Optimal Location of Cross-Docking Centers for a Distribution Network in Argentina

by

Fernando G. Dobrusky

M. Sc. in Industrial Engineering, Columbia University, 2001 Industrial Engineer, University of Buenos Aires, Argentina, **1999**

Submitted to the Engineering Systems Division In Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Logistics

at the

Massachusetts Institute of Technology

June 2003

@ 2003 Fernando **G.** Dobrusky, **All** rights reserved The author hereby grants to M.I.T permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part.

> **Engineering Systems Division** May 9th, **2003**

Certified **by**

Cynthia Barnhart Professor of Civil & Environmental Engineering and Engineering Systems Co-Director, Center for *//yansportation and Logistics* Thesis Supervisor

Accepted **by**

Signature of Author

Yossi Sheffi Professortof Civit & Environmental Engineering and Engineering Systems Co-Director, Center for Transportation and Logistics

 $-2-$

 $\label{eq:2.1} \frac{1}{2} \sum_{i=1}^n \frac{1}{2} \sum_{j=1}^n \frac{$

Optimal Location of Cross-Docking Centers for a Distribution Network in Argentina

by

Fernando **G.** Dobrusky

Submitted to the Engineering Systems Division on May **9, 2003** in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Logistics at the Massachusetts Institute of Technology.

ABSTRACT

The objective of this thesis is to define an optimal distribution network for Argentina using Cross Docking Centers. The products to be delivered are in a Central Warehouse in Buenos Aires, the main city and port of Argentina. These products have to **be** distributed to hundreds of locations in the interior of the country. These locations have lower levels of demand and are at great distance from Buenos Aires.

In order to achieve efficient distribution, cross docking centers and a third party carrier distribution are utilized. To find the optimal number, size and location of cross docking centers and which cross docking center or carrier should supply each location, two models are developed.

The first model is a location-allocation model in which capacities of the cross docking centers are not considered constraining. In this case, the model is complemented **by** a heuristic approach that is used to find a near-optimal feasible solution. The second model, a capacited location model, is more complex, taking into account the demands of each location and defining the optimal location of cross docking centers and their respective capacities.

Both models are analyzed with the data representing the distribution of pharmaceutical products in Argentina in **1999.** The models' solution generates savings of **5%,** compared to the current network that was designed based on intuition and other external factors, without the use of an optimization tool.

Thesis Supervisor: Cynthia Barnhart

Title: Professor of Civil **&** Environmental Engineering and Engineering Systems Co-Director, Center for Transportation and Logistics

to my wife, **Gabi** who successfully optimizes our happiness every day since **1988**

 \sim

 $-6-$

Table of Contents

 $\frac{1}{2} \left(\frac{1}{2} \right)$

Acknowledgments

First of all, **I** want to thank the Fulbright Commission, the YPF Foundation, the Interamerican Development Bank and the De Sa Interamerican Foundation, who supported my learning experience at Columbia University in New York and here, at MIT.

There are many people that made my academic experience at MIT and the writing of this thesis an excellent experience in my life. **I** want to thank all my professors and classmates that taught me so much in such a short time.

I want to thank Yasmine **El Alj,** who helped me learn and understand OPL Studio. **I** also want to thank Attilio Bellman who gave me good Italian advice.

Especially, **I** want to thank my thesis advisor Cynthia Barnhart. She is not only a brilliant professor, she is also a wonderful person. **I** want to thank her for being my advisor, for teaching me so much and for being an excellent guide during these times. **I** want to thank her for all the time she dedicated to help me.

Finally, as **I** stated before, **I** dedicate this thesis to my wife, Gabi.

 $-10-$

 $\label{eq:2.1} \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\right)^{2} \left(\frac{1}{\sqrt{2}}\right)^{2} \left(\$

 $\label{eq:2.1} \frac{1}{\sqrt{2}}\int_{0}^{\infty}\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{2\pi}}\right)^{2\alpha} \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{2\pi}}\right)^{\alpha} \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}$

1. Introduction

1.1. Thesis Overview

The objective of this thesis is to define an optimal distribution network for Argentina using Cross Docking Centers. The products to be delivered are in a Central Warehouse and Distribution Center in Buenos Aires, the main city and port of Argentina. These products have to be distributed to hundreds of locations in the interior of the country, each with lower levels of demand and at great distance from Buenos Aires.

In order to achieve efficient distribution, one solution is the utilization of Cross Docking Centers **(CDC),** where products are received from truck loads and reloaded to smaller trucks and vans. Another solution is to distribute through a carrier and pay the carrier's tariff. Probably the optimal solution is a combination of both, with some locations supplied **by** a **CDC** and others **by** a carrier.

To find the optimal number of CDC's, where they should be located, which size they should be, and which locations should be supplied **by** which **CDC** or if they should be supplied **by** a carrier, two models are developed.

The first model is a location-allocation model where the capacities of the CDC's are not considered as a constraint. In this case, the model is complemented **by** a heuristic approach that is used to find a feasible solution that is close to optimal.

The second model, a capacited location model, is a more complex model that takes into account the demands of each location and defines the optimal location of CDC's and their capacity.

Both models are analyzed with the data of the distribution of products from the pharmaceutical industry in Argentina. Compared to the current network, the results of the models show savings of **5%** of the total cost when the Cross Docking network is optimized. The current network was designed based on intuition and other external factors, without the use of any optimization tools.

1.2. Introduction to cross-docking

Cross Docking is a well known approach to distribute products in a network. It can, in some industries, generate important improvements in service and cost reductions. **A** Cross Docking Center is also known as a transshipment point.

The basic concept is that a **CDC** is an open space with docks on both sides. Once products get to the **CDC,** the trucks are unloaded and the products are moved from the docks IN to different docks **OUT,** depending on the final destination. From the dock **OUT,** other trucks are loaded and they distribute the products to their final destination.

One common example of cross docking is the one used **by** Wal-Mart. At the CDC's, vendors deliver their products with their own trucks at one side. The products are divided and assigned to different docks, depending on the supermarket which they are destined, and then they are loaded on Wal-Mart trucks that deliver products from different vendors to each supermarket.

In other cases cross-docking might include other processing steps. It is not rare to find at CDC's, operations like stamping, repacking, bundling, or performing quality controls.

There are many papers that focus on analyzing models to define the best shape and number of docks for a **CDC,** depending on the flux of products.

In the case of this thesis, the type of **CDC** that will be analyzed is one sometimes known as "one source". In this case, all products are from the same source and the objective of the **CDC** is to divide the cargo into smaller trucks and vans and improve distribution, both in service time and in cost.

The improvement generated **by** this type of **CDC** is described in the following example:

When a truck is delivering orders, a stop for a delivery can take from a few minutes to hours, depending on the destination. If a large truck has 400 orders to deliver, it can take more than **10** days to finish with all deliveries. If those 400 orders are divided into **10** smaller vehicles, all orders can be delivered in one day. If all the destinations are relatively close to the origin, the **10** smaller vehicles can deliver from the central location. If the distance from the origin to the locations is long (like in the case analyzed here), going from the original location (in the case of Argentina, from Buenos Aires) to the destination with **¹⁰** small vehicles can be very uneconomical. The **CDC** strategy, of traveling with the large truck to a mid point close to the destinations and transshipping the orders to smaller trucks that make the deliveries, minimizes the cost.

Other factors, like infrastructure of roads and regulations, can render delivery with large trucks impossible, also necessitating use of the **CDC.**

1.3. Optimization of cross-docking operations

Cross docking operations are a recurring research topic in Optimization. In the design of a **CDC** operation, there are many aspects that can be optimized:

- Number of CDC's
- **"** Locations of CDC's
- **"** Allocation of destinations to a **CDC**
- **Optimal number of docks (both in and out)**
- * Optimal shape (a line with docks in both sides, a "T" shape, etc)
- **"** Optimal distribution from **CDC** to destinations (similar to a traveling salesperson problem)
- \bullet Etc.

Usually, these questions are answered separately, or just two questions addressed simultaneously. Once a solution is found for one question, then it is taken as fixed when addressing the next optimization question.

The following figure depicts a common sequence used in addressing design questions of a **CDC** operation:

This sequential optimization approach has a distinct disadvantage: a global optimum will likely not be found for the system. As an example, the optimal distribution from a **CDC** to the customers might be improved if one customer is re-allocated to another **CDC.** The fact that the allocation was defined before, however, makes re-allocation infeasible.

Now, with tremendous computer power available at low cost, achieving global optimality is perhaps within reach.

1.4. Literature Review

There is a vast literature that deals with optimization for location models and Cross Docking, and it would be impossible to go over all in these few paragraphs. This section is just a list of some of the most relevant for the topic of the thesis.

There is a recently edited book **by** Zvi Drezner and Horst W. Hamacher called "Facility Location, Applications and Theory" which is a compilation of different papers that deal with different aspects of Location problems. The first chapter of the book, a paper **by** Zvi Drezner, Kathrin Klamroth, Anita Schobel and George **0.** Wesolowsky is called "The Weber Problem". In this chapter they describe the history of the location problem. They start with Pierre de Fermat **(1601-1665)** who stated the problem as follows: "given three points in the plane, find a fourth point such that the sum of its distances to the three given points is a minimum". The authors show different solutions for the problem presented through the history of Science.

The solution of using Cross Docking Centers to solve a transportation problem is analyzed **by** John **G.** Klincewicz in his paper "Solving a Freight Transport Problem Using Facility Location Techniques". In this paper, the author refers to a **CDC** as a Consolidation Terminal and uses it to replace direct shipping. He includes in the problem variation in the inventory costs and proposes a heuristic algorithm to solve the problem.

Another paper that analyzes the location of transshipment points is **"A** Branch-And-Bound Algorithm for the Transportation Problem with Location of **p** Transshipment Points" **by** Alfredo Marin and Blas Pelegrin. In this paper, the authors analyze the solution for a predetermined number of transshipment points but they add the fact that there is flow of material from the customers to the transshipment point (return).

Yuri Levin and Adi Ben-Israel, in their paper **"A** Heuristic Method for Large-Scale Multi-Facility Location problems", develop an algorithm that locates the facility and reassigns customers to facilities using the heuristic of Nearest Center Reclassification.

The combination of facility location, customer allocation and routing is analyzed **by** Gilbert Laporte, Eric Gourdin and Martine Labbe, in their paper "The Uncapacitated Facility Location Problem with Client Matching". There, they propose an algorithm to match clients in one route, in order to get a better solution than solving the different parts of the problem separately.

The use of a set partitioning approach for a capacitated location problem is developed **by** Roberto Baldacci, Eleni Hadjiconstantinou, Vittorio Maniezzo and Aristide Mingozzi and the model is described in the paper **"A** New Method for Solving Capacitated Location Problems Based on a Set Partitioning Approach".

With respect to optimization within a Cross Docking Center, there is current research at the Georgia Institute of Technology, led **by** John **J.** Bartholdi III and Kevin R. Gue. In their paper "Reducing Labor in an LTL Cross-docking Terminal" they analyze the optimal movements of materials in order to minimize labor costs. John **J.** Bartholdi III and Kevin R. Gue are also researching the optimal shape of a Cross Docking Center. Their research in this field can be found at the Georgia Tech Website.

-18-

2. Problem definition

2.1. Introduction

The basic problem that is being assessed in this thesis is the distribution of goods from a central location to other cities located at a great distance from it. It is a transportation problem, and the objective function to minimize is the total cost of the distribution.

The model will be applied to the distribution of pharmaceutical products in Argentina, with its Central Location in Buenos Aires.

When initially approaching the problem, many possible solutions can be analyzed:

One solution is to have regional warehouses and distribute to the cities from there. For many products, this is the standard solution, but not for the pharmaceutical industry in Argentina. This alternative could be analyzed from an economical point of view, but the large number of SKUs for this industry, the high cost of the products, the **highly** regulated storage conditions and the perishable characteristic of these products make it uneconomical.

Another solution is to distribute the products from the central location directly to all the cities with a private fleet. This would also be an uneconomical solution, because most of the cities have very low demand, thus requiring a very large fleet with very low levels of utilization.

Another solution is to use a carrier. In some cases this can be the most economical solution, but in other cases, the premium charged **by** the carrier might be too high.

The best solution for most of the cities, and the one generally used in practical situations, is the establishment of cross-docking centers **(CDC).** These CDC's are located at several points in the country. Daily, the orders supplied **by** the **CDC** to the cities are shipped from Buenos Aires **by** Truck Loads (TL). In the **CDC,** the truck is unloaded and the orders for each city are re-located in other docks. There, they are loaded into smaller trucks or vans that distribute them to the cities, at a lower cost.

The objective of this thesis is to define a model to find the optimal solution to these questions:

- How many CDC's are needed?
- Where should they be located?
- * What is the optimal size of each **CDC?**
- * Which cities should be supplied **by** each **CDC?**
- * Which cities should be supplied **by** a carrier from the central location?

2.2. The use of cross-docking centers in Argentina

Argentina, as shown in the following map, is a large country. Buenos Aires is the main city, in which is located the main port of the country that connects Argentina with the rest of the world.

The maximum linear distance from North to South is 3,694 km **(2,295** miles), and from East to West 1,423 km **(884** miles).

These distances, combined with mountainous geography in the North and West, make the distribution of goods a complicated problem. This problem is even

more complicated when infrastructure conditions are considered. Some roads that interconnect cities are not appropriate for certain types of trucks.

In order to understand the logistics in Argentina, it is important to understand where the population is located. Argentina has **37,812,817** people (July 2002 est.) with more than 12,000,000 people living in Buenos Aires and its surroundings. From a distribution perspective, the Buenos Aires metropolitan area usually accounts for 40 to 50% of the country's demand for most types of products.

Argentina is a very centralized country. Buenos Aires is the capital city and its main port. The transportation infrastructure (the main rail and roads) are centered in Buenos Aires and, as a consequence, while there is a connection from other cities to the center (Buenos Aires) there is no good interconnections among other cities.

As a result of these factors, most Central Distribution Centers for most companies operating at a national level in Argentina are located in the Buenos Aires area.

The problem of distribution to the interior has always been a challenge for most companies. Many solutions have been implemented: from regional warehouses in other important cities like Cordoba or Mendoza, to the regular use of LTL (less than truckload) carriers for locations with less demands to many other more creative solutions. **Of** course, all the solutions differ from industry to industry.

One common solution for some types of industries (among them the pharmaceutical industry) has been the use of Cross Docking Centers **(CDC).**

In these operations, where the central warehouse and distribution center is in Buenos Aires, truckloads are sent from Buenos Aires to the CDC's. Once they get to the **CDC,** the trucks are unloaded and the products are moved from the dock IN to different docks **OUT,** depending on the region of the final destination. From the dock **OUT,** smaller trucks or vans are loaded and distribute the products to the final destination.

Historically, the decision of where to locate these CDC's is made based on the importance of the cities and historical factors related to the companies. It is the objective of the thesis to review the benefits of having optimal **CDC** locations.

2.3. Modeling the operation

The next step in this thesis is to create a model that describes the operation. This model, defined **by** parameters and variables, will be used to identify the optimal network for the operation.

In order to start modeling the operation, it is useful to have a diagram of the network:

Small City **-** Weekly Delivery

Small City Distribution- From **CDC** to small city once a week

Large City Distribution **-** From **CDC** to Large city, daily

Carrier Transportation **-** From BS **AS** to destination

TL **-** From BS **AS** to **CDC**

It can be seen in the model that there are two different types of destinations: a small city, which needs weekly delivery; and a large city, which has a larger demand and needs daily delivery.

There are two different ways in which the product can be delivered from Buenos Aires to the final destination:

- a) Through a **CDC,** as described before; or
- **b)** Directly through a carrier from Buenos Aires.

One objective of the model is to consider both possibilities and find the best choice for each destination.

2.4. Parameter definition

Values of demand and cost have been slightly changed to maintain confidentiality with the source of the data. This does not affect, however, the results or analysis of the problem. The data belongs to a Third Party Logistics (3PL) firm in Argentina that in **1999** was in charge of the warehouse operations and distribution for the five most important pharmaceutical laboratories.

The cities: For the purpose of this analysis, we consider **687** Argentinean cities (called CITY in the model) with consumption of pharmaceutical products in the year **1999.** We exclude from our analysis Buenos Aires and its surroundings. These locations have direct distribution from the central location. (For a list of cities considered, see Appendix **1.)**

The possible CDC's: The 24 most important cities (called in the model **CDC's)** account for **80%** of the total demand. Because the most important savings are generated **by** distributing locally from the **CDC** (the distribution from the **CDC** to the city where the **CDC** is located), these 24 locations are defined as potential CDC's. (For potential **CDC** locations, see Appendix **1.)**

The demand: The demand is based on orders from each city for pharmaceutical products from the **5** leading laboratories, defined in **kg.** Because distribution is direct to the pharmacies, there is small variability in the size of the orders. So it is reasonable to say that demand is proportional to the space (volume) needed in the truck and in the **CDC,** and the number of orders/deliveries. (For a list of demands see Appendix **1.)**

Frequency: In the actual distribution, there are different frequencies which often change. In this model, we assume that cities that are potential CDC's must be supplied every working day (20 times/month). Moreover, we assume that the rest of the cities must be served once every week (4 times/month). (See list of frequencies in Appendix **1.)**

Fixed cost of CDC's: This cost, for the model, will include the cost of the TL transportation from Buenos Aires to the **CDC** and all costs associated with the operation of the **CDC.** The costs, defined in Argentinean pesos per month, vary from potential **CDC** to potential **CDC** (different distances from Buenos Aires, etc.) and for each potential **CDC,** depending on capacity.

In order to calculate the costs and size for each potential **CDC,** we take into account the following costs:

- . Fixed general (supervising cost, independent of the location and capacity);
- Administration (depends on the capacity);
- Rent (proportional to the square root of the capacity);
- Labor (proportional to the capacity); and

. Inbound transportation (depends on the distance to Buenos Aires and the capacity. Capacity defines the number and size of trucks in the fleet assigned to the **CDC** for the Buenos Aires **- CDC** route.).

These costs are divided into possible capacity levels, from **50,000** to **300,000 kg** per month in intervals of **50,000 kg.** The data is attached in Appendix 2.

Distribution costs from a CDC: Because the exact distances **by** road from each point to each different point are difficult to get, they are calculated as the linear distances from each potential **CDC** to each city. Longitude and Latitude coordinates are used for each city and the Euclidean distance is calculated. The total cost is a function of the frequency; the demand of the city; and the distance. The distance influences the cost of moving the products from the **CDC** to the city in a van or a small truck. The demand influences the number of deliveries and stops the driver has to make. And finally, the frequency influences the number of trips per month each driver has to make along these routes.

Distribution costs from Buenos Aires by a carrier: To calculate these costs, the structure is the same as the distribution from a **CDC,** but the parameters are different. The distance is not as important for the carriers (they already have in place the network), but they charge more for each delivery.

 $-28-$

3. The methodology

In order to solve the **CDC** optimization problem, two different models are implemented:

- **"** a Location-Allocation Model; and
- **"** a Capacited Location Model.

In the following sections, both models are described and the results for each model are presented.

3.1. Location-Allocation Model

3.1.1. Introduction

In this model the capacity constraint is ignored. The model solution defines the optimal number and location of CDC's, and the allocation of the destinations to each of the CDC's.

The model has variables X **(ij)** that take value **1** if CITY i is assigned to **CDC j,** or **0** otherwise. As constraints, once source must be selected for each destination; all destinations must be supplied; and if a **CDC** is used, its fixed cost must be included in the total cost. The model's objective is to minimize total costs.

In order to deal with shipments **by** a carrier, the carrier is considered as **CDC** 1 **(j=1).** In this case the fixed cost is **\$ 1** (the idea is to "open" the carrier **CDC** only if it is needed, but to keep it from influencing the objective). The costs of distribution from **CDC** to the destination for the carrier are the prices that the

carrier charges to go from Buenos Aires to the destination, as described in chapter 2.

To standardize costs and to have a model that makes valid calculations, all cost parameters for the CDC's and distribution for each CDC-Destination pair are computed on a per month basis.

We complement this model a heuristic process that allows us to produce a feasible solution satisfying the CDC's capacities.

3.1.2. Description

This model, **a** Location-Allocation model, determines the **CDC** locations and allocates each destination to one **CDC.**

In this model, the following sets and parameters are defined:

- **"** CITY is the set of **687** cities that need to be supplied;
- . **CDC** is the set of **25** potential CDC's (24 locations and the carrier);
- . cf **(j)** is the fixed cost if **CDC (j)** is opened; and
- **"** c(i,j) is the cost to distribute from **CDC j** to CITY i.

In this model, the **CDC** is considered to have unlimited capacity and the relation between the allocated demand for a **CDC** and its capacity is reviewed after a solution is obtained from the model.

Any **CDC j** that has at least one positive value for an X **(i,j)** is considered to be an open **CDC** and the value for Z **U)** is **1.** The fixed cost, in these cases, is added to the objective. This situation is defined **by** the constraint:

forall(i in CITY) forall (j in **CDC)** $X[i,j] \leq Z[j];$

All cities i must be served completely (that is, the sum of $X(i,j)$ for all j must be 1 for each i). The model could require that all demand to each city be supplied from a single source (making X(i,j) a binary variable), but the LP relaxation, without these requirements, satisfies these constraint. This occurs because the model does not include **CDC** capacities, so once a route is the best solution for part of the demand of a city (going through a specific **CDC** and not another, and not with a carrier), then it is the best route for all of the demand.

Model in OPL Studio:

In order to run the model, OPL studio is used. Below is the code, which also provides a full description of the model:

*/** constants ***/**

range **CDC [1..25];** range CITY **[1..687];**

*/** constraint declaration ***/**

constraint coverage [CITY]; constraint open [CITY,CDC];

*/** Parameters */

 $float + cf[CDC] = ...;$

float+ c [CITY,CDC] **=** ...;

*/** Variables*/

var float+ X [CITY,CDC]; var float+ Z **[CDC];**

```
/* Objective Function*/
```

```
minimize (sum (i in CITY) sum (j in CDC) X[i,j] * c[i,j] + sum (j in CDC) Z[j] * cf
[j)
```

```
/*constraints*/
```
subject to{

forall(i in CITY) coverage[i]: sum(j in CDC) $X[i,j] = 1;$

```
forall(i in CITY) forall (j in CDC)
open [i,j]: X[i,j] \leq Z[j];
```
};

*/** Output display */

display(i in CITY, j in CDC $: X[i,j] > 0$) <c[i,j]>; $display(j in CDC : Z[j] > 0) cf[j];$

3.1.3. Heuristic approach to achieve feasibility with respect to capacity constraints

This Location-Allocation Model is simple and gives a solution, but it considers the capacity of the **CDC** to be infinite, which is not true.

In order to solve this problem, a predetermined capacity is assigned to each **CDC** (which is implicit in the fixed cost of the **CDC)** and then it is reviewed if the optimal solution is a valid solution.

If the solution is not a valid solution, because the aggregated demand for a **CDC** is more than the pre-supposed capacity, the capacity of that **CDC** is incremented **(by** changing the fixed cost) and the model is re-run.

The first time, the model is run with the minimum capacity for each **CDC** and then it is successively run with increased capacities until a valid solution is found.

An important limitation with this approach is that it will not be known if the solution obtained is the optimal solution. Because, for example, a **CDC** that was discarded **by** the model because of being expensive could be part of the optimal solution if the value of capacity (and thus the cost) was lower.

A new model called the Capacited Location Model, which considers all the demands and all the capacity/cost curves for each **CDC** is analyzed in the second part of this chapter.

3.1.4. Results

After running the model, below are the results obtained:

Solving the problem: first run

In the first iteration, it is not known which **CDC** will be used in the optimal solution, and how much demand each one will carry. In order to avoid assigning extra capacity to a **CDC,** a run of the model is made with the fixed costs for all CDC's assigned with the minimal capacity.

The result will not be feasible for the demands assigned to each **CDC** (in some cases **50,000 kg** per month will not be enough capacity even to supply the city where the **CDC** is located). But, in any case, it provides a first idea about the potential outcome.

Once the first result is obtained from OPL Studio, **18** out of 24 possible CDC's are opened. The quantity of the demands of the cities i that have $X(i,j) = 1$ are assigned to **CDC j.**

In the following table, the selected CDC's with their monthly demands (that is sum of all monthly demands of the cities assigned to each **CDC)** are shown. Also shown are the CDC's that comply with the constraint on the demand; the new capacity for the CDC's that do not comply; and the re-calculated fixed cost for the subsequent run.

CDC 1, Buenos Aires, is the demand that would be shipped **by** a carrier.

The total cost in this run is: **\$ 816,615.**

Iterations of the model

Based on the result of the first iteration, the model is run again with the new fixed costs values. For the second iteration, the total cost is: **\$ 877,109** and now only CDC's **6, 10,** 12 and **18** need more capacity.

Once more the same steps are repeated with the new values and for the 3rd run the solution is:

In this case the total cost is: **\$ 894,927.**

For the 4th and final run, the results are:

The 4th run gives a solution that is feasible. The aggregated demand for each **CDC** is less than or equal to the capacity associated with the fixed cost used.

The total cost for the 4th and final run is: **\$ 904,988.**

The total number of CDC's opened is **17.**

This solution is considered to be the best solution that can be obtained with this model and the application of the described heuristics.

3.2. Capacited Location Model

3.2.1. Introduction

In this model, the capacity of the **CDC** is part of the model. Each possible capacity for each possible **CDC** is considered as an alternative.

In this case, unlike the uncapcitated case, the model cannot be cast as a linear program, but rather must be formulated as a Mixed Integer Program (MIP). This is because the model's structure no longer ensures that the linear programming relaxation will produce integer decisions. Instead, it is possible to achieve solutions to the LP relaxation in which demands are served **by** opening CDC's partially, limiting costs to the maximum extent possible **by** paying only for exactly enough capacity to serve the demands.

This new model has more than **100,000** variables and constraints. To ensure tractability of this MIP, we add a constraint to create a more robust LP relaxation.

Unlike our previous model, in this case an optimal solution can be obtained, given the parameters. There is no heuristic involved.

3.2.2. Description

In our capacitated location model the following sets and parameters are defined:

- * CITY is the set of **687** cities that need to be supplied;
- **" CDC** is the set of **25** potential CDC's (24 locations and the carrier);
- **"** CAPACITY is the capacity of the **CDC.** They are set as six different alternative capacities (in increments of **50,000 kg** per month) from **50,000** to **300,000 kg** per month;
- * cf **(j,k)** is the fixed cost if **CDC (j)** is opened with a capacity **k;**
- . c(i,j) is the cost to distribute from **CDC j** to CITY i; and
- \bullet d(i) is the demand for city i in kilograms per month.

The non-negative, continuous variables X represent the proportion of demand at a city served **by** a particular **CDC.** X variables can be fractional when more than one **CDC** is needed to serve a city. While we could require the X variables to be binary, we believe that this would increase the complexity of the model and the change would not be a good tradeoff.

As in the location-allocation model, for **j=1** the **CDC** is considered to be distribution from Buenos Aires **by** a carrier. The associated distribution costs are the carrier costs, and there are no fixed costs. (In running the model, the carrier option is assigned a fixed cost of **\$1** to ensure that the model includes the carrier option in the solution only if it is needed).

Each **CDC j** with capacity **k** with **X(i,j,k)** strictly positive for at least one i is considered to be an open **CDC.** The value of Z **(j,k)** in these cases is 1 and the fixed cost of the **CDC j** with capacity **k** is added to the objective. This is defined **by** the constraint:

forall(i in CITY) forall (j in **CDC)** forall **(k** in CAPACITY) open **[i,j,k]: X[i,j,k] <= ZIj,k];**

And **by** the fact that Z **(j,k)** is a binary variable. If the Z-variables were defined as continuous variables, then the excess demand of one **CDC** (for example **0.10** of city i) could be served **by 0.1** of a different **CDC,** thereby incurring only **10%** of the fixed cost of this **CDC.** Enforcing the binary constraints ensures that even if X **(ij,k)** is strictly positive but small, **CDC j** at capacity **k** is fully opened.

Another constraint requires that each city i is served completely:

```
forall(i in CITY)
coverage[i]: sum(j in CDC) sum (k in CAPACITY) X[i,j,k] = 1;
```
Another constraint that was not in the previous model is the demand satisfaction constraint. In this case, the capacity of an open **CDC** must be greater than the sum of all the demands supplied from that **CDC:**

forall (j in **CDC)** forall **(k** in CAPACITY) demand **[j,k]:** Z[j,k]*cap[k] **>=** sum (i in CITY) X **[ij,k] *** d[i] ;

There is another constraint:

forall (j in **CDC)** onesize **[]:** sum **(k** in CAPACITY) **Z[j,k] <=1**

This constraint specifies that if a **CDC** of a specific capacity is opened in a location, there will not be another **CDC** opened at that location. This is a redundant constraint: no **CDC** would need more than **300,000** kg/month capacity (maximum capacity available) and opening two CDC's of capacities **A** and B at the same location would always be more expensive than opening one **CDC** of capacity **C,** with **C=A+B.** The reason to add this constraint to the model is to make it more robust. With this constraint, solutions that open up small portions of a second **CDC** at a location to handle overflow are excluded. This improves the bound on the optimal solution value provided **by** the LP relaxation, making the problem easier to solve with a **PC.**

Model in OPL Studio:

/* constants ***/** range **CDC [1 .. 25];** range CITY **[1..687];** range CAPACITY **[1..6];**

*/** constraints ***/**

constraint coverage [CITY]; constraint open [CITY,CDC, CAPACITY]; constraint demand **[CDC,** CAPACITY]; constraint onesize **[CDC];**

```
/* Parameters */
float + cf[CDC, CAPACITY] = ...;float+ c [CITY, CDC] = ...;float + d [CITY] = ...;float+ cap [CAPACITY] = ...;
```
*/** Variables*/

var float+ X [CITY, CDC, CAPACITY]; var int Z **[CDC,** CAPACITY] in **0..1;**

```
/* Objective Function*/
```
minimize (sum (i in CITY) sum (j in **CDC)** sum **(k** in CAPACITY) **X[i,j,k]** * c[i,j] + sum (j in **CDC)** sum **(k** in CAPACITY) **Z[j,k] *** cf **[j,k])**

subject to{

```
forall(i in CITY)
coverage[i]: sum(j in CDC) sum (k in CAPACITY) X[i,j,k] = 1;
```

```
forall(i in CITY) forall (j in CDC) forall (k in CAPACITY)
open [i,j,k]: X[ij,k] <= Z[j,k];
```

```
forall (j in CDC) forall (k in CAPACITY)
demand U,k]: Z[j,k]*cap[k] >= sum (i in CITY) X [i,j,k] *d[i
```

```
forall (j in CDC)
onesize [j]: sum (k in CAPACITY) Z[j,k] < = 1};
```

```
/* Output Printing*/
```

```
display(i in CITY, j in CDC, k in CAPACITY: X[i,j,k] > 0) <X[i,j,k], Z[j,k]>;
display(j in CDC, k in CAPACITY: Z[j,k] > 0) <Z[j,k]>;
```
3.2.3. Results

In this case, the following results are obtained:

In this case, the volume is adjusted to the capacity. That can be observed in the case of **CDC 3** (volume **199,985** and capacity 200,000), **CDC 9, 19** and 21. **By** allowing more possibilities for the capacity of CDC's, these results can be further improved, as discussed in the next section.

The value of the objective (total cost) in this case is **\$ 897,059.89.** When compared with the objective value of the Location-Allocation Model and the Heuristics described in the previous section, it is a reduction of **0.88%.**

3.3. Post processing

Once the solution is obtained, there are ways to increase the number of options and generate a better solution.

One way to do so is to generate more alternative capacities for the CDC's. It is expected that new possible capacities could improve the solution when: **1)** the total demand assigned to a **CDC** is very close to the CDC's capacity; and 2) the total demand assigned to a **CDC** is very far from its capacity.

In the case of **CDC 3** with assigned volume **199,985** and capacity 200,000, there are some cities that could be supplied from **CDC 3,** but the associated savings is not sufficient to enlarge **CDC 3.** Having the possibility of a 210,000 or 220,000 kg/month **CDC** at **3** would probably generate savings in distribution.

On the other hand, when the volume is much smaller than the capacity, for example, **CDC 25** with volume 19,146 and capacity **50,000,** a capacity of 20,000 would save fixed costs at the **CDC.**

To get the monthly costs for CDC's of these sizes, a new analysis would have to be made to assess the cost of the inbound transportation, labor, and all the **CDC** costs described in chapter 2.

Another possible improvement is to refine the cost calculation for distribution from the **CDC** to the cities. These costs are calculated as a function of the distance, volume and frequency. The true distribution costs can be obtained after the routes from the **CDC** to different cities are designed.

The problem of designing these routes can be formulated as a "traveling salesperson problem **(TSP)"** (Croes, **1958).** The objective is to minimize the distance (cost) when going from one point to several others, passing one time **by** each of them.

These post-processing improvements are not analyzed in this thesis but are recommended for models that are implemented in practice.

-46-

4. Analysis of Results

4.1. Introduction

Before comparing the results of the models, it is important to have an understanding of how decisions were made in Argentina. After this process is described, the results of the two models presented are compared. Then, the network defined in **1999** is compared to the best solution obtained **by** the models.

Here is the decision making process for **CDC** location for the pharmaceutical industry in Argentina in **1999** as narrated **by** the project leader of the team in charge of the decision.

The story of 1999

In **1999,** the team in charge of the project, which I lead, had little time and little expertise to make the decision. The operation had to be running after a few months and many activities depended on the solution: renting the warehouses that would act as CDC's, preparing them, recruiting and training workers and managers, purchasing of forklifts, IT integration, procedures definition, etc.

We talked to a consultant about it, but he said he needed at least a month to collect the data, analyze it and come up with the optimal locations. We didn't have a month to spend, or, at least, that is what we thought.

The decision was made **by** the sales manager, the operations manager, the transportation manager and me. We were in a conference room with a big map

of Argentina on the wall with the most important locations of the customers marked with red pins and the less important marked with green pins.

In that meeting, comments like "We need another **CDC** in the Andean Region, too much red there;" or "We have to have one **CDC** in Mar del Plata, I know someone that lives there that would be a great **CDC** manager;" or "We are already distributing for a chemical company in that region, even if we can't have the products in the same physical space, we should be able to get some synergy with both operations, at least in the administration", were the basis for the decisions made.

At the end of the three hour meeting, the decision was made: only 14 CDC's were selected. Which locations would be supplied **by** each **CDC** would be a decision to make in the future, once **CDC** sizes were defined. And the size was going to depend on the deals that the company would make renting the space.

In the following two years, two new CDC's were opened, one was closed and two were expanded.

In the following sections, the results of the models are compared with the costs of the decisions made **by** the team in **1999.** The costs defined here are applied to the decisions made **by** the team at that time in order to have a fair comparison.

4.2. Comparison of the models

In this section, the objective is to compare both models, the Location-Allocation Model (LAM) and the Capacited Location Model (CLM).

Comparing the complexity of the models, the LAM is a much simpler one. It is a linear program with **17,200** variables and **17,862** constraints. The CLM model is a MIP (Mixed Integer Programming) Model and has **103,200** variables and **103,912** constraints.

This complexity increase is reflected in the solutions times in OPL Studio: **3.2** seconds for LAM and 342 seconds for CLM.

Regarding optimality, CLM gets to an optimal value but LAM, used in combination with the heuristics, does not get to an optimal solution.

When the values of the objectives are compared, these are the results:

LAM: **\$ 904,988**

CLM: **\$ 897,060**

CLM generated a reduction in total cost of **\$ 7,928,** which is a reduction of **0.88%.**

This reduction is not so significant, especially when compared with the reduction obtained **by** the models compared to the costs of the decisions made in **1999,** as analyzed in the next section.

CLM is the best model to solve this problem, if problem size is manageable. To illustrate, when the CLM model is tested with **15** more possible CDC's, OPL Studio runs out of memory and does not produce a solution. In this case, the LAM and heuristic is the most effective approach for the problem.

4.3. Analysis for the pharmaceutical industry in Argentina

In order to analyze the cost reduction that is provided with our models, it is first necessary to compute the costs of the **1999** solution in a manner that is consistent with that of the LAM and CLM models. The **CDC** locations selected in **1999** are:

To get an allocation of cities to these CDC's and to define the **CDC** sizes needed, a LAM model is run with added constraints forcing the model to choose only the selected locations for CDC's. To do this in a simple way, very high fixed costs are allocated to the CDC's not included and the total number of CDC's (sum of $Z(j)$) is forced to be **15** (14 locations and the carrier).

The result obtained is:

The total cost is: **\$ 943,008.**

If this total monthly cost is compared with the total cost of the optimized result, there is a reduction of **\$** 45,948, which represents almost a **5%** improvement.

To analyze if this reduction is good enough to relocate the CDC's, a project evaluation should be made. The investment should be compared with the present value of the savings. It is not the objective of this thesis to make that decision.

In any case, it is clear from this case study that the use of optimization models can improve the design of a Cross Docking Network.

5. Conclusions

5.1. Introduction

In the first chapter of this thesis, the introduction, a cross docking operation is explained. **A** cross docking center is basically a transshipment point where products are unloaded from a vehicle and reloaded in another vehicle to continue the distribution without being stored. In that chapter is analyzed why, for certain industries, a Cross Docking Operation is a better alternative for distribution than direct shipping or regional warehouses. Also in the chapter, there is a brief analysis of how a Cross Docking Network can be optimized and why it is a recurrent topic in the Operations Research literature. At the end of the first chapter, there is a literature review providing further information for readers interested in this topic.

In the second chapter, defining the problem addressed in this thesis, the situation and the objectives of the analysis are thoroughly described in order to set the groundwork for modeling and optimizing the problem. There, it is defined that the solution must consider the use of a Cross Docking Center **(CDC)** or a direct shipment through a third party carrier when it is more economical. The questions that the model answers are:

- How many CDC's are needed?
- Where should they be located?
- **"** What is the optimal size of each **CDC?**
- * Which cities should be supplied **by** each **CDC?**
- . Which cities should be supplied **by** a carrier from the central location?

Since the model is applied to the distribution of pharmaceutical products in Argentina, there is also a review of the geography of Argentina. It is explained that Buenos Aires is the main city of the country and the most important Argentinean port on the Atlantic Ocean, and that Buenos Aires and its surroundings account for 40 to **50%** of the demand of most products. The rest of the demand is dispersed around hundreds of cities with much lower demand and spread within large distances.

Also in the second chapter, the cross docking operation is modeled and different sets and parameters are defined. There is a detailed explanation on how possible locations for CDC's were chosen, how costs were calculated and how the levels of demand were obtained.

The third chapter describes the methodology used to solve the problem. Here two different models are proposed to solve it. The first model is a Location-Allocation Model (LAM) and the second is a Capacitated Location Model (CLM).

The LAM is a linear model that considers the CDC's as uncapacitated and decides whether to open them or not. For the ones opened, the cities (destinations) are allocated and the fixed and distribution costs are added to the objective. The alternative of direct shipment through a carrier is included in the model as a **CDC** without fixed costs and the distribution costs are the carrier's tariffs. This model is combined with a heuristic algorithm. The algorithm is needed to find a feasible solution where the capacity of the **CDC** (associated with the fixed cost) is greater than or equal to the aggregated demand of the cities assigned to the **CDC.**

The CLM is a mixed integer programming (MIP) model. In this model the demand of the cities and the capacities of the CDC's are considered. The variables X **(i,j)** that in the previous model indicated if a city i is allocated to a **CDC j,** now is X **(ij,k)** and indicates that the city i is allocated to the **CDC j** that has a capacity **k.** This increases the complexity of the model and transforms it into a MIP formulation, because the LP relaxation opens only portions of CDC's. This model gets an optimal solution, but has the limitation of being solvable only for a limited number of possible CDC's.

Also in chapter three, there is a description on how the results of the model can be improved. This post processing can be done based on the results of the models. Two further improvements are proposed, the fine tuning of the **CDC** capacities and the optimization of the distribution from each **CDC** to the cities. The capacities can be adapted to the values of the demand, either **by** decreasing the capacity and saving fixed costs for the CDC's with low demand allocation; or **by** increasing the capacity for those which the assigned aggregated demand was equal to the capacity in order to lower distribution costs. The distribution from the **CDC** to the cities, originally calculated as a function of the demand, the distance and the frequency, can be optimized once the assignment to the **CDC** is done. This optimization of the routing from a point to other different locations is known in the Operations Research literature as a Traveling Salesperson Problem.

The fourth chapter is the analysis of the results obtained. It reviews how the original network design was achieved in **1999** in Argentina. Then the results of both methods are compared both on their results and computational time consumed in order to get to the solution. The CLM obtains an optimal solution at the cost of higher complexity. The result of the LAM is a suboptimal solution with a total cost **0.88%** higher than the CLM.

When the results are compared with the network designed in **1999,** the costs defined in the problem are assigned to the network designed in order to make the values comparable. To get these values, the LAM model is run with extra constraints that force the solution to choose the defined CDC's. The monthly cost for the **1999** network is **\$ 943,008** and the optimized monthly cost is **\$ 897,060,**

a reduction of almost **5%:** a very significant value in the third party logistics industry which generally has low levels of profits.

In the remainder of this conclusion, the benefits of the model, the applicability and other applications are reviewed.

5.2. Benefits of the model

Shown in the following map of Argentina is the location of the **CDC** as defined **by** the optimal solution and the area supplied **by** a carrier.

If the results are compared between the optimal and the real network, it can be seen that the optimal network has more CDC's. One possible explanation is that, in general, managers are risk averse. The set up of a **CDC** makes cost less sensitive to variations in demand: labor, rent, and all other costs need to be paid if there is a slowdown in sales. On the other hand, distribution costs are more flexible. If there is a drop in the demand, distribution can be easily reorganized and the costs will be reduced. Many managers would decide to pay more for

distribution in good times and be able to show to their superiors cost reductions when times are not so good.

When comparing both solutions in the total cost, as stated before, the use of the model can provide savings of **5%** and probably more if further improvements are also implemented. These savings can be generated if the model is used at the time of the design of the network. **If** the change is analyzed once the network is in place, other extra costs need to be included in the analysis, costs associated with closing a **CDC** (labor, lease, etc.).

If there is a plan to reengineer the network, the model still can be applied, but extra constraints would need to be added. For example, if the closure of a **CDC** demands an investment, this value can be transformed into a monthly cost (for example, as the value of a monthly payment if the company gets a loan for that amount) that can be added to the model if that **CDC** is not selected as part of the solution.

The model can be also used to analyze the effects of increasing demand. If, for example, the 3PL company gets a new Laboratory as a client and wants to see which changes would need to be made to the network in order to handle the new demand and how much the cost would increase.

From a socio economical point of view, these types of optimizations are very important in Argentina. Argentina is a country that in the beginnings of the $20th$ century was among the richest nations of the world. It is not the objective of this thesis to analyze the reasons, but it is a fact that now Argentina is in an extremely difficult condition. Argentina today has one of the largest national debts in the world, almost **60%** of the population lives in poverty and there is more than 20% unemployment. The situation keeps getting worse every year. This problem is more acute in the interior of the country. In the newspapers

there are almost daily reports of children from the provinces that die due to starvation and lack of medicine.

The improvement of a logistics network that can serve better the needs of the population and reduce prices can improve the quality of life of those who more need it. **Of** course, the most important problems of the country need to be solved in other areas, but the negative effects of social policies can be diminished **by** a more efficient logistics system. And **by** the time better social policies are in place, an efficient logistics network can be very useful in the implementation of those policies.

5.3. Other applications and further research

These models are being used in many different fields. In logistics and transportation, not only for the location of Cross Docking Centers, but also for the location of regional warehouses, offices for customer service, etc. Generalizing, the models can be applied to any type of distribution where there are destination points that require a certain amount of resources from a supplier and a decision has to be made about the number, size and location of the suppliers. Other uses of the models also include Telecommunication networks.

Software companies use these models to provide standard and customized solutions for all these problems. These software products in general collect the data, and based on the parameters entered, provide a solution and sensitivity analysis capabilities. Most of these products are not transparent to the user, so the user needs to trust that the solution is a good solution without really knowing if it is the best solution or how close to the optimal it is. For the design of a distribution network, because it is a task that companies undertake every several years, there are consulting firms that provide the service of collecting and analyzing the data and solve the problem for the companies.

With regards to further research in the field, efforts should be taken for the interaction of the different aspects of a Cross Docking Network. Instead of solving the location of CDC's, allocation of customers, distribution from the **CDC** to customers and geometry of the **CDC** as separate problems, they could be solved together as one problem.

Appendices

Appendix 1. List of cities with coordinates, demand and frequency:

 $\label{eq:2.1} \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\right)^{2} \left(\frac{1}{\sqrt{2}}\right)^{2} \left(\$

 $-76-$

Appendix 2: List of the CDC Costs:

-78-

Biographical Reference

Fernando Dobrusky is an MIT Master of Engineering in Logistics Candidate. Prior to moving to Cambridge, Fernando worked as Business Analyst at the Information Technology Services Division of the United Nations in New York. He was involved in the re-engineering of the **UN** inventory, the creation of a global database for projects and the optimization of the scheduling and resource allocation process for conferences.

Before relocating in the States, Fernando worked as the Project Leader in Logistics Engineering in Tasa Logistica, Buenos Aires, where was responsible for different logistic projects of the company including the centralization of warehouse's operations in the North of Buenos Aires and the re-organization of the cross-docking network around the country. He also worked with Argenmag as a consultant, advising the **CEO** on several organization and quality matters.

Fernando holds a **M.S.** in Industrial Engineering from Columbia University, where he also was a **TA** for the course Supply Chain Management at Columbia Business School. In addition, he is an Industrial Engineer from the University of Buenos Aires, where he subsequently was an adjunct professor, teaching Thesis Project and Commercial Logistics until moving to United States.

Fernando has been awarded the Fulbright Fellowship, the Interamerican Development Bank scholarship, and the Barsa award. Given his academic performance at Columbia University, he also received an award from the Japanese Government, to visit Japan to meet politicians, business leaders and university professors.

Bibliographical References

Baldacci, R., Hadjiconstantinou, **E.,** Maniezzo, **V.,** Mingozzi, **A.** (2002) **"A** New Method for Solving Capacitated Location Problems Based on a Set Partitioning Approach" Computers Ops. Res., Vol. **29, pp** *365-386.*

Bartholdi, **J. J.,** Gue, K. R. **(1999)** "Reducing Labor in an LTL Cross-docking Terminal" February *25,* **1999,** in Press.

Croes, **G. A.** *(1958)* **"A** Method for Solving Traveling-Salesman Problems". *1958.* Operations Research, Volume **6,** Issue **6** (Nov. **-** Dec., **1958), pp. 791-812**

Drezner, **Z.,** Hamacher, H. W. (2002) **"** Facility Location, Applications and Theory". Springer-Verlag Berlin Heidelberg New York

Klincewicz, **J. G. (1990)** "Solving a Freight Transport Problem Using Facility Location Techniques" Operations Research, Volume **38,** Issue 1 (Jan. **-** Feb., **1990), pp. 99-109**

Laporte, **G.,** Gourdin, **E.,** Labbe, M. (2000) "The Uncapacitated Facility Location Problem with Client Matching". 2000. Operations Research, Volume 48, number *5* (Sept. **-** Oct. 2000), **pp. 0671-0685.**

Levin, **Y.,** Ben-Israel, **A.** (2002) **"A** Heuristic Method for Large-Scale Multi-Facility Location problems" Computers and Operations Research, Article in Press.

Marin, **A.,** Pelegrin, B. **(1997) "A** Branch-And-Bound Algorithm for the Transportation Problem with Location of **p** Transshipment Points" Computers Ops. Res., Vol. 24, No. **7, pp 659-678.**

Websites consulted:

CIA - The World Fact book 2002 **-** Argentina. http://www.cia.gov/cia/publications/factbook/geos/ar.html Universidad Arturo Prat from Chile. http://www.unap.cl/imagenes/inte/mapa-argentina-gig.gif

Hotel Guia. Argentina http://www.hotelguia.com/argentina.jpg

Georgia Institute of Technology http://www.isye.gatech.edu/~jjb/wh/book/crossdock/design.html

Microsoft Clipart http://dgl.microsoft.com/mgolen/home.asp