to exit it) and limits on capacity and waiting times. Those are detailed in the section "*4.3. Model Formulation*". Note that the last LTL legs from country hubs to the customer do not have any routing variable assigned, as they are already defined by the inbound flow, i.e. all packids entering the BCN hub (b_{kl} = 1 or ob_{kl}=1) will travel on the last LTL leg from the BCN hub to the customer.

The solution of the problem is therefore determined by the 9 "routing" variables shown in **Figure 5**. However, for modelling purposes we have added two new sets of binary variables, the so called "prima" and the "departure" ones. Summarizing, we have:

- Routing variables: determine whether a packid *k* travels through this route on the *l*–th truck.

- Prima variables: determine if we get to use the l^{th} truck of the route. This is the variable that will be used in the minimization function to calculate the total FTL cost.
  $d'_{i=1, j=1, l=3}$ = 1 means we have used at least 3 FTL direct trucks from factory *i=1* to customer *j=1*.
  $ob'_{l=2}$ = 1 means we have used at least 2 trucks from the OCO to the BCN hub.

- Departure variables: determine on which day trucks depart from their origin.
  $dDept_{i=1,j=1,l=1,s=1}$ = 1 if the first (*l=1*) FTL direct truck from factory *i=1* to customer *j=1* departs the factory on the day *s=1*.
  $obDept_{l=1,s=1}$ = 1 if the first (*l=1*) truck from the OCP to the BCN hub departs the OCO on day *s=1*.

4.1.4. Objective Function

The goal of the model is to minimize the total cost:

Min (total cost) = Min (cost of FTL trucks + cost of LTL + handling costs)

<u>Cost of the FTL trucks</u>

With the defined prima variables, we can calculate the total cost spent on FTL trucks:

- FTLcost: cost of all FTL direct trucks from factory *i* to the customers in zone *p.*
- FTXcost: cost of all FTL trucks from factory *i* to the country hubs (BCN, MAD, ZAR).
- toOCPcost: cost of all FTL trucks from factory *i* to the origin consolidation hub (OCP).

- OFTLcost: cost of all FTL direct trucks from the OCP to the customers in zone *p*.
- OFTXcost: cost of all FTL trucks from the OCP to the country hubs.

Below a couple examples on the formulation of these costs are given, the rest can be found in section *"4.3.3. Model formulation – Objective function".*

$$FTLcost = \sum_{\substack{ijpl \\ zone(j)=p}} D_{ip} * d'_{ijl} \tag{1}$$

$$OFTXcost = OB * \sum_{l} ob'_{l} + OM * \sum_{l} om'_{l} + OZ * \sum_{l} oz'_{l} \tag{2}$$

<u>Cost of the LTL trucks</u>

All the PIDS that are not routed on a FTL customer direct truck will have a last leg charged at a €/kg cost based on their waybill weight. Section *"4.2.1. Linearization of the LTL cost function"* provides detailed explanation of the calculation of this LTL cost:

Minimum costs (applies if a waybill exists, i.e. has a weight >0).

$$mincost = mincost\boldsymbol{f} + mincost\boldsymbol{b} + mincost\boldsymbol{ob} + mincost\boldsymbol{m} + mincost\boldsymbol{om} + mincost\boldsymbol{z} + mincost\boldsymbol{oz}$$

Adder costs (difference between the min. cost and the actual waybill cost):

ltl**dir** = LTL adder from factory *i* to a customer in zone *p*.

ltl**bcn**= LTL adder from the BCN hub to a customer in zone *p*. Considers those PIDS originally routed in FTL trucks from factory to BCN hub.

ltl**mad** = LTL adder from the MAD hub to a customer in zone *p*. Considers those PIDS originally routed in FTL trucks from factory to MAD hub.

ltl**zar** = LTL adder from the ZAR hub to a customer in zone *p*. Considers those PIDS originally routed in FTL trucks from factory to ZAR hub.

ltl**obcn** = LTL adder from the BCN hub to a customer in zone *p*. Considers those PIDS originally routed in FTL trucks from factory to OCP and then to BCN hub.

ltlomad = LTL adder from the MAD hub to a customer in zone *p*. Considers those PIDS originally routed in FTL trucks from factory to OCP and then to MAD hub.

ltlozar = LTL adder from the ZAR hub to a customer in zone *p*. Considers those PIDS originally routed in FTL trucks from factory to OCP and then to ZAR hub.

<u>Handling costs</u>

In this study case we have assumed there are only handling charges at the origin consolidation point (OCP):

$$Ohandlingcost = \ OO * \sum_{kl} w_k * o_{kl}\,, \qquad (3)$$

where *OO* are the handling costs [€/kg], and the sum covers the weight [kg] of all PIDs routed via the OCP.

4.2. AUXILIARY VARIABLES TO SUPPORT FORMULATION

Including non-linearities in a mathematical programming formulation results in formulations typically much harder to solve, since they require the use of more involved methods than those applicable to problems where all functions are linear. For this reason, it is a common practice to linearize any nonlinear parts in the problem, whenever possible, even if it is at the cost of introducing additional auxiliary variables that are not directly related to a decision.

This section contains the details of the linearizations that have been used in the study case.

4.2.1. Linearization of the LTL cost function

As explained previously, the LTL rate structure consists on a range of unit costs (€/kg) depending on the bin range of the waybill weight. The structure of this cost function is represented in Figure 6.

Waybill weight	*LTL Rate*
minimum cost	*10 eur*
0-30kg	*1 eur/kg*
30-200kg	*0.8 eur/kg*
…	
Max Cost	*4000 eur*

equivalent to →

Wb weight [kg]	*Intervals [kg] (def. by upper bound)*	*LTL Rate*
$0 - t_0$	$tram_{n=0}$	$rate_{n=0}$ *=min cost [€]*
$t_0 - t_1$	$tram_{n=1}$	$rate_{n=1}$ *[€/kg]*
$t_1 - t_2$	$tram_{n=2}$	$rate_{n=2}$ *[€/kg]*
…	…	…
$t_{Nn-1} - t_{Nn}$	$tram_{n=Nn}$	$rate_{n=Nn}$ *[€/kg]*

Figure 6. Structure of the LTL cost function. Source: self-elaboration.

Below, **Figure 7** represents the associated LTL cost graph. Note that the row {$rates_n$} has a first value ($rate_{n=0}$) that represents an absolute value [€] equal to the minimum cost, while all the rest ($rate_{n=1, \dots, Nn}$) refer to the slope of the cost function [€/kg]. These slopes decrease as the waybill weight increases. Therefore, except for the initial segment of the minimum cost, the LTL cost function is concave. There is also a maximum cost [€], defined by the last point of the x-axis: $tram_{n=Nn}$.

A different set of LTL rates is defined for each:

- Origin point of the LTL route: any of the three country hubs or the *Ni* origin factories.
- Destination zone *p*. Spain is divided into 10 zones depending on the postal code of the customer.

Therefore, as input parameters we will have all the following value sets, where *i* defines the origin factory, *p* the destination zone of the customer, and *n* the tram of the waybill weight bin range. Note that the tram value equals the upper bound of the waybill weight's bin range:

$tramBB_{pn}$ *and* $rateBB_{pn}$

$tramMM_{pn}$ *and* $rateMM_{pn}$

$tramZZ_{pn}$ *and* $rateZZ_{pn}$

$tramFF_{ipn}$ *and* $rateFF_{ipn}$

Figure 7 provides a visual example (from BCN hub) of the LTL linearized function logic. It also includes two-point examples of alpha values, a new variable explained right below.

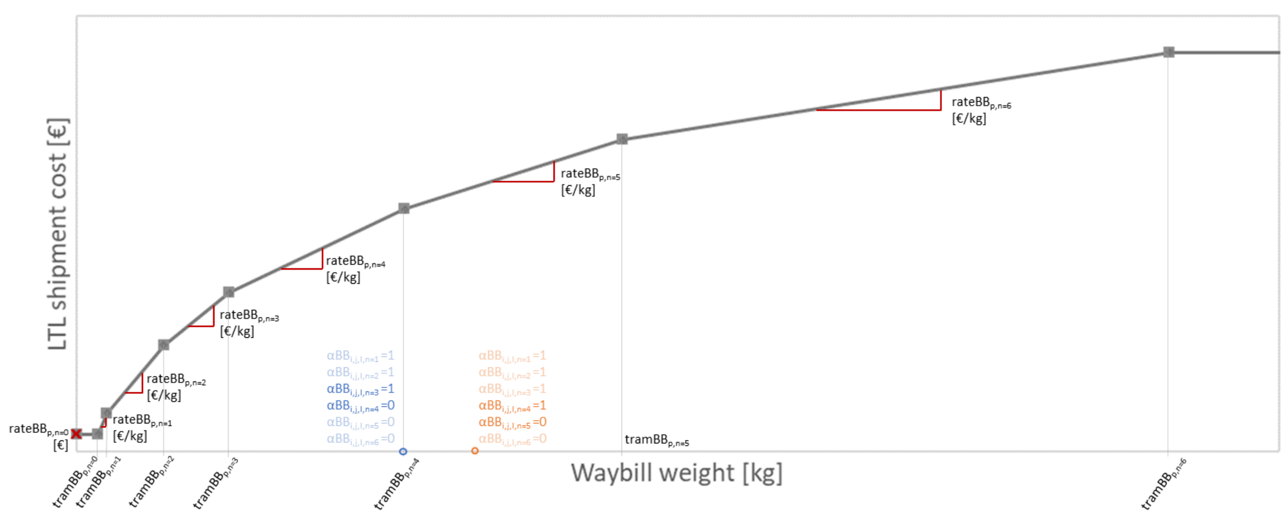

Figure 7. Graphic visualization of the LTL function and the alpha variables.

To obtain the cost of each waybill based on its weight, we have defined two new sets of binary variables:

- Alpha variables α_{ijln}: determine if the waybill weight exceeds the upper limit of the n-th cost interval, $tram_n$. They also contain the indices related to factory, customer, and truck. As defined previously, a waybill consists of all packIDs travelling together e2e (i.e. same factory, origin, and truck).

  Taking as an example the PIDS routed from factory to BCN hub with an FTL, and then from BCN hub to the customer with the LTL for which we need to define the cost function:

$$\alpha BB_{ijln} = 1 \Leftrightarrow \sum_{\substack{k\ tq \\ fact_k = i \\ customer_k = j}} (w_k * b_{kl}) > tramBB_{pn} \quad \forall\, p, l, n \quad zone_j = p \tag{4}$$

  which is equivalent to the combination of the following two inequalities,

$$tramBB_{pn} + \left(C - tramBB_{pn}\right) * \alpha BB_{ijln} \geq \sum_{\substack{k \\ fact_k = i \\ customer_k = j}} w_k * b_{kl} \qquad \forall i, j, p, l, n \quad st. \quad zone_j = p \tag{5}$$

$$\sum_{\substack{k \\ fact_k = i \\ customer_k = j}} (\mathrm{w}_k * b_{kl}) \geq (\mathrm{tramBB}_{pn} + \varepsilon) * \alpha\mathrm{BB}_{ijln} \forall\, i, j, p, l, n \quad zone_j = p \tag{6}$$

  being C the truck capacity defined in the input parameters and ε an auxiliary small value.

  Note that equation (6) should be a strict inequality. However, standard optimization methods do not allow them. Therefore, it has been converted to a non-strict inequality by adding this small enough value ε. (To fix its value, it is enough to use any value a bit below the accuracy of the weights of the different orders).

  We assume no consolidation capabilities in the hubs: the only place where we allow to consolidate PackIds into bigger waybills is at the origin factory, so if two PIDS originating from different factories and traveling to the same customer meet in a hub, they would not be consolidated together. This means that the last LTL leg would still consider two fully different shipments, resulting in a higher cost:

LTL cost (shipment 1) + LTL cost (shipment 2) ≥ LTL cost (shipment 1 + shipment 2)

- Waybill variables WV_{ijl}: determine if a waybill uses a certain LTL route (=1 if yes, =0 if no). Going back to the BCN example (defined by the routing variable b_{kl} that will force an LTL leg out of the Barcelona hub):

$$\sum_k (\mathrm{w}_k * b_{kl}) \geq 0 \Rightarrow WVb_{ijl} = 1 \quad \forall i,j,l \quad st.\ fact_k = i \quad and \quad customer_k = j \qquad (7)$$

Note that a restriction on the other direction (⇐) is not needed as we are addressing a minimization problem, and it will never pay to set WVB_{ijl} to 1 if it is not really necessary.
Equation (7) can be expressed as the following constraint:

$$WVb_{ijl} * \sum_k (\mathrm{w}_k * b_{kl}) \geq \sum_k (\mathrm{w}_k * b_{kl}) \quad \forall i,j,l \quad st.\ fact_k = i \quad and \quad customer_k = j \qquad (8)$$

However, this has a product of variables, so it has been linearized by creating a new binary variable:

$$L_{WVb_{b\,kl}} = WVb_{ijl} * b_{kl} \qquad \forall i,j,k,l\ st. fact_k = i \ \ and\ customer_k = j \qquad (9)$$

More details in section *"4.2.2. Linearization of binary variables products"*. Finally, we can now write:

$$\sum_k \left(\mathrm{w}_k * \mathrm{L}_{\mathrm{WVb}_{\mathrm{b}\,kl}}\right) \geq \sum_k (\mathrm{w}_k * b_{kl}) \qquad \forall l \qquad (10)$$

Summarizing, for each waybill (fixed indices *i,j,l*) the *WV* variable determines whether we apply the minimum LTL cost, while the row of alphas {α_n} is used to calculate the cost increase through the *n* intervals based on the waybill weight. With those variables we can define a piecewise function to calculate the LTL costs.

Taking as an example the LTL costs related to all PIDS that went from the factory to the BCN hub in a FTL truck, and then from BCN to customer via LTL trucking:

Minimum cost:

$$mincostb = \sum_{\substack{ijpl \\ zone_j=p}} \left[rateBB_{p,n=0} * WVb_{ijl}\right] \tag{11}$$

Adding the cost increase based on waybill weight:

$$ltlbcn = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha BB_{ijln} * rateBB_{pn} * \left(tramBB_{p,n} - tramBB_{p,n-1}\right)$$

$$+ rateBB_{pn} * \left(\alpha BB_{i,j,l,n-1} - \alpha BB_{ijln}\right) * \Big(\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (b_{kl} * w_k) - tramBB_{p,n-1}\Big)] \tag{12}$$

First line of equation (12) calculates the cost adder of all the intervals that the waybill weight has fully completed. It activates for all the *n* intervals with α_n=1 (meaning the waybill weight exceeds them), and then does: rate * ($tram_n$-$tram_{n-1}$) = slope * distance in the x-axis = [€/kg *kg] = [€].

Second line of the equation (12) calculates the cost of the remaining weight. It activates when the weight has surpassed the *n-1* interval, but has not achieved the upper limit of the *n-th* interval (i.e. α_{n-1} =1 and α_n =0). Then: rate * (current weight – upbound of last surpassed interval) = slope * distance in the x-axis = [€/kg * kg] = [€].

To facilitate the understanding, **Figure 8** provides a graphic visualization of the equation (12).

Note that equation (12) has again multiplication of variables, for which we have defined a new variable:

$$L_{alphas_{b_{kl}}} = b_{kl} * \left(\alpha BB_{ijln-1} - \alpha BB_{ijln}\right) \quad \forall\, i,j,k,l,n\; s.t.\, fact_k = i, customer_k = j, n > 1 \tag{13}$$

Finally, we can now write the linear equation:

$$ltlbcn = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha BB_{ijln} * rateBB_{pn} * \left(tramBB_{p,n} - tramBB_{p,n-1}\right)$$

$$+ rateBB_{pn} * (\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (L\_alphas\_b_{kln} * w_k) - (\alpha BB_{i,j,l,n-1} - \alpha BB_{ijln}) * tramBB_{p,n-1})] \quad (14)$$

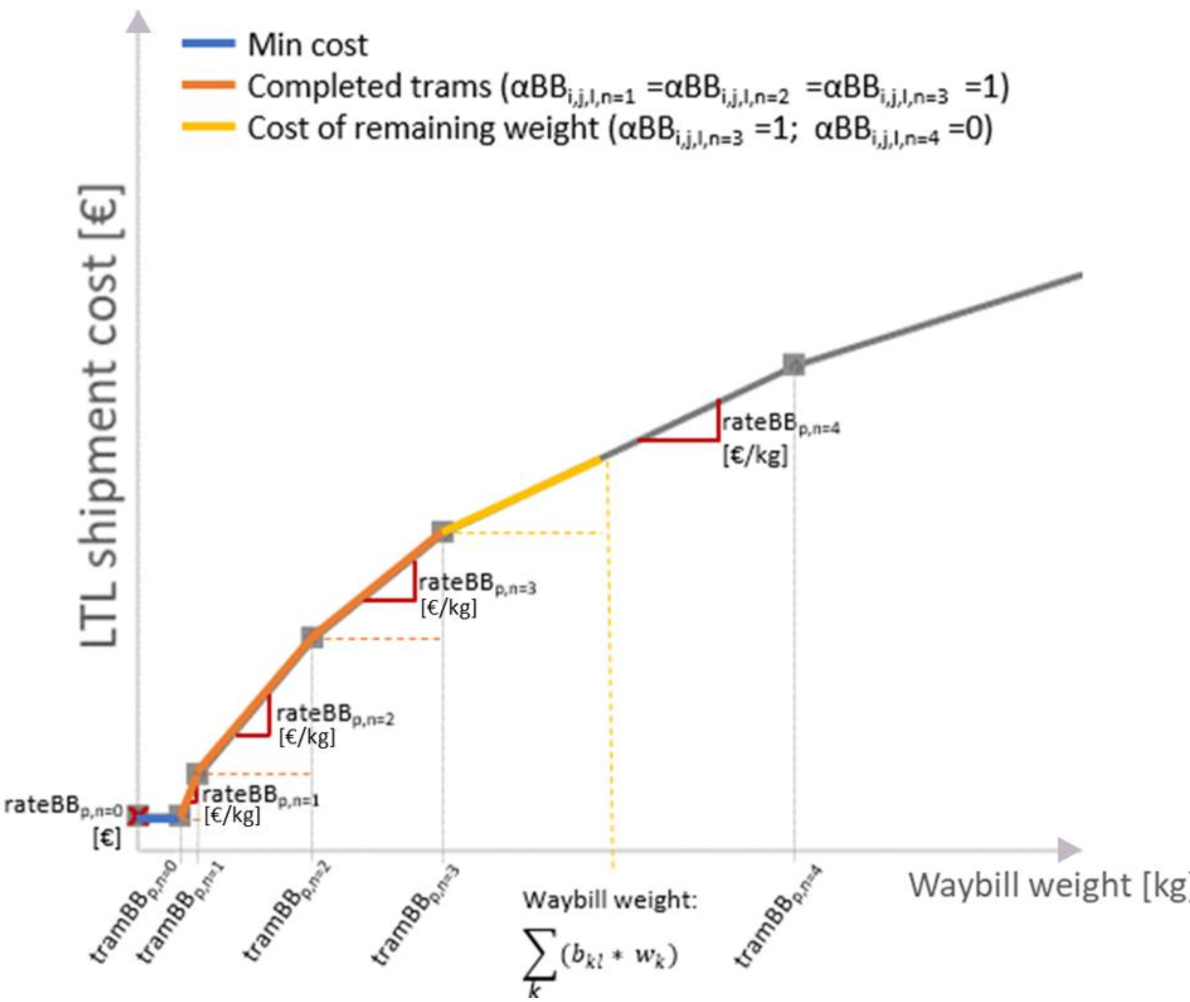

Figure 8. Graphic visualization of the LTL cost function.

4.2.2. Linearization of binary variables product

The objective is to derive a mixed integer linear formulation. Therefore, products of variables should be avoided or, if necessary, linearized. In this section we show how we do it.

Being a and b two of the model binary variables, in all those cases where they appear multiplying each other we have defined a third binary variable $L_{ab} = a * b$. To ensure that the value of this new variable is consistent with the values of a and b, we need some extra constrains as, for instance, the following set, known as Fortet inequalities:

$$L_{ab} \leq a$$

$$L_{ab} \leq b \; L_{ab} \geq a + b - 1$$

a,b,c ∈ [0,1]

4.3. MODEL FORMULATION

4.3.1. Input: shipment data and parameters

- Indices:

  k = {1,2,…, N_k} is the PackID index

  i = {1,2,…, N_i} is the origin factory index

  j = {1,2,…, N_j} is the destination customer index

  p = {1,2,…, N_p} is the destination zone index

  s = {1,2,…, N_s} is the day index

  l = {1,2,…, N_l} is the truck index. The truck number (l) is tied to each FTL route. I.e. there will be so many trucks with l=1 as number of different FTL routes.

  n = {0,1,2,…, N_n} is the interval index (of the linearized LTL cost function).

- Shipments data:

  w_k = weight of the PID k [kg]

  $fact_k$ = i-origin factory of the PID k

  $customer_k$ = j-destination customer of the PID k

  r_k = r requested factory ship date of the PID k

  $zone_j$ = p-destination zone of the j customer

- FTL Trucking costs:

  D_{ip} = FTL direct cost from the origin factory i to a customer in zone p [€/truck].

  B_i = FTL truck cost from the origin factory i to the Barcelona hub [€/truck].

  M_i = FTL truck cost from the origin factory i to the Madrid hub [€/truck].

  Z_i = FTL truck cost from the origin factory i to the Zaragoza hub [€/truck].

  O_i = FTL truck cost from the origin factory i to the Origin consolidation hub [€/truck].

  OD_p = FTL direct cost from the OCP to a customer in zone p [€/truck].

  OB = FTL truck cost from the OCP to the Barcelona hub [€/truck].

  OM = FTL truck cost from the OCP to the Madrid hub [€/truck].

  OZ = FTL truck cost from the OCP to the Zaragoza hub [€/truck].

- <u>LTL Trucking costs:</u>

  $tramBB_{pn}$ = upbound waybill weight of the interval *n* of the LTL linearized cost function from the BCN hub to a customer in zone *p*.

  $tramMM_{pn}$ = upbound waybill weight of the interval *n* of the LTL linearized cost function from the MAD hub to a customer in zone *p*.

  $tramZZ_{pn}$ = upbound waybill weight of the interval *n* of the LTL linearized cost function from the ZAR hub to a customer in zone *p*.

  $tramFF_{ipn}$ = upbound waybill weight of the interval *n* of the LTL linearized cost function from the factory *i* to a customer in zone *p*.

  $rateBB_{pn}$ = cost of the interval *n* of the LTL linearized function, from the BCN hub to a customer in zone *p*. *n=0* contains the minimum cost [€], while the rest (*n=1, 2, …, Nn*) contain the slope [€/kg].

  $rateMM_{pn}$ = cost of the interval *n* of the LTL linearized function, from the MAD hub to a customer in zone *p*. *n=0* contains the minimum cost [€], while the rest (*n=1, 2, …, Nn*) contain the slope [€/kg].

  $rateZZ_{pn}$ = cost of the interval *n* of the LTL linearized function, from the ZAR hub to a customer in zone *p*. *n=0* contains the minimum cost [€], while the rest (*n=1, 2, …, Nn*) contain the slope [€/kg].

  $rateFF_{ipn}$ = cost of the interval *n* of the LTL linearized function, from the factory *i* to a customer in zone *p*. *n=0* contains the minimum cost [€], while the rest (*n=1, 2, …, Nn*) contain the slope [€/kg].

- <u>Other parameters:</u>

  C = truck capacity [kg].

  OO = handling cost at the origin consolidation hub (OCP) [€/kg].

  WF = maximum waiting time at any origin factory; between PID requested and actual shipment date [days].

  WO = maximum waiting time at the OCP; between arrival and departure from OCP [days].

  $TimeO_i$ = travel time from factory *i* to the OCP [days].

4.3.2. Decision variables

- Routing variables: determine the routing of each PID, i.e. the problem's solution.

  ∀*k,l* :

  d_{kl} =1 if packID *k* goes from the factory direct to the customer on the *l* FTL truck.

  =0 otherwise.

  f_{kl} =1 if packID *k* goes from the factory direct to the customer on the *l* LTL truck.

  =0 otherwise.

  b_{kl} =1 if packID *k* goes from the factory to the BCN hub on the *l* FTL truck.

  =0 otherwise.

  m_{kl} =1 if packID *k* goes from the factory to the MAD hub on the *l* FTL truck.

  =0 otherwise.

  z_{kl} =1 if packID *k* goes from the factory to the ZAR hub on the *l* FTL truck.

  =0 otherwise.

  od_{kl} =1 if packID *k* goes from the OCP direct to the customer on the *l* FTL truck.

  =0 otherwise.

  ob_{kl} =1 if packID *k* goes from the OCP to the BCN hub on the *l* FTL truck.

  =0 otherwise.

  om_{kl} =1 if packID *k* goes from the OCP to the MAD hub on the *l* FTL truck.

  =0 otherwise.

  oz_{kl} =1 if packID *k* goes from the OCP to the ZAR hub on the *l* FTL truck.

  =0 otherwise.

- Auxiliary routing variable:

  ∀*k,l* :

  o_{kl} =1 if packID *k* goes via OCP (either OCP +direct/+BCN/+MAD/+ZAR), on the *l* FTL truck from factory to the OCP.

  =0 otherwise.

- Prima variables: determine whether FTL trucks are used.

  ∀*i,j,l* :

  d'_{ijl} =1 if we use the *l* direct truck from factory *i* to customer *j*.

  =0 otherwise.

  b'_{il} =1 if we use the *l* truck from factory *i* to the Barcelona hub.

  =0 otherwise.

  m'_{il} =1 if we use the *l* truck from factory *i* to the Madrid hub.

  =0 otherwise.

  z'_{il} =1 if we use the *l* truck from factory *i* to the Zaragoza hub.

  =0 otherwise.

  o'_{il} =1 if we use the *l* truck from factory *i* to the OCP.

  =0 otherwise.

  od'_{jl} =1 if we use the *l* direct truck from the OCP to the customer *j*.

  =0 otherwise.

  ob'_{l} =1 if we use the *l* truck from the OCP to the Barcelona hub.

  =0 otherwise.

  om'_{l} =1 if we use the *l* truck from the OCP to the Madrid hub.

  =0 otherwise.

  oz'_{l} =1 if we use the *l* truck from the OCP to the Zaragoza hub.

  =0 otherwise.

- Departure variables: determine the truck's departure day from its origin.

  ∀*i,j,l,s* :

  $dDept_{ijls}$ =1 if the *l* direct truck (FTL) from factory *i* to customer *j* departs the factory on day *s*.

  =0 otherwise.

  $fDept_{ijls}$ =1 if the *l* direct truck (LTL) from factory *i* to customer *j* departs the factory on day *s*.

  =0 otherwise.

  $bDept_{ils}$ =1 if the *l* truck from the factory *i* to the BCN hub departs the factory on day *s*.

  =0 otherwise.

$mDept_{ils}$=1 if the *l* truck from the factory *i* to the MAD hub departs the factory on day *s*.

=0 otherwise.

$zDept_{ils}$ =1 if the *l* truck from the factory *i* to the ZAR hub departs the factory on day *s*.

=0 otherwise.

$oDept_{ils}$ =1 if the *l* truck from the factory *i* to the OCP departs the factory on day *s*.

=0 otherwise.

$odDept_{jls}$=1 if the *l* direct truck from the OCP to customer *j* departs the OCP on day *s*.

=0 otherwise.

$obDept_{ls}$=1 if the *l* truck from the OCP to the BCN hub departs the OCP on day *s*.

=0 otherwise.

$omDept_{ls}$=1 if the *l* truck from the OCP to the MAD hub departs the OCP on day *s*.

=0 otherwise.

$ozDept_{ls}$=1 if the *l* truck from the OCP to the ZAR hub departs the OCP on day *s*.

=0 otherwise.

- <u>WV variables</u>: determine whether a minimum LTL cost applies to a waybill.

  $\forall i,j,l$:

  WVf_{ijl} =1 if there is a waybill from factory *i* to customer *j* travelling on the *l* LTL truck from factory direct to customer.

  =0 otherwise.

  WVb_{ijl} =1 if there is a waybill from factory *i* to customer *j* travelling on the *l* truck from the factory to the BCN hub (meaning it will go on LTL from BCN hub to customer).

  =0 otherwise.

  WVm_{ijl} =1 if there is a waybill from factory *i* to customer *j* travelling on the *l* truck from the factory to the MAD hub (meaning it will go on LTL from MAD hub to customer).

  =0 otherwise.

  WVz_{ijl} =1 if there is a waybill from factory *i* to customer *j* travelling on the *l* truck from the factory to the ZAR hub (meaning it will go on LTL from ZAR hub to customer).

  =0 otherwise.

$WVob_{ijl}$ =1 if there is a waybill from factory *i* to customer *j* travelling on the *l* truck from the OCP to the BCN hub (meaning it will go on LTL from BCN hub to customer).

=0 otherwise.

$WVom_{ijl}$ =1 if there is a waybill from factory *i* to customer *j* travelling on the *l* truck from the OCP to the MAD hub (meaning it will go on LTL from MAD hub to customer).

=0 otherwise.

$WVoz_{ijl}$ =1 if there is a waybill from factory *i* to customer *j* travelling on the *l* truck from the OCP to the ZAR hub (meaning it will go on LTL from ZAR hub to customer).

=0 otherwise.

- Alpha variables: determine the interval of the LTL cost function based on the waybill weight.

  $\forall i,j,l,n$:

  αFF_{ijln} =1 if the weight of the waybill (defined by factory *i* to customer *j* travelling on the *l* LTL truck from factory direct to customer) exceeds interval *n*.

  =0 otherwise.

  αBB_{ijln} =1 if the weight of the waybill (defined by factory *i* to customer *j* travelling on the *l* truck from the factory to the BCN hub) exceeds interval *n*.

  =0 otherwise.

  αMM_{ijln} =1 if the weight of the waybill (defined by factory *i* to customer *j* travelling on the *l* truck from the factory to the MAD hub) exceeds interval *n*.

  =0 otherwise.

  αZZ_{ijln} =1 if the weight of the waybill (defined by factory *i* to customer *j* travelling on the *l* truck from the factory to the ZAR hub) exceeds interval *n*.

  =0 otherwise.

  αOBB_{ijln} =1 if the weight of the waybill (defined by factory *i* to customer *j* travelling on the *l* truck from OCP to the BCN hub) exceeds interval *n*.

  =0 otherwise.

  αOMM_{ijln}=1 if the weight of the waybill (defined by factory *i* to customer *j* travelling on the *l* truck from OCP to the MAD hub) exceeds interval *n*.

  =0 otherwise.

  αOZZ_{ijln} =1 if the weight of the waybill (defined by factory *i* to customer *j* travelling on the *l* truck from OCP to the ZAR hub) exceeds interval *n*.

=0 otherwise.

- Linearized variables to avoid products of variables: *L_a_b=a*b={0,1}.* As all model variables are binary {0,1}, all those linearized variables *L_a_b* have value 1 if both variables *a=1* and *b=1*. More info in section "*4.2.2. Linearization of binary variables products*".

  ∀*k,i,j,l,s,n>0* :

  $L\_WVf\_f_{kl} = WVf_{ijl} * f_{kl}$

  $L\_WVb\_b_{kl} = WVb_{ijl} * b_{kl}$

  $L\_WVm\_m_{kl} = WVm_{ijl} * m_{kl}$

  $L\_WVz\_z_{kl} = WVz_{ijl} * z_{kl}$

  $L\_WVob\_ob_{kl} = WVob_{ijl} * ob_{kl}$

  $L\_WVom\_om_{kl} = WVom_{ijl} * om_{kl}$

  $L\_WVoz\_oz_{kl} = WVoz_{ijl} * oz_{kl}$

  $L\_oDept\_o_{kls} = oDept_{ils} * o_{kl}$

  $L\_odDept\_od_{kls} = odDept_{jls} * od_{kl}$

  $L\_obDept\_ob_{kls} = obDept_{ls} * ob_{kl}$

  $L\_omDept\_om_{kls} = omDept_{ls} * om_{kl}$

  $L\_ozDept\_oz_{kls} = ozDept_{ls} * oz_{kl}$

  $L\_alphas\_f_{kln} = [\alpha FF_{i,j,l,n-1} - \alpha FF_{i,j,l,n}] * f_{k,l}$

  $L\_alphas\_b_{kln} = [\alpha BB_{i,j,l,n-1} - \alpha BB_{i,j,l,n}] * b_{k,l}$

  $L\_alphas\_m_{kln} = [\alpha MM_{i,j,l,n-1} - \alpha MM_{i,j,l,n}] * m_{k,l}$

  $L\_alphas\_z_{kln} = [\alpha ZZ_{i,j,l,n-1} - \alpha ZZ_{i,j,l,n}] * z_{k,l}$

  $L\_alphas\_ob_{kln} = [\alpha OBB_{i,j,l,n-1} - \alpha OBB_{i,j,l,n}] * ob_{k,l}$

  $L\_alphas\_om_{kln} = [\alpha OMM_{i,j,l,n-1} - \alpha OMM_{i,j,l,n}] * om_{k,l}$

  $L\_alphas\_oz_{kln} = [\alpha OZZ_{i,j,l,n-1} - \alpha OZZ_{i,j,l,n}] * oz_{k,l}$

4.3.3. Objective function

$$Min(TotalCost)$$

$$TotalCost = FTLcost + FTXcost + toOCPcost + OFTLcost + OFTXcost + Ohandlingcost + mincost + ltldir + ltlbcn + ltlmad + ltlzar + ltlobcn + ltlomad + ltlozar$$

with:

$$FTLcost = \sum_{\substack{ijpl \\ zone_j=p}} D_{ip} * d'_{ijl}$$

$$FTXcost = \sum_{i} [B_i * \sum_{l} b'_{il} + M_i * \sum_{l} m'_{il} + Z_i * \sum_{l} z'_{il}]$$

$$toOCPcost = \sum_{i} [O_i * \sum_{l} o'_{il}]$$

$$OFTLcost = \sum_{\substack{jpl \\ zone_j=p}} [OD_p * od'_{jl}]$$

$$OFTXcost = OB * \sum_{l} ob'_l + OM * \sum_{l} om'_l + OZ * \sum_{l} oz'_l$$

$$Ohandlingcost = OO * \sum_{kl} w_k * o_{kl}$$

$$mincost = \sum_{\substack{ijpl \\ zone_j=p}} [rateFF_{i,p,n=0} * WVf_{ijl} + rateBB_{p,n=0} * (WVb_{ijl} + WVob_{ijl})$$
$$+ rateMM_{p,n=0} * (WVm_{ijl} + WVom_{ijl}) + rateZZ_{p,n=0} * (WVz_{ijl} + WVoz_{ijl})]$$

$$ltldir = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha FF_{ijln} * rateFF_{ipn} * \left(tramFF_{ipn} - tramFF_{i,p,n-1}\right)$$
$$+ rateFF_{ipn} * \left(\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (L\_alphas\_f_{kln} * w_k) - (\alpha FF_{i,j,l,n-1} - \alpha FF_{ijln}) * tramFF_{i,p,n-1}\right)]$$

$$ltlbcn = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha BB_{ijln} * rateBB_{pn} * \left(tramBB_{p,n} - tramBB_{p,n-1}\right)$$
$$+ rateBB_{pn} * \left(\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (L\_alphas\_b_{kln} * w_k) - (\alpha BB_{i,j,l,n-1} - \alpha BB_{ijln}) * tramBB_{p,n-1}\right)]$$

$$ltlmad = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha MM_{ijln} * rateMM_{pn} * \left(tramMM_{p,n} - tramMM_{p,n-1}\right)$$
$$+ rateMM_{pn} * \left(\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (L\_alphas\_m_{kln} * w_k) - (\alpha MM_{i,j,l,n-1} - \alpha MM_{ijln}) * tramMM_{p,n-1}\right)]$$

$$ltlzar = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha ZZ_{ijln} * rateZZ_{pn} * \left(tramZZ_{p,n} - tramZZ_{p,n-1}\right)$$
$$+ rateZZ_{pn} * \left(\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (L\_alphas\_z_{kln} * w_k) - (\alpha ZZ_{i,j,l,n-1} - \alpha ZZ_{ijln}) * tramZZ_{p,n-1}\right)]$$

$$ltlobcn = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha OBB_{ijln} * rateBB_{pn} * \left(tramBB_{p,n} - tramBB_{p,n-1}\right)$$
$$+ rateBB_{pn} * \left(\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (L\_alphas\_ob_{kln} * w_k) - (\alpha OBB_{i,j,l,n-1} - \alpha OBB_{ijln}) * tramBB_{p,n-1}\right)]$$

$$ltlomad = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha OMM_{ijln} * rateMM_{pn} * \left(tramMM_{p,n} - tramMM_{p,n-1}\right)$$

$$+ rateMM_{pn} * (\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (L\_alphas\_om_{kln} * w_k) - (\alpha OMM_{i,j,l,n-1} - \alpha OMM_{ijln}) * tramMM_{p,n-1})]$$

$$ltlozar = \sum_{\substack{ijpln \\ zone_j=p \\ n>0}} [\alpha OZZ_{ijln} * rateZZ_{pn} * \left(tramZZ_{p,n} - tramZZ_{p,n-1}\right)$$

$$+ rateZZ_{pn} * (\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (L\_alphas\_oz_{kln} * w_k) - (\alpha OZZ_{i,j,l,n-1} - \alpha OZZ_{ijln}) * tramZZ_{p,n-1})]$$

4.3.4. Constraints

R1. Each PackID must use one (and only one) route/truck.

$$\sum_{l} d_{kl} + f_{kl} + b_{kl} + m_{kl} + z_{kl} + o_{kl} = 1 \qquad \forall k$$

R2. Constraints to limit trucks' capacity.

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * d_{kl} \leq C \quad \forall i,j,l$$

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * f_{kl} \leq C \quad \forall i,j,l$$

$$\sum_{\substack{k \\ fact_k=i}} w_k * b_{kl} \leq C \quad \forall i,l$$

$$\sum_{\substack{k \\ fact_k=i}} w_k * m_{kl} \leq C \quad \forall i,l$$

$$\sum_{\substack{k \\ fact_k=i}} w_k * z_{kl} \leq C \quad \forall i,l$$

$$\sum_{\substack{k \\ fact_k=i}} w_k * o_{kl} \leq C \quad \forall i,l$$

$$\sum_{\substack{k \\ customer_k=j}} w_k * od_{kl} \leq C \quad \forall j,l$$

$$\sum_{k} w_k * ob_{kl} \leq C \quad \forall l$$

$$\sum_{k} w_k * om_{kl} \leq C \quad \forall l$$

$$\sum_{k} w_k * oz_{kl} \leq C \quad \forall l$$

R3. Define prima variables: force their activation if we use the *l* truck of the FTL route.

$$d_{kl} \leq d'_{ijl} \quad \forall i,j,k,l \ \ st. \ fact_k = i, customer_k = j$$

$$b_{kl} \leq b'_{il} \quad \forall i,k,l \ \ st. \ fact_k = i$$

$$m_{kl} \leq m'_{il} \quad \forall i,k,l \ \ st. \ fact_k = i$$

$$z_{kl} \leq z'_{il} \quad \forall i,k,l \ \ st. \ fact_k = i$$

$$o_{kl} \leq o'_{il} \quad \forall i,k,l \ \ st. \ fact_k = i$$

$$od_{kl} \leq od'_{jl} \quad \forall j,k,l \ \ st. \ customer_k = j$$

$$ob_{kl} \leq ob'_{l} \quad \forall k,l$$

$$om_{kl} \leq om'_{l} \quad \forall k,l$$

$$oz_{kl} \leq oz'_{l} \quad \forall k,l$$

R4. Define prima variables: allow its activation only if the *l* truck of the FTL route is used (not needed because we are minimizing, however it helps the algorithm).

$$\sum_{\substack{k \\ fact_k = i \\ customer_k = j}} d_{kl} \geq d'_{ijl} \quad \forall i,j,l$$

$$\sum_{\substack{k \\ fact_k = i}} b_{kl} \geq b'_{il} \quad \forall i,l$$

$$\sum_{\substack{k \\ fact_k = i}} m_{kl} \geq m'_{il} \quad \forall i,l$$

$$\sum_{\substack{k \\ fact_k = i}} z_{kl} \geq z'_{il} \quad \forall i,l$$

$$\sum_{\substack{k \\ fact_k = i}} o_{kl} \geq o'_{il} \quad \forall i,l$$

$$\sum_{\substack{k \\ customer_k=j}} od_{kl} \geq od'_{jl} \quad \forall j,l$$

$$\sum_{k} ob_{kl} \geq ob'_{l} \quad \forall l$$

$$\sum_{k} om_{kl} \geq om'_{l} \quad \forall l$$

$$\sum_{k} oz_{kl} \geq oz'_{l} \quad \forall l$$

R5. Define prima variables: avoid symmetries, otherwise different index orders would be read by the algorithm as different solutions, drastically increasing the computational time.

$$d'_{ijl} \leq d'_{i,j,l-1} \quad \forall i,j,l \;\; st. \; l > 1$$

$$b'_{il} \leq b'_{i,l-1} \quad \forall i,l \;\; st. \; l > 1$$

$$m'_{il} \leq m'_{i,l-1} \quad \forall i,l \;\; st. \; l > 1$$

$$z'_{il} \leq z'_{i,l-1} \quad \forall i,l \;\; st. \; l > 1$$

$$o'_{il} \leq o'_{i,l-1} \quad \forall i,l \;\; st. \; l > 1$$

$$od'_{jl} \leq od'_{jl-1} \quad \forall j,l \;\; st. \; l > 1$$

$$ob'_{l} \leq ob'_{l-1} \quad \forall l \;\; st. \; l > 1$$

$$om'_{l} \leq om'_{l-1} \quad \forall l \;\; st. \; l > 1$$

$$oz'_{l} \leq oz'_{l-1} \quad \forall l \;\; st. \; l > 1$$

R6. Constraints on truck shipping date: if truck *l* has PIDs assigned, it has to depart on a day that aligns with the requirements of all its assigned PIDs.

$$d_{kl} \leq \sum_{\substack{s \\ r_k \leq s \\ s \leq r_k+WF}} dDept_{ijls} \quad \forall i,j,k,l \;\; st. \; fact_k = i, customer_k = j$$

$$f_{kl} \leq \sum_{\substack{s \\ r_k \leq s \\ s \leq r_k+WF}} fDept_{ijls} \quad \forall i,j,k,l \;\; st. \; fact_k = i, customer_k = j$$

$$b_{kl} \leq \sum_{\substack{s \\ r_k \leq s \\ s \leq r_k + WF}} bDept_{ils} \quad \forall i,k,l \;\; st.\; fact_k = i$$

$$m_{kl} \leq \sum_{\substack{s \\ r_k \leq s \\ s \leq r_k + WF}} mDept_{ils} \quad \forall i,k,l \;\; st.\; fact_k = i$$

$$z_{kl} \leq \sum_{\substack{s \\ r_k \leq s \\ s \leq r_k + WF}} zDept_{ils} \quad \forall i,k,l \;\; st.\; fact_k = i$$

$$o_{kl} \leq \sum_{\substack{s \\ r_k \leq s \\ s \leq r_k + WF}} oDept_{ils} \quad \forall i,k,l \;\; st.\; fact_k = i$$

R7. Constraints on truck shipping date: force that Departure variables are only activated when the truck has at least one PID assigned. Not needed in a minimization problem but helps the algorithm.

$$dDept_{ijls} \leq \sum_{\substack{k \\ fact_k = i \\ customer_k = j}} d_{kl} \quad \forall i,j,l,s$$

$$fDept_{ijls} \leq \sum_{\substack{k \\ fact_k = i \\ customer_k = j}} f_{kl} \quad \forall i,j,l,s$$

$$bDept_{ils} \leq \sum_{\substack{k \\ fact_k = i}} b_{kl} \quad \forall i,l,s$$

$$mDept_{ils} \leq \sum_{\substack{k \\ fact_k = i}} m_{kl} \quad \forall i,l,s$$

$$zDept_{ils} \leq \sum_{\substack{k \\ fact_k = i}} z_{kl} \quad \forall i,l,s$$

$$oDept_{ils} \leq \sum_{\substack{k \\ fact_k = i}} o_{kl} \quad \forall i,l,s$$

$$odDept_{jls} \leq \sum_{\substack{k \\ customer_k=j}} od_{kl} \qquad \forall j,l,s$$

$$obDept_{ls} \leq \sum_{k} ob_{kl} \qquad \forall l,s$$

$$omDept_{ls} \leq \sum_{k} om_{kl} \qquad \forall l,s$$

$$ozDept_{ls} \leq \sum_{k} oz_{kl} \qquad \forall l,s$$

R8. Constraints on truck shipping date: a given truck can only go one day.

$$\sum_{s} dDept_{ijls} \leq 1 \qquad \forall i,j,l$$

$$\sum_{s} fDept_{ijls} \leq 1 \qquad \forall i,j,l$$

$$\sum_{s} bDept_{ils} \leq 1 \qquad \forall i,l$$

$$\sum_{s} mDept_{ils} \leq 1 \qquad \forall i,l$$

$$\sum_{s} zDept_{ils} \leq 1 \qquad \forall i,l$$

$$\sum_{s} oDept_{ils} \leq 1 \qquad \forall i,l$$

$$\sum_{s} odDept_{jls} \leq 1 \qquad \forall j,l$$

$$\sum_{s} obDept_{ls} \leq 1 \qquad \forall l$$

$$\sum_{s} omDept_{ls} \leq 1 \qquad \forall l$$

$$\sum_{s} ozDept_{ls} \leq 1 \qquad \forall l$$

R9. Constraints to define the auxiliary OCP variables.

$$\sum_{l} o_{kl} = \sum_{l} od_{kl} + ob_{kl} + om_{kl} + oz_{kl} \qquad \forall k$$

$$\sum_{l} L\_oDept\_o_{kls} \leq \sum_{\substack{l,s' \\ s' \geq s+TimeO_i \\ s' \leq s+TimeO_i+WO}} L\_odDept\_od_{kls'} + L\_obDept\_ob_{kls'} + L\_omDept\_om_{kls'} + L\_ozDept\_oz_{kls'} \quad \forall i,j,k,s \;\; st. fact_k = i, customer_k = j$$

R10. Constraints to define the alpha variables: if the waybill weight surpasses interval *n*, the alpha variable gets activated. A waybill is defined by all PIDS that travel end to end together (from same factory *i* to same customer *j* in same truck *l*).

$$tramFF_{ipn} + (C - tramFF_{ipn}) * \alpha FF_{ijln} \geq \sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * f_{kl} \qquad \forall i,j,p,l,n \;\; st. \;\; zone_j = p$$

$$tramBB_{pn} + (C - tramBB_{pn}) * \alpha BB_{ijln} \geq \sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * b_{kl} \qquad \forall i,j,p,l,n \;\; st. \;\; zone_j = p$$

$$tramMM_{pn} + (C - tramMM_{pn}) * \alpha MM_{ijln} \geq \sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * m_{kl} \qquad \forall i,j,p,l,n \;\; st. \;\; zone_j = p$$

$$tramZZ_{pn} + (C - tramZZ_{pn}) * \alpha ZZ_{ijln} \geq \sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * z_{kl} \qquad \forall i,j,p,l,n \;\; st. \;\; zone_j = p$$

$$tramBB_{pn} + (C - tramBB_{pn}) * \alpha OBB_{ijln} \geq \sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * ob_{kl} \qquad \forall i,j,p,l,n \;\; st. \;\; zone_j = p$$

$$tramMM_{pn} + (C - tramMM_{pn}) * \alpha OMM_{ijln} \geq \sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * om_{kl} \qquad \forall i,j,p,l,n \;\; st. \;\; zone_j = p$$

$$tramZZ_{pn} + (C - tramZZ_{pn}) * \alpha OZZ_{ijln} \geq \sum_{\substack{k \\ fact_k=i \\ customer_k=j}} w_k * oz_{kl} \qquad \forall i,j,p,l,n \;\; st. \;\; zone_j = p$$

R11. Constraints to define the alpha variables: force that they are only activated if the waybill weight surpasses interval *n*.

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (\mathrm{w}_k * f_{kl}) \geq (\mathrm{tramFF}_{ipn} + \varepsilon) * \alpha\mathrm{FF}_{ijln} \qquad \forall\, i,j,p,l,n \;\; zone_j = p$$

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (\mathrm{w}_k * b_{kl}) \geq (\mathrm{tramBB}_{pn} + \varepsilon) * \alpha\mathrm{BB}_{ijln} \qquad \forall\, i,j,p,l,n \;\; zone_j = p$$

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (\mathrm{w}_k * m_{kl}) \geq (\mathrm{tramMM}_{pn} + \varepsilon) * \alpha\mathrm{MM}_{ijln} \qquad \forall\, i,j,p,l,n \;\; zone_j = p$$

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (\mathrm{w}_k * z_{kl}) \geq (\mathrm{tramZZ}_{pn} + \varepsilon) * \alpha\mathrm{ZZ}_{ijln} \qquad \forall\, i,j,p,l,n \;\; zone_j = p$$

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (\mathrm{w}_k * ob_{kl}) \geq (\mathrm{tramBB}_{pn} + \varepsilon) * \alpha\mathrm{OBB}_{ijln} \qquad \forall\, i,j,p,l,n \;\; zone_j = p$$

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (\mathrm{w}_k * om_{kl}) \geq (\mathrm{tramMM}_{pn} + \varepsilon) * \alpha\mathrm{OMM}_{ijln} \qquad \forall\, i,j,p,l,n \;\; zone_j = p$$

$$\sum_{\substack{k \\ fact_k=i \\ customer_k=j}} (\mathrm{w}_k * oz_{kl}) \geq (\mathrm{tramZZ}_{pn} + \varepsilon) * \alpha\mathrm{OZZ}_{ijln} \qquad \forall\, i,j,p,l,n \;\; zone_j = p$$

R12. Constraints to define alphas: avoid symmetries.

$$\alpha FF_{ijln} \leq \alpha FF_{i,j,l,n-1} \quad \forall i,j,l,n > 0$$
$$\alpha BB_{ijln} \leq \alpha BB_{i,j,l,n-1} \quad \forall i,j,l,n > 0$$
$$\alpha MM_{ijln} \leq \alpha MM_{i,j,l,n-1} \quad \forall i,j,l,n > 0$$
$$\alpha ZZ_{ijln} \leq \alpha ZZ_{i,j,l,n-1} \quad \forall i,j,l,n > 0$$
$$\alpha OBB_{ijln} \leq \alpha OBB_{i,j,l,n-1} \quad \forall i,j,l,n > 0$$
$$\alpha OMM_{ijln} \leq \alpha OMM_{i,j,l,n-1} \quad \forall i,j,l,n > 0$$
$$\alpha OZZ_{ijln} \leq \alpha OZZ_{i,j,l,n-1} \quad \forall i,j,l,n > 0$$

R13. Constraints to define the WV variables (its linearization): WV variables get activated if there is a waybill with weight>0 from factory *i* to customer *j* in the truck *l*.

$$\sum_k (w_k * L\_WVf\_f_{kl}) \geq \sum_k (w_k * f_{kl}) \quad \forall l$$

$$\sum_k (w_k * L\_WVb\_b_{kl}) \geq \sum_k (w_k * b_{kl}) \quad \forall l$$

$$\sum_k (w_k * L\_WVm\_m_{kl}) \geq \sum_k (w_k * m_{kl}) \quad \forall l$$

$$\sum_k (w_k * L\_WVz\_z_{kl}) \geq \sum_k (w_k * z_{kl}) \quad \forall l$$

$$\sum_k (w_k * L\_WVob\_ob_{kl}) \geq \sum_k (w_k * ob_{kl}) \quad \forall l$$

$$\sum_k (w_k * L\_WVom\_om_{kl}) \geq \sum_k (w_k * om_{kl}) \quad \forall l$$

$$\sum_k (w_k * L\_WVoz\_oz_{kl}) \geq \sum_k (w_k * oz_{kl}) \quad \forall l$$

R14. Constraints to define linearization variables:

$$L\_oDept\_o_{kls} \leq oDept_{ils} \quad \forall k,i,l,s, \;\; st. \; fact_k = i$$
$$L\_oDept\_o_{kls} \leq o_{kl} \quad \forall k,l,s$$
$$L\_oDept\_o_{kls} \geq oDept_{ils} + o_{kl} - 1 \quad \forall k,i,l,s, \;\; st. \; fact_k = i$$

$$L\_odDept\_od_{kls} \leq odDept_{jls} \quad \forall k,j,l,s, \;\; st. \; customer_k = j$$
$$L\_odDept\_od_{kls} \leq od_{kl} \quad \forall k,l,s$$
$$L\_odDept\_od_{kls} \geq odDept_{jls} + od_{kl} - 1 \quad \forall k,j,l,s, \;\; st. \; customer_k = j$$

$$L\_obDept\_ob_{kls} \leq obDept_{ls} \quad \forall k,l,s$$
$$L\_obDept\_ob_{kls} \leq ob_{kl} \quad \forall k,l,s$$
$$L\_obDept\_ob_{kls} \geq obDept_{ls} + ob_{kl} - 1 \quad \forall k,l,s$$

$$L\_omDept\_om_{kls} \leq omDept_{ls} \quad \forall k,l,s$$
$$L\_omDept\_om_{kls} \leq om_{kl} \quad \forall k,l,s$$
$$L\_omDept\_om_{kls} \geq omDept_{ls} + om_{kl} - 1 \quad \forall k,l,s$$

$$L\_ozDept\_oz_{kls} \leq ozDept_{ls} \quad \forall k,l,s$$
$$L\_ozDept\_oz_{kls} \leq oz_{kl} \quad \forall k,l,s$$
$$L\_ozDept\_oz_{kls} \geq ozDept_{ls} + oz_{kl} - 1 \quad \forall k,l,s$$

$$L\_WVf\_f_{kl} \leq WVf_{ijl} \quad \forall i,j,k,l \;\; st. \;\; fact_k = i, customer_k = j$$
$$L\_WVf\_f_{kl} \leq f_{kl} \quad \forall k,l$$

$$L\_WVf\_f_{kl} \geq WVf_{ijl} + f_{kl} - 1 \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$

$$L\_WVb\_b_{kl} \leq WVb_{ijl} \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$
$$L\_WVb\_b_{kl} \leq b_{kl} \quad \forall k,l$$
$$L\_WVb\_b_{kl} \geq WVb_{ijl} + b_{kl} - 1 \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$

$$L\_WVm\_m_{kl} \leq WVm_{ijl} \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$
$$L\_WVm\_m_{kl} \leq m_{kl} \quad \forall k,l$$
$$L\_WVm\_m_{kl} \geq WVm_{ijl} + m_{kl} - 1 \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$

$$L\_WVz\_z_{kl} \leq WVz_{ijl} \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$
$$L\_WVz\_z_{kl} \leq z_{kl} \quad \forall k,l$$
$$L\_WVz\_z_{kl} \geq WVz_{ijl} + z_{kl} - 1 \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$

$$L\_WVob\_ob_{kl} \leq WVob_{ijl} \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$
$$L\_WVob\_ob_{kl} \leq ob_{kl} \quad \forall k,l$$
$$L\_WVob\_ob_{kl} \geq WVob_{ijl} + ob_{kl} - 1 \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$

$$L\_WVom\_om_{kl} \leq WVom_{ijl} \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$
$$L\_WVom\_om_{kl} \leq om_{kl} \quad \forall k,l$$
$$L\_WVom\_om_{kl} \geq WVom_{ijl} + om_{kl} - 1 \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$

$$L\_WVoz\_oz_{kl} \leq WVoz_{ijl} \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$
$$L\_WVoz\_oz_{kl} \leq oz_{kl} \quad \forall k,l$$
$$L\_WVoz\_oz_{kl} \geq WVoz_{ijl} + oz_{kl} - 1 \quad \forall i,j,k,l \quad st. \quad fact_k = i, customer_k = j$$

$$L\_alphas\_f_{kln} \leq f_{kl} \quad \forall k,l,n$$
$$L\_alphas\_f_{kln} \leq \left(\alpha FF_{i,j,l,n-1} - \alpha FF_{ijln}\right) \qquad \forall i,j,k,l,n \;\; st. \;\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_f_{kln} \geq f_{kl} + \left(\alpha FF_{i,j,l,n-1} - \alpha FF_{ijln}\right) - 1 \qquad \forall i,j,k,l,n \;\; st. \;\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_b_{kln} \leq b_{kl} \quad \forall k,l,n$$
$$L\_alphas\_b_{kln} \leq \left(\alpha BB_{i,j,l,n-1} - \alpha BB_{ijln}\right) \qquad \forall i,j,k,l,n \;\; st. \;\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_b_{kln} \geq b_{kl} + \left(\alpha BB_{i,j,l,n-1} - \alpha BB_{ijln}\right) - 1 \qquad \forall i,j,k,l,n \;\; st. \;\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_m_{kln} \leq m_{kl} \quad \forall k,l,n$$
$$L\_alphas\_m_{kln} \leq \left(\alpha MM_{i,j,l,n-1} - \alpha MM_{ijln}\right) \qquad \forall i,j,k,l,n \;\; st. \;\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_m_{kln} \geq m_{kl} + \left(\alpha MM_{i,j,l,n-1} - \alpha MM_{ijln}\right) - 1 \qquad \forall i,j,k,l,n \;\; st. \;\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_z_{kln} \leq z_{kl} \quad \forall k,l,n$$
$$L\_alphas\_z_{kln} \leq \left(\alpha ZZ_{i,j,l,n-1} - \alpha ZZ_{ijln}\right) \qquad \forall i,j,k,l,n \;\; st. \;\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_z_{kln} \geq z_{kl} + \left(\alpha ZZ_{i,j,l,n-1} - \alpha ZZ_{ijln}\right) - 1 \qquad \forall i,j,k,l,n \;\; st. \;\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_ob_{kln} \leq ob_{kl} \quad \forall k,l,n$$
$$L\_alphas\_ob_{kln} \leq \left(\alpha OBB_{i,j,l,n-1} - \alpha OBB_{ijln}\right) \qquad \forall i,j,k,l,n \;\; st.\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_ob_{kln} \geq ob_{kl} + \left(\alpha OBB_{i,j,l,n-1} - \alpha OBB_{ijln}\right) - 1 \qquad \forall i,j,k,l,n \;\; st.\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_om_{kln} \leq om_{kl} \quad \forall k,l,n$$
$$L\_alphas\_om_{kln} \leq \left(\alpha OMM_{i,j,l,n-1} - \alpha OMM_{ijln}\right) \qquad \forall i,j,k,l,n \;\; st.\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_om_{kln} \geq om_{kl} + \left(\alpha OMM_{i,j,l,n-1} - \alpha OMM_{ijln}\right) - 1 \qquad \forall i,j,k,l,n \;\; st.\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_oz_{kln} \leq oz_{kl} \quad \forall k,l,n$$
$$L\_alphas\_oz_{kln} \leq \left(\alpha OZZ_{i,j,l,n-1} - \alpha OZZ_{ijln}\right) \qquad \forall i,j,k,l,n \;\; st.\; fact_k = i, customer_k = j, n > 0$$

$$L\_alphas\_oz_{kln} \geq oz_{kl} + \left(\alpha OZZ_{i,j,l,n-1} - \alpha OZZ_{ijln}\right) - 1 \qquad \forall i,j,k,l,n \;\; st.\; fact_k = i, customer_k = j, n > 0$$

4.4. RUNNING THE MODEL

4.4.1. Software and algorithm

In the previous sections we have presented an integer programming formulation of the particular problem studied in detail in this project. The standard method for solving this type of programs is the branch & bound method, which is an implicit enumeration method that discards entire portions of the feasible region based on bounds of the best value that can be attained on these regions. Indeed, most commercial solvers implement branch & cut algorithms, that combine branch & bound with the gradual inclusion of generic valid inequalities. These inequalities help pushing the bounds closer to the actual values of the integer solutions, speeding up the optimization process.

Several solvers are available for addressing this type of problems. The most competitive ones include Xpress, Cplex, or Gurobi. In all cases, the available platforms allow to introduce separately the formal model and the instance data, so that solving the problem with different datasets (different instances of the same problem) only requires modifying the data files, but does not require modifying the formulation.

The solver chosen in this project has been Xpress[m1], because its ease of use and also because student licenses are available. Although some tests have been performed with different policies for exploring the branch and bound tree, finally the algorithm has been run with its default settings.

Appendix A contains the XPRESS code with the model's formulation, together with the input parameters and data for the sample provided in the following section.

4.4.2. Sample

We have run an example with 30 PIDS, from 4 factories to 10 customers, in the period of 2 days:

<u>Input</u>

- Shipment data as shown in Table 1.

Table 1. Input Shipment data for the example with 30PIDS.

PIDS	w	r	fact	customer	zone
1	2618	1	2	7	6
2	2452	1	2	3	7
3	5	1	4	10	6
4	4828	1	2	2	10
5	6388	1	2	10	6
6	2998	1	2	8	2
7	1346	1	2	2	10
8	4215	1	2	5	8
9	978	1	2	10	6
10	3902	1	2	8	2
11	478	1	2	2	10
12	5220	1	2	2	10
13	315	1	2	1	1
14	927	1	2	7	6
15	3248	1	2	10	6
16	5388	1	2	4	1
17	5934	2	2	2	10
18	2220	2	2	8	2
19	4292	2	2	7	6
20	5935	2	2	2	10
21	10	2	2	6	2
22	5821	2	3	8	2
23	5387	2	3	9	6
24	1737	2	3	9	6
25	6421	2	3	7	6
26	262	2	3	6	2
27	3845	2	1	2	10
28	1036	2	1	3	7
29	5606	2	1	5	8
30	3624	2	1	3	7

- Costs: detailed visualization of all rates can be found in *Appendix B*.
- Other parameters:
  - Truck capacity: C=6500 kg
  - Maximum time that shipments can wait at factory: WF=0 days
  - Maximum time that shipments can wait at OCP: WO=1 day
  - Routing time from factories to OCP:

    2 days from factory 1

    1 day from factories 2,3,4.

Output

The solver has found a global optimum, with a total cost of 50577.91€.

The solution is represented in Figure 9 and Table2. It has routed some shipments direct from factory to customer, while others have used the OCP or the BCN hub. The MAD and ZAR hubs have not been used. It is to be noted that the OCP has been used to achieve better consolidation (ex PIDS 10 and 18, that had different factory departure days without allowance to wait there), but also to achieve better rates: due to supply/demand balance, in some cases it is cheaper to route via the OCP than direct (ex. Fact2 to customers in Zone10).

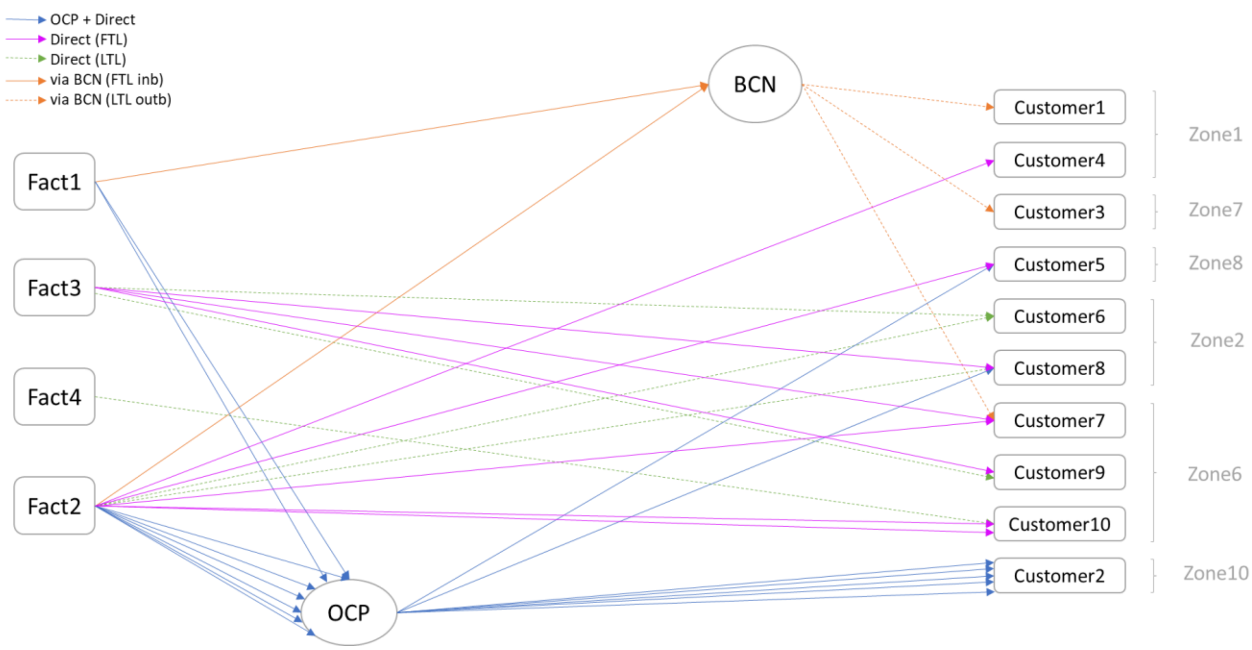

Figure 9. Visual representation of the model solution, example with 30PIDS.

Table 2. Model solution (detailed), example with 30PIDS.

			ID			DAY		WEIGHT [kg]			COST [€]			
	Truck from	Truck To	Truck ID	Waybill ID	PID ID	Truck dept (ex truck origin)	PID req (ex fact)	PID	Waybil	Truck	FTL cost [€/truck]	LTL cost [€/waybill]	*LTL Min cost [€/wb]*	*LTL Adder cost [€/wb]*
FACT --> OCP	Fact1	OCP	fa1_o1	fa1_cu5_od1	29	2	2	5,606	5,606	5,606	€ 1,050			
			fa1_o2	fa1_cu2_od1	27	2	2	3,845	3,845	3,845	€ 1,050			
	Fact2	OCP	fa2_o1	fa2_cu2_od2	17	2	2	5,934	5,934	5,934	€ 350			
			fa2_o2	fa2_cu8_od1	10	1	1	3,902	3,902	3,902	€ 350			
			fa2_o3	fa2_cu2_od5	4	1	1	4,828	6,174	6,174	€ 350			
					7		1	1,346						
			fa2_o4	fa2_cu2_od4	11	1	1	478	5,698	5,698	€ 350			
					12		1	5,220						
			fa2_o5	fa2_cu2_od3	20	2	2	5,935	5,935	5,935	€ 350			
			fa2_o6	fa2_cu8_od1	18	2	2	2,220	2,220	2,220	€ 350			
OCP --> CUSTOMER	OCP	Customer2 (Zone10)	cu2_od1	fa1_cu2_od1	27	4	2	3,845	3,845	3,845	€ 1,450			
			cu2_od2	fa2_cu2_od2	17	4	2	5,934	5,934	5,934	€ 1,450			
			cu2_od3	fa2_cu2_od3	20	4	2	5,935	5,935	5,935	€ 1,450			
			cu2_od4	fa2_cu2_od4	11	3	1	478	5,698	5,698	€ 1,450			
					12		1	5,220						
			cu2_od5	fa2_cu2_od5	4	3	1	4,828	6,174	6,174	€ 1,450			
					7		1	1,346						
	OCP	Customer5 (Zone8)	cu5_od1	fa1_cu5_od1	29	5	2	5,606	5,606	5,606	€ 4,550			
	OCP	Customer8 (Zone2)	cu8_od1	fa2_cu8_od1	10	3	1	3,902	6,122	6,122	€ 2,250			
					18		2	2,220						
FACT--> CUSTOMER (FTL)	Fact2	Customer10 (Zone6)	fa2-cu10_d1	fa2_cu10_d1	5	1	1	6,388	6,388	6,388	€ 1,900			
			fa2-cu10_d2	fa2_cu10_d2	9	1	1	978	4,226	4,226	€ 1,900			
					15		1	3,248						
	Fact2	Customer4 (Zone1)	fa2-cu4_d1	fa2_cu4_d1	16	1	1	5,388	5,388	5,388	€ 1,950			
	Fact2	Customer5 (Zone8)	fa2-cu5_d1	fa2_cu5_d1	8	1	1	4,215	4,215	4,215	€ 4,550			
	Fact2	Customer7 (Zone6)	fa2-cu7_d1	fa2_cu7_d1	19	2	2	4,292	4,292	4,292	€ 1,900			
	Fact3	Customer7 (Zone6)	fa3-cu7_d1	fa3_cu7_d1	25	2	2	6,421	6,421	6,421	€ 2,050			
	Fact3	Customer8 (Zone2)	fa3-cu8_d1	fa3_cu8_d1	22	2	2	5,821	5,821	5,821	€ 2,250			
	Fact3	Customer9 (Zone6)	fa3-cu9_d1	fa3_cu9_d1	23	2	2	5,387	5,387	5,387	€ 2,050			
FACT--> CUSTOMER (LTL)	Fact2	Customer6 (Zone2)	LTL	fa2_cu6_f10	21	LTL	2	10	10	LTL		€ 11.97	€ 9.57	€ 2
		Customer8 (Zone2)		fa2_cu8_f2	6		1	2,998	2,998			€ 2,018.85	€ 9.57	€ 2,009
	Fact3	Customer6 (Zone2)		fa3_cu6_f9	26		2	262	262			€ 232.45	€ 9.57	€ 223
		Customer9 (Zone6)		fa3_cu9_f9	24		2	1,737	1,737			€ 1,280.35	€ 8.61	€ 1,272
	Fact4	Customer10 (Zone6)		fa4_cu10_f5	3		1	5	5			€ 9.04	€ 9.04	€ -
FACT-->BCN	Fact1	BCN	fa1_b1	fa1_cu3_b1	28	2	2	1,036	4,660	4,660	€ 2,300			
					30		2	3,624						
	Fact2	BCN	fa2_b1	fa2_cu1_b1	13	1	1	315	315	6,312	€ 1,850			
				fa2_cu3_b1	2		1	2,452	2,452					
				fa2_cu7_b1	1		1	2,618	3,545					
					14		1	927						
BCN --> CUSTOMER	BCN	Customer3 (Zone7)	LTL	fa1_cu3_b1	28	LTL	2	1,036	4,660	LTL		€ 2,505	€ 25.92	€ 2,479
					30		2	3,624						
		Customer1 (Zone1)		fa2_cu1_b1	13		1	315	315			€ 119	€ 8.32	€ 111
		Customer3 (Zone7)		fa2_cu3_b1	2		1	2,452	2,452			€ 1,496	€ 25.92	€ 1,470
		Customer7 (Zone6)		fa2_cu7_b1	1		1	2,618	3,545			€ 775	€ 6.87	€ 768
					14		1	927						

4.4.3. Adaptability of the model

The defined program serves to solve a very specific case; but the power of having this built-in algorithm is to provide a baseline code as well as familiarity and understanding of the program language and solving algorithm. Then the model can be applied to any other dataset as long as the network remains the same, or can be more or less easily adapted to similar networks. Also, additional aspects that are relevant in practice could be easily added. For example, we could include a new PID parameter with its latest allowed delivery date (and a restriction to enforce it). One could also include certain norms, such as minimum volume used by a route, or an "opening cost" (i.e. fix cost for a route if it is used at least one time, relating to potential set up investment). If necessary, instead of the weight, the sizes of the PIDs and the trucks capacity could be measured in volume (loading meters), etc. And of course, it could be considered the option to include the input data and parameters not as fixed, but as stochastic variables better representing reality. But this, clearly, would result in a much harder problem in the area of stochastic optimization.

Thinking broader, one could even consider this model as a baseline for factory location assignment: we could set up artificial origin nodes with 0-cost legs to different factories, which would have assigned certain node cost equal to production cost (plus the cost of the transportation route afterwards). **Figure 10** represents such scenario.

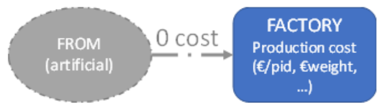

Figure 10. Artificial nodes to model factory location assignment decisions.

It has also been mentioned in earlier sections that this case study considers only the economic cost, however this could be easily extrapolated to any other quantitative measure (i.e. environmental impact: Co2eq).

4.4.4. Computational limitations

We are in front of an integer linear problem (binary to be more specific), where for each indivisible PID we must choose whether to take or not to take a certain route. This results in a (computationally) very expensive exact program, not applicable at the scale of a multinational network. The above is an example of the maximum complexity that this exact algorithm can cope with, which is extremely far from any real-world application in logistics. Certain adjustments have been done to help the algorithm: providing an initial solution, ensuring no symmetries apply, playing with constraints so that the model can easily recognize null variables… but improvements are rather negligible considering how distant we are from the order of magnitude need for real production.

4.4.5. Next Steps - Extrapolating the model to a practical solution

As science advances, so does the power of computational systems (i.e. parallel computing), so one could imagine a future with enough computational strength to solve the prior exercise at a real scale (millions of shipments/variables) in minimum, practical time. However, we are still far away from this, so for the time being other solutions shall be applied to use the prior model for real applications. The use of heuristics would be very interesting in this case. However, to have an idea of the expectable % gap with respect to the optimum, and evaluate whether the solutions provided by a particular heuristic represent a good trade-off between speed and solution quality, it is important to count on exact methods that allow to compare heuristic and exact solutions at least for small instances.

Another solution would be to split the problem into two steps:

1- Continuous relaxation allowing the division of the PIDS. This would determine which lanes should be opened.
2- Assignment of this "opened" lanes to each PID. This would solve the same integer exact problem formulated previously, but with a drastic decrease in the number of variables.

As in heuristics, here we would again be trading accuracy for speed.

4.4.6. Additional note on the exact formulation

It is to be mentioned that the feasible region of the formulation of the exact model (*section 4.3*.) contains solutions that are not really feasible for our problem, but this is not an issue, as is explained next to avoid misunderstandings.

As defined, we do not allow waybills to split at the last FTL truck used. In case of the OCP there are two FTL trucks (fact->ocp->country hub). Hence, in a feasible solution of our formulation might wrongly allow waybills to split in the first FTL truck from factory to the OCP. It could also happen that a waybill is wrongly consolidating two PIDS that do not even travel in the same truck at all, like in the following example:

- PID *k=1* (from fact *i=1* to customer *j=1*) which travels in the *l=1* truck from factory to BCN hub: *b(1,1)=1*.
- PID *k=2* (from fact *i=1* to customer *j=1*) which travels in the *l=1* truck from OCP to BCN hub: *ob(1,1)=1*.

In case this would happen, the LTL cost of this waybill would be charged twice: both *ltlbcn* and *ltl***o***bcn* would be rated considering the total waybill weight (*PID1+PID2*). Since we are minimizing, such a solution will never appear as the optimal solution of the formulation. So, it is no necessary to increase the size of the formulation by including new constraints that forbid such solutions.

4.5. ENVIRONMENTAL IMPACT[1]

The company involved in this analysis represents an industry whose products require materials extraction and significant manufacturing processes. In addition, through their lifecycle those products will also consume a significant amount of energy. Given all this, it is to be considered that their transportation and logistic activities only represent around 2% of the total environmental footprint. Although this is a small fraction, any improvement towards sustainability is always welcome.

The goal of the study case analysed is to reduce network inefficiencies. Although the objective function in this thesis has been defined as an economic cost, this could easily be extrapolated to any other quantitative measure such as greenhouse gas emissions.

Even if minimizing cost, we can consider that both economic cost and greenhouse gas emissions are positively correlated with the shipped tones-km. In that case we can consider that minimizing cost will also externalize to Co2 emissions.

On the company's case, 15,000 tones of products are shipped annually, covering an average distance of 2,000km with standard truck mode (Co2 factor: 140g/tkm, see EEA2021). If the model could be used to improve the network by 10%, this would result in 420 tones of Co2 savings per year.

Note the above is a very rough approximation that only intends to provide an order of magnitude.

4.6. BUDGET[1]

The cost of developing this study accounts for 25,000€ mainly driven by human resources, as other costs (amortization of computer's hardward and software such as licenses, energy, etc) are negligible.

- Project author: 400h * 50€/h = 20,000€
- Project director: 50h * 100€/h = 5,000€

[1] *The "Environmental Impact" and "Budget" sections have been included in this document only to comply with the formal guidelines of ETSEIB's master thesis. However, given the nature of this thesis those sections do not really apply.*

5. OTHER INVESTIGATIONS ON ACADEMIC APPLICATIONS TO REAL LOGISTICS NETWORK DESIGN

This project has been very much focused on an exact algorithm to optimize the network operation, however there are multiple other approaches that could as well proof very insightful. The below list summarizes the best initiatives that have been identified in this regard by reviewing the literature, and from the interviews carried out:

- Digital twins on dynamic transport optimization
  Extrapolate Wardrop's traffic principles (see En.wikipedia.org. 2021) to the logistics field: creating digital twins of the loads and letting them decide dynamically and selfishly on their best routing alternative (user equilibrium). Compared to the system optimal, this user equilibrium may be easier to achieve while also providing very good results.

- Resilient planning enabled by probabilistic simulation
  Having a good forecast is key to optimize the transportation network and avoid inefficiencies. Probabilistic simulation allows parameters to be random variables, hence the result is not a plan that best fits a unique situation but rather a set of potential future scenarios. Given today's VUCA world, having this risk-assessed planning is key.

- Forecasting with Artificial Intelligence
  Also relating to the importance of having a precise forecast, there are currently multiple initiatives that intend to use AI to find very in-depth features that can improve the definition of the forecast pattern.

- Real time visibility
  Today, the main focus of the majority of companies is to achieve a proper visualization platform that can support shipment tracking at any moments. Although this relates more to IT disciplines (farther away from operations research), it is listed here given the criticality it has proven to have during the Covid Pandemic. The conclusion is clear, disruptions are frequent and in such global SC having a proper visibility is key for agility, resilience, and overall Supply Chain success.

CONCLUSIONS AND NEXT STEPS

It is clear that the ability to use mathematical models to get the maximum insights from data is very powerful in helping companies make the right decisions. This is especially true in the case of logistics network design, where, although highly volatile and unpredictable, a theoretical optimum exists. While typical simulation tools can provide good local insights, optimization models capable of finding the most efficient global network given any new input can be very promising.

Inspired by a case from a real company, this project has defined an exploitation problem of a small section of a multinational network. It has formulated an exact program that, through the brunch and cut algorithms of the FICO XPRESS software, has proven to successfully provide the best routing of each PID so that total system's cost is minimum, balancing between direct routing and shipment consolidation. This program shall be used as a baseline and calibrator of further heuristic models that could deal with larger datasets and be applied to real life.

This is just one of the many initiatives that aim to improve the use of data models and technology to gain better insights in the logistics business. After all, the university continues to progress, and with it the opportunity to better process information and reduce inefficiencies; as long as we are able to integrate academic advances to business practice.

BIBLIOGRAPHY

BIBLIOGRAPHIC REFERENCES

Books and papers

Darvish, M. and Coelho L.C. (2018) *Sequential versus integrated optimization: Production, location, inventory control and distribution*. European Journal of Operational Research 268(1), pages 203-214.

Lekić, M., Rogić, K., Boldizsár, A., Zöldy, M. and Török, Á. (2021) *Big Data in Logistics*. Periodica Polytechnica Transportation Engineering, 49(1), pages 60-65.

Pires, M.C., Parreira, R. and Frazzon, E.M. (2021) *Integrated operational supply chain planning in Industry 4.0*. International Journal of Integrated Supply Management 14(1), pages 28-49.

Tordecilla, R. D., Juan, A. A., Montoya-Torres, J.R. , Quintero-Araujo, C. L. and Panadero, J. (2021) *Simulation-optimization methods for designing and assessing resilient supply chain networks under uncertainty scenarios: A review*. Simulation Modelling Practice and Theory 106, 102166.

Used web pages

EEA 2021. *European Environment Agency*. [online] Available at: <https://www.eea.europa.eu/> [Accessed July 2021].

En.wikipedia.org 2021. *John Glen Wardrop - Wikipedia*. [online] Available at: <https://en.wikipedia.org/wiki/John_Glen_Wardrop> [Accessed June 2021].

Freightos 2021. *What is FTL (Full Truck Load)?* [online] Freightos. Available at: <https://www.freightos.com/freight-resources/ftl-full-truck-load-freight-term-glossary/> [Accessed June 2021].

COMPLEMENTARY BIBLIOGRAPHY

To compliment my prior knowledge in combinatorial optimization I have used some classic references as well as more recent works dealing with problems similar to the object of this thesis:

Ahuja, R., Magnanti, T. and Orlin, J. (1993) *Network flows*. Upper Saddle River. Prentice-Hall.

Contreras, I. (2017) *Hub location problems*. PhD thesis. Universitat Politècnica de Catalunya.

Eitiveni, I., Kurnia, S. and Buyya, R. (2017) *Sustainable Supply Chain Management: Taxonomy, Gaps, and Future Directions*. School of Computing and Information Systems. The University of Melbourne, Australia.

Heipcke, S. (2010) *Modelling and problem solving with Xpress-Mosel. FICO Xpress Training*. Available at: <http://www.fico.com/xpress> [Accessed 2020, 2021].

Hu, T.C. Kahng (2018) *Linear and integer programming made easy*.

Martí, R. (2010) *Algoritmos Heurísticos en Optimización Combinatoria.*

Pastor, R., Corominas, A., Coves, A., Lusa, A. (2015-2016) *Mètodes Quantitatius a la Cadena de Subministrament, Transparències.* ETSEIB, UPC.