1. Abstract

The current and future state of the infrastructure represents a key factor to analyze and develop logistical projects. However, predicting the end of planned infrastructure projects is a challenge considering the uncertainty related with the development of this sort of projects in Colombia. In this study, a robust design of a cross-docking network is introduced to deal with the inherent uncertainty of the Colombian infrastructure projects. For this purpose, the current and projected Colombian infrastructure are analyzed using different criteria. Then a Heuristic Simulated Based Optimization model is developed to design the cross-docking network considering the most viable infrastructural scenarios. In addition, a simulation model is proposed to validate other important variables for decision making, such as the hierarchy of the selected locations, the total cost of the network, and the usage level of the different resources. Finally, through a financial and qualitative analysis, the viability of the proposed solution is evaluated looking for calculating the return of the investment of the projects, as well as the most suitable locations for building the hubs that will support the cross-docking network.

Firstly, this methodology starts by collecting different input data such as city clusters (group of similar cities located nearly), actual and predicted demand of parcel services, current and projected infrastructure information like lengths, construction time estimations, sizes and capabilities among other parameters necessary for this model. Considering the large number of cities in Colombia, a method to select the most important clusters of cities was used. The method proposed for this purpose is divided in three stages, first a Pareto analysis displays the cities that group the 80% of the Colombian population, then a Complete-linkage clustering is run multiple times to estimate the average number of clusters, and finally, the K-means clustering technique is used to define the clusters for the study, using the average number of clusters resulting from the Complete-linkage clustering technique Regarding the infrastructure, the data that causes more uncertainty is the new highways’ (known as 4G projects in Colombia) delivery times. To estimate such times, a probability distribution was linked to every project according with the three points estimations of the Program Evaluation and Review Technique (PERT). A review of the available data and the expert criteria is used
to forecast the execution time of each project, from a pessimistic, regular and optimistic perspective.

Secondly, the Heuristic Simulated Based Optimization was executed employing a genetic algorithm. To obtain more details of the solution behavior and validate the Heuristic Simulated Based Optimization results, a deeper simulation is conducted. In this simulation more, detailed costs are calculated considering aspects as different truck sizes and costs. Also, this simulation allows to calculate variables as total traveled distance, number and truck loads, man hours and processed demand in each HUB. This validation helped to visualize how the designed cross-docking network impacts the logistics of a parcel services company in Colombia. The results of the model indicate that the load consolidation generates relevant impacts at a financial and operational level in comparison with a real local company. Despite the uncertainty of the 4G projects, the cross-docking distribution model should be implemented to generate efficiencies in the number of travels between cities, and maximize the use of the installed truck capacity.

2. Justification

The rapid economic growth worldwide and recent developments, such as the increasing amount and diffusion of ICTs (information and communication technologies) have burgeoned new business models and revolutionary industries (Klaus Schwab, World Economic Forum, 2015). For instance, industries, consumers and even governments demand an efficient flow of products and services seeking a suitable level of competitiveness. Therefore, the development and implementation of new logistic strategies have a crucial role in the global market. Logistics performance in international and domestic trade is of utmost importance to the economic growth and competitiveness of countries, thus the logistics sector is now recognized as one of the core pillars of economic development (The World Bank, 2017).

The current state of the Colombian infrastructure represents a key factor to analyze and develop logistic projects. Over the last years, industrialists have pointed out that the lack of adequate infrastructure in Colombia is withholding the development of the country, being the freight transportation one of the most affected sectors (ANIF1, 2013). According to the WEF (World Economic Forum), Colombia is placed in the 97th position over 120 countries in relation with the km/thousands of inhabitants of land routes, and it has a 0.013 km/km2 density of paved roads, being one of the worst qualities of infrastructure in Latin America (Cardenas, Gaviria & Melendez, 2006). Multilateral entities such as the Inter-American Development Bank (BID) and the World Bank have recommended to raise the investment over 6% of the annual Gross Domestic Product (GDP) in the next decade, to overcome the backwardness, compared to other countries (ANIF, 2013). When analyzing the railway transport infrastructures in Colombia, it can be noted that the number of kilometers over thousands of inhabitants’ index is 0.07, evidencing the lack of development. Referring to

---

1 Agencia Nacional de Instituciones Financieras.
Facultad de Ingeniería

INGENIERÍA INDUSTRIAL

Trabajo de grado – Primer Semestre 2018

aerial infrastructure, the country has 546 public and private airports, where the average of the mobilized load over the last years was 631 thousand tons per year (Cardenas et al, 2006).

Additionally, Colombia is in the 94th place of The Logistics Performance Index (The World Bank, 2016), which means that the country needs to improve in different fields of logistics such as infrastructure planning, service provision, and transport facilitation. Because of this critical situation, the government has committed the execution of important projects to improve the transportation infrastructure to be able to accomplish different FTA (Free trade agreements) that Colombia has signed with different countries, such as the USA and the European Union. As a result, there are various changes in the transport composition. The government has reported that there will be investments up to 3% of the annual GDP to face such backwardness, when according to the ANIF. In average, the previous investments were 1% of the annual GDP (ANIF, 2013).

In numbers, the government will generate contributions of 2900 million (USD) in railway, fluvial, aerial and maritime transport, in addition to the 4333 million (USD) for connecting the regions, which join to other projects, add up to 37333 million(USD) in investments by 2020 (ANIF, 2014). However, because of the national reality, these projects have a high coefficient of uncertainty. To illustrate this, based on the Colombian index of public works, only thirteen out twenty-five 4G projects approved, are susceptible to be accomplished (ANIF, 2013). Other issues are related to the establishment of public-private associations, problems with the acquisition of lands, engineering and construction problems due to internal security, and juridical, environmental, or other social problems. Therefore, the distribution of goods coming from different regions of Colombia, at the minor cost and keeping the service promise (the agreed time) is certainly a challenge for Colombian logistic companies because of the difficult Colombian geography and the use of a developing infrastructure.

3. **Theoretical background**

To satisfy the requirements of clients and get the highest profits in the third-party logistic business, it is necessary to find a solution to the problem of locating distribution centers and routing for connecting clients with those centers. These problems have been one of the most important and often discussed economic problems by researchers in logistics all over the world. They have proposed different kinds of methods to solve these problems (Richards, 1962). The main purpose of these methods is to minimize the cost of locating facilities and transporting goods. Already in 1905, Alfred Webber, a German economist published his *Theory of the Selection of the Location of Industries*, which wanted to cut the cost of transportation based on the weight and the distance of travel. He defined the problem as follows:

---

2 4G represents the fourth generation of concessions that the government has given to contractors to build transportation infrastructure.
“The question of the location of industries is a part of the general problem of the local distribution of economic activities. In each economic organization and in each stage of technical and economic evolution there must be a "somewhere" as well as a "somehow" of production, distribution, and consumption." (Webber, 1909)

Even though the definition above is a good introduction to the location problem, it does not have any real applicability because its oversimplification of the reality (Reid, 1968). Afterwards, with the arrival of operations research, new quantitative tools came up to solve this problem. The most frequently used model of linear programming rests on a set of nodes (that represent locations that demand goods and are a possible location for the distribution centers), and archways (that represent connections among the nodes). In addition, these models have parameters as the cost of localization and maintaining a location and the cost of sending a unit of product through an archway conforming a PL Network Model (Taha, 2012).

From this basic abstraction came all the different models, for instance, M. Daskin (1995) presented a model without restriction of the capability but including constraints to guarantee the demand satisfaction. A. Klose and A. Drexl (2005) included restrictions for evaluating if the fulfillment of the demand does not exceed the DC’s (Distribution center) capability. In addition, S. Syam (1997) considered a constraint of a maximum number of locations. The previous models are just a few among several amounts of variations of the basic model, which want to depict the reality to achieve better results.

In accordance with the problem, it may be useful to consider aspects of qualitative and quantitative nature, just as multiple linear optimization models do in the resolution of the MDCL (Multiple distribution center location) problem. For instance, the fuzzy distribution center location model (Yang, Ji, Gao & Li. 2006) and the multi-criteria decision making based on Fuzzy theory model proposed a method for decision-making based on linear integer programming (Wang, Xiong & Jiang. 2014). They had basic quantitative parameters, considering external and internal qualitative parameters, which are quantified by the implementation of fuzzy models, such as the fuzzy p-media models (Canós, Ivorra & Liern. 1999). Furthermore, among the set of the MDCL solution models, it is possible to find the optimization models with probabilistic parameters. Zhuge, Yu, Zhen and Wang (2016) proposed a dual-stage stochastic optimization model, identifying the location of the distribution centers in the first stage, and modeling the detailed distribution plan at the probabilistic behavior of the demand in the second stage. The stochastic linear optimization models tend to control the uncertainty of the parameters by establishing scenarios (Santoso, Ahmed, Goetschalckx, & Shapiro, 2005; Klibi & Martel, 2012; Zhen & Wang, 2015). Drezner and Scott (2013) proposed a linear programming model that integrates the location of the distribution centers, with the inventory analysis and probabilistic demand, found on the newsboy model problem (Drezner & Scott. 2013). Similarly, a proposed approach for the MDCL problem resolution was achieved by multi-objective linear optimization, where a multi-objective mixed integer linear programming (MOMILP) is used, weighting two objective functions that evaluate different aspects of the problem, as proposed by Sopha, Sri, Pradana, Guwan, Karuniawati (2016). Another approach for this type of problem is the dual-
stage model, which optimizes the location of the distribution center, using different information in a two-stage algorithm (Cupic, Teodorovic. 2013).

However, given the limitations of the LP (Linear Programming) to solve problems with combinatorial explosion, other methods have been developed to guarantee practical solutions to the DCLP (distribution center location problem) considering the different aspects to be evaluated in the decision making which depend on the context of interest. Therefore, multiple heuristics have been proposed for the resolution of MDCL type of problems, with variations in the parameters or implementation according to the type of problem studied. In practice, the authors base their models in the center of mass heuristic (Dobrusky. 2003). To illustrate this, the heuristic method for large-scale multi-facility location problems proposes an algorithm that identifies the optimal location point and reassigns customers to the logistic network by using the nearest neighbor heuristic (Dobrusky. 2003). Hua, Ju, and Yuan (2016) proposed the use of particle swarm optimization (PSO) heuristic, in the MDCM context, using non-linear parameters to calculate the fitness function (Hua, Ju & Yuan. 2016). Zak and Weglinski proposed the use of a logistic center location heuristic, in an established geographical context, based on the implementation of the ELECTRE III/V process, where the decision process is done by a binary association, evaluating sets of variables and information (Zak & Weglinski. 2014).

These solutions are suitable for the MDCL problem; however, this project aims to use a cross-docking logistics strategy, which uses these models and it is applied by many companies in different industries. The basic idea behind cross-docking is to transfer incoming shipments directly to outgoing vehicles without storing them in between (Jan Van Belle, 2012). There is a vast literature dealing with optimization for models and cross-docking (Dobrusky, 2003), thus it could be analyzed from different perspectives and using numerous strategies, which are based on the type of problem they are dealing with. For example, Anthony Ross (2008) described two heuristics that generate global feasible, near-optimal distribution system design and utilization strategies utilizing simulated annealing, and the problem is characterized by multiple product families and multiple cross-docking and distribution center sites. Later, Rami Musa (2010) approached the transportation problem of cross-docking network applying a nonlinear programming method, a novel ant colony optimization algorithm that meant to minimize the transportation cost in a network by loading trucks in the supplier locations and then route them either directly to the customers or to cross-docking facilities. Also, another strategy was used by S. Meysam Mousavi, (2014), who introduced two novel deterministic mixed-integer linear programming models that are integrated for the location of cross-docking centers and proposed a hybrid fuzzy stochastic programming solution approach attempting to incorporate two kinds of uncertainties into mathematical programming models. Recently, (Asefeh Hasani Goodarzi, 2016) developed a mixed-integer nonlinear programming formulation called biogeography-based algorithm for the problem in which the cross-docks location is determined without considering simultaneously the transportation strategy.
According to the aforementioned models, the resolution of the cross-docking problems focuses mainly on the vehicle routing aspects. Besides, the existing models are developed over downstream supply chains and do not apply for parcel services structures (characterized by multi-directional logistic network), and the existing models do not consider the uncertainty of the infrastructure development (parcel service). Therefore, taking into account the importance of the logistic industry, the global markets and the current and projected infrastructure in Colombia, this project looks for answering the following research questions, *how can the parcel flow be improved in Colombia? Where must be located the cross-docking distribution centers to improve the business model of a parcel company in Colombia? How could affect the different possible scenarios of infrastructure to the given cross-docking network in Colombia?*

4. **Problem statement and objectives**

The main objective of this project is *to propose a methodology for the robust design of a cross-docking network for a parcel service company in Colombia using optimization models and taking infrastructure uncertainty into account.* This main objective will be achieved by accomplishing the following specific objectives:

- Identify the different criteria that affect a cross-docking network in Colombia, particularly referred to the infrastructure uncertainty
- Determine the possible scenarios\(^3\) that affect the design of a robust cross-docking network in Colombia.
- Design and develop a robust cross-docking network taking infrastructure uncertainty into account.
- Validate the proposed approach on the basis of the case study and the scenarios defined using a simulation model.

4.1. **Design requirements:**

- The selected model considers both sorts of variables: quantitative and qualitative.
- The selected model takes into account the uncertainty of the different infrastructure scenarios.
- The selected model can adapt the number of possible locations.
- The selected model considers the level of importance of the location in the cross-docking network.
- The simulation environment compares the performance of the proposed solution of different scenarios with the one of the current situation.
- The simulation environment generates a performance indicator for the solution of each scenario.

---

\(^3\) In this project scenarios are the different possible combinations of criteria and parameters such as a given layout of transport infrastructure in certain moment in Colombia.
• The solution of scenarios determines an approximately location for Cross Docking centers location (Hubs).

   The resultant model enables to change the number of possible locations by setting a budget. Also, the model determines the order of opening based on the importance and saves generated for each hub. The simulation tool generates performance indicators as average used capacity and average transported unit cost that allows to make a benchmark between different the proposed solution and other logistic networks. Finally, though a financial and qualitative analysis of the proposed locations, the most qualified cities are chosen as cross-docking hubs. Considering the final characteristics of the proposed model that were explained before, we can ensure the design requirements were accomplished.

4.2. Design constraints:

• For the construction of solutions and simulation of scenarios, only available data of parcel services operations in Colombia was considered.
• For the construction of solutions and simulation of scenarios, available data of current and projected Colombian transportation infrastructure was considered, as available.
• Routing decisions were not considered.
• The proposed approach considered the total demand (orders) in units of mass (kg) from cluster to cluster. However, the specific characteristics of each order such as size and weight were not considered.
• The financial analysis was calculated based on the saves generated by the reduction of fixed costs related to truck’s number and capacity
• The distance between cities used in the data prep phase were estimated by a regression analysis
• The demand between the cross-docking network nodes is estimated by a regression analysis
• The simulation of the HSBO are made with less parameters and variables than the validation results simulation due to the algorithm performance.

   A regression analysis was made to estimate the road distance between cities of the study due to the complexity of collect all these data. Also, the demand was estimated from information of the Colombian Postal Sector Study (2015) using its data to determine the number of parcels that are sent from a cluster to another. About the developing and current infrastructure data, it was collected from the National Infrastructure Agency (Colombia).

4.3. Norms and standards

   The methodology that was used in this project is described by the ISO 13053-1 for the quantitative methods in process improvement. This methodology typically comprises five

5. Methodology

As was mentioned above, this methodology comprises five phases: Define, Measure, Analyze, Improve and Control. Each is stage is explained below.

a. Define

The parcel services industry faces a particularly challenge in Colombia. The uncertainty related to the infrastructure projects obstructs the proper long-term planning associated with the construction of strategic facilities that allow a more efficient future operation. This project aims to develop a methodology capable of design a cross-docking network, dealing with the uncertainty of the infrastructure projects in Colombia and seeking an optimal performance. During this phase, the problem was clearly defined, particularly for the parcel service in Colombia. Also, during this phase the authors looked for the sources of information, the data required and the requirements for the final solution. The final solution must identify those locations where a cross-docking hub should be built, considering the available budget, and the proper time to invest in the planning horizon.

b. Measure

Given the context of the problem, it was necessary to assess the current state of the infrastructure projects in Colombia that could affect the operation of a parcel services company. Therefore, the main mode of freight transport was identified. Road transportation represents 73% of the total freight transport in Colombia (Mintrasporte, 2016), for this reason only road projects were considering. 4G projects are the main road projects, they will cover over 7000 km and require an investment of more than 16666 million USD (ANI, 2017). 4G projects are currently in two phases pre-construction and construction. Although, there are other road projects in Colombia, they were not considered because of their limited scope and lack of information.

Once the infrastructure projects to be evaluated were identified, the cities or towns that were considered in the model were established as well, given that there are approximately 1122 towns in Colombia. To simplify the analysis and provide a more accurate model, a Pareto analysis was proposed to determine those cities and towns that have the most significant population and represent the 80% of the total amount of inhabitants of the country. In addition, a clustering process was developed to define groups of cities that are geographically close. The clustering methodology proposed for this problem is described in two general stages, complete link hierarchical clustering (stage one) and k-means clustering (stage two).

The First stage of the clustering uses a complete link hierarchical clustering methodology. This methodology uses as input the geographical location of all the Pareto
cities identified previously. Initially each city is in a cluster of its own. The cities are then sequentially combined by randomly selecting an origin city \((W)\) and a destination city \((T)\), and then evaluating the distances between points, if the distance between the origin city and the destination city is greater than a defined parameter of maximum distance, the cities are separated in two different clusters; if the distances between the selected cities is lesser than the parameter of maximum distance, both cities are grouped in the same cluster. This process is executed until all the cities are group into clusters.

In order to define de final number of clusters, the first stage is executed in \(N\) iterations, and using the number of final number of clusters, an average is calculated \((K)\).

The second stage of the clustering methodology uses the \(K\) parameter calculated in the first stage, and groups the cities into \(K\) clusters, according to de maximum distance parameter. The maximum distance parameter is defined by using the promise of service of the parcel company.

The previous clustering analysis facilitated the demand, costs and other parameters estimation because the bigger the town or city the easier to get more accurate data available. The demand between clusters is calculated based on a model to estimate an origin-destination matrix for demand of passengers in urban transportation (Amézquita, Duran, & Fajardo, 2016).The parameters of the model were adapted, so that public data could be used. The used parameters were: Total parcel services load transported in 2015 and its growth rate, population and PIB. The total parcel services load transported in 2015 was 88.5 million of packages/year with an annually growth rate of 1,7%. (Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia, 2016). The transportation cost used were the ones from a local parcel services company in Colombia.

c. Analyze

Given that the final decision is whether to open a hub facility in a specific location, the parameters defined above and the design constraints, an optimization model arose as the most convenience way to deal and solve the problem. A linear programming model (LPM) was proposed as a first approach to the defined and analyzed the problem. The developed optimization linear model is described in the following:

Sets
Set \(\{1,2, \ldots, n\}\) : Set of clusters in the model.

Variables
Var \(X\{i \in I\}\): Binary variable that indicates if the cluster \(i\) is a cross-docking hub.
Var \(Y\{i \in I, j \in I\}\): Binary variable that indicates if the cluster \(i\) is part of the \(j\) Hub.
Var \(R\{i \in I, j \in I, k \in I, m \in I\}\): Binary variable that indicates that the demand between the clusters \(i - m\) is satisfied by going through the clusters \(j - k\).
Parameters

Demand \{i \in I, j \in I\} : Parameter that indicates the demand between clusters \(i - j\).

Cost \{i \in I, j \in I\} : Parameter that indicates the cost of transporting a unit of demand between the clusters \(i - j\).

Objective Function

Minimize the sum of the cost of the distribution of the parcels from \(i\) to \(j\), by adding the multiplication of the cost \(c\) of transportation with the demand, by consolidating the load at the load at the hubs in \(r\).

\[
\sum_{i,j,k,m \in I} \text{Demand}[i,j] \times (\text{Cost}[i,j] \times r[i,j,m,k])
\] (1)

Constrains

Constrain that indicates that each city \(i\) can only be part of one cross-docking center \(j\):

R1: \(\sum_{j \in I} Y[i,j] = 1\quad \forall i \in I\) (2)

Constrain that indicates that the cities \(i\) can only be part of a hub \(j\):

R2: \(Y[i,j] \leq x[j]\quad \forall i, j \in I\) (3)

Constrain that indicates the demand that must go through the cross-docking centers:

R3: \(Y[i,k] + Y[j,m] \leq 2 + r[i,k,jm]\quad \forall i, j, k, m \in I\) (4)

Constrain that restricts the number of cross-docking centers in the network:

R4: \(\sum_{i \in I} x[i] \leq n\) (5)

Results

The following four clusters were obtained:

- A1: Madrid, Cundinamarca
- A2: Puerto Tejada, Cauca
- A3: Envigado, Antioquia
- A4: Sitionuevo, Magdalena.

After executing this LPM, it was evident that using this kind of model, the uncertainty related to the 4G project could not be considered and even though integer programming optimization approaches are applied to solver small hub problems, larger instances of hub location problems need to be solved by heuristic procedures or meta – heuristics procedures (2013, R.Z. Farahani, M. Hekmatfar, A.B. Arabani, E.Nikbakhsh). Therefore, it was
necessary to build a heuristic model capable of considering and processing all the relevant parameters that were defined, based on the linear programming model.

d. Improve

As a result of the analysis described above, the following Heuristic Simulation Based Optimization (HSBO) was proposed, based on a Genetic Algorithms (GA). GA are one of the most popular heuristic algorithms that represents a powerful and robust approach for developing heuristic for complex and large – scale combinatorial optimization problems (2005, C.B Cunha, M.R Silva). Therefore, the general structure of a genetic algorithm was implemented with some particular change. The steps involved in the HSBO proposed to solve the hub location problem considering infrastructure uncertainty are as follows:

The initial population was randomly generated considering the number of hubs facilities that could be built with the available budget to invest. Each solution was generated so that every cluster could become a hub. In the chromosome, if a cluster is a hub, it was represented as “1” and “0” otherwise, in this last case the cluster must consolidate load in the nearest hub, as it is represented in the Figure 1 below. For instance, in that case, cluster 1 is a hub and cluster 7 is not a hub and must consolidate load in the hub 8.

![GA Chromosome Solution Codification](image)

<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
<th>Cluster 4</th>
<th>Cluster 5</th>
<th>Cluster 6</th>
<th>Cluster 7</th>
<th>Cluster 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1. GA- Solution representation.*

After the initial population was generated the model followed the process describes in Figure 2. Although the general structure of a genetic algorithm was implemented, a simulation technique was included in the fitness evaluation stage seeking a more reliable and robust solution, where a considerable amount of different possible scenarios was considered.
In the context of fitness evaluation stage, a Monte Carlo simulation was used, and two hundred of different scenarios were created. Each scenario was a possible combination of the 4G road projects that are ended simultaneously in a certain year. Those scenarios were generated applying the PERT method (Rodríguez, 2016), which use the beta probability distribution function. Each project had its own beta function with the following parameters (expected value and standard deviation):

\[ E = \frac{P + 4MI + O}{6} \]  \hspace{1cm} (6)

\[ DS = \frac{(O - P)}{6} \]  \hspace{1cm} (7)
Where, P was the most pessimistic case, M was the most likely case and O the most optimistic case. Those values were estimated based on expert concept. A random number was calculated for each 4G road project and then, it was compared with the accumulated beta probability distribution for the project in the selected year. If the random value was lower than the accumulated distribution then that project is finished in this scenario. As result, two hundred of vectors with length 26 (which is the number of 4G road projects) were created. Given that the mixed of the concluded project was different for each scenario, two hundred of distance and time matrix were calculate in every solution, considering that the building of new roads and their combination had a positive effect in travel times between to clusters.

In each solution evaluation, every cluster (that was not a hub) was assign to the nearest hub and the routes between hubs were set. The load consolidation, which is the main characteristic in a cross-docking network, was made through a matrix that added the load between clusters considering the load coming from other clusters that were associated to the same hub. Having the total load that will be transported from hub to hub, it was possible to calculate the number of trucks to be used. Two kinds of cost were estimated, fixed and variable costs. Fixed cost was composed of hub building costs and trucks purchase and operation. Variable costs were calculated based on the addition of the simulated scenarios which provide a value that considers two hundred of possible cases, given robustness to the model. The fitness is the result of dividing one by the addition of variable and fixed cost.

Figure 3. summarizes the key processes in the simulation-based optimization model. Once this model had been executed and the approximate location of the hub had been defined, it was necessary to indicate a more detailed position for the cross-docking center in the cluster.

**Genetic Algorithm - Fitness Evaluation**

![Diagram](image)

*Figure 3. Fitness Evaluation.*
Based on the intuitionistic fuzzy hierarchical group decision-making model for the cross-docking location selection by Mousavi & Vahdani (2016), and the ELECTRE cross-docking center location method by Uysal & Yavuz (2014), the following qualitative criteria was proposed to evaluate each town in the clusters (only those selected as a hub):

<table>
<thead>
<tr>
<th>Main Criteria</th>
<th>Sub-Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>Labor Cost</td>
</tr>
<tr>
<td></td>
<td>Land Cost</td>
</tr>
<tr>
<td></td>
<td>Transportation Cost</td>
</tr>
<tr>
<td></td>
<td>Financial Incentives</td>
</tr>
<tr>
<td>Markets</td>
<td>Travel time between cross-docking centers</td>
</tr>
<tr>
<td></td>
<td>Availability of utilities</td>
</tr>
<tr>
<td>Government Influence</td>
<td>The certainty and stability of laws</td>
</tr>
<tr>
<td></td>
<td>Government restrictions and policies</td>
</tr>
<tr>
<td></td>
<td>Security</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Quality and reliability of modes of transportation</td>
</tr>
<tr>
<td></td>
<td>Free zones</td>
</tr>
<tr>
<td></td>
<td>Telecommunication systems</td>
</tr>
<tr>
<td>Labor resource</td>
<td>Skill level</td>
</tr>
<tr>
<td></td>
<td>Availability of labor</td>
</tr>
</tbody>
</table>

Table 1. Qualitative Criteria.

The weighting factor of each main criterion was defined by an assessment from experts and the literature review of:

<table>
<thead>
<tr>
<th>Main Criteria</th>
<th>Factor Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>0.267</td>
</tr>
<tr>
<td>Markets</td>
<td>0.2</td>
</tr>
<tr>
<td>Government Influence</td>
<td>0.178</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>0.289</td>
</tr>
<tr>
<td>Labor resource</td>
<td>0.067</td>
</tr>
</tbody>
</table>

Table 2. Weighted qualitative criteria.

The experts where selected from the National Planning Direction of Colombia, the decision-making process was explained to each of them. Each expert evaluated each city, by
assigning a qualification for each main criteria. The selected hubs where validated with the director of infrastructure of a local parcel company.

For the selection of the location of the cross-docking center in the selected clusters, each town was evaluated by experts. The decision of the best location for the cross-docking center was selected from the towns with the best weighted score of the criteria by the expert’s assessment.

e. Control

Once the genetic algorithm was executed, the result obtained from the developed model was validated throughout the comparison with two cases: case with no load consolidation (Case 1) and case of a local parcel services company in Colombia (Case 2). The case one does not have locations where the load is consolidated, instead of this, all the cluster send load directly between them. The case two uses the hub location of a local parcel services company of Colombia. These locations are Bogotá, Cali y Barranquilla (TCC, 2018). Another simulation model was proposed for that purpose. This time, a thousand scenarios were generated and the variable cost was more specific because of the use of a different size of trucks. In the same way, the simulation provides data such as: total traveled distance, required workforce time, number, size and occupation of the trucks and the processed load in each hub, always satisfying the demand. Those parameters worked as a performance indicator and are compare with the case 1 and case 2 mentioned above.

Finally, the financial capacity of the company was considered. Considering the budget available to invest at the same time to stablish the selected cross-docking centers, therefore an opening construction order was defined.

The ordering methodology used a simulation-based optimization approach to define the opening order of the cross-docking centers. This approach is separated in two stages: the first stage is to define the order of opening of the centers, and the second stages indicates the time between each opening.

The first stages use as parameter, the results of the previous methodology of selection. Knowing the location of the cross-docking centers; a horizon of planning (h) is defined as parameter and divided in equal periods of time (t) according to the number of cross-docking centers selected (N), where \( t = \frac{h}{N} \).
In each $T_0 + t \times (n)$, where $n$ is the number of cross docking centers that are open, a simulation is made, where the cost of the network is calculated by simulating the opening of $n$ cross-docking centers. For each simulation, the cross-docking center with the lowest cost of the network is selected. For $n > 1$, the simulation is made by calculating the cost of the network, if the previous cross-docking centers are already opened and selects the greatest aggregated saving. This methodology evaluates the cost savings by opening cross-docking centers and sorts the opening from the greatest saving in cost to the lowest aggregate saving in the horizon of planning ($h$).

The second stage of the ordering defines the time between openings of each cross-docking center. For this stage, the defined order from the previous stage is used. Starting at year $T_0$, an analysis of the network cost is completed by comparing the cost of the network with $n$ cross docking centers open, with the cost of the network with $n-1$ cross docking centers open, starting at $n=1$. The decision of the time of opening the center is made by evaluating the cost of the initial investment for the opening and its projected payback at a commercial rate, and evaluating the projected savings of the network cost, with the cross-docking center open. Assuming that no cross-docking centers are open simultaneously, the time maximum amount of years that the company should wait, is the calculated payback time for each hub. This evaluation was made for each cross-docking center, in a separate way.

6. Results

In this section, results are reported in the same order that each step was presented in the methodology.

- Current state of the 4G road projects in Colombia

Most of the relevant information about the 4G road projects in Colombia was consulted from the ANI (Agency Nacional de Infrastructure), the key information of each project extracted from ANI’s available and public resources was: total length, geographical location,
transportation cost savings, travel time savings, agreed construction time (starting and finishing dates). Appendix 2 summarizes the available information for each 4G project in Colombia

- Main cities

As a result of the Pareto analysis, 250 cities were identified as the most populated in the country. Figure 5 is a visual representation of the executed pareto analysis. The 250 cities are shown in the Appendix 3.

![Figure 5. Pareto Analysis Representation. Taken from google maps.](image)

- Clusters.

After applying the two-stage clustering methodology proposed, 32 clusters of cities were calculated. The graphical representation of the clusters obtained is the following:
The 32 clusters and the cities which belonging to each cluster are exposed in the Appendix 4.

- **Parameters and criteria.**

  The available data and input criteria defined as relevant to consider in the Heuristic Simulated Based Optimization model are presented in the Appendix 5.

- **Scenarios**

  As it was mentioned above, in each scenario a different combination of completed projects is evaluated, to do this, all projects must have a beta probability function with different parameters in a certain year. Appendix 6 shows the beta probability function for each project and its parameters evaluated in the year 2025 because, in average, this is the most likely year when all the projects should be finish. The hundred evaluated scenarios are shown as well. There are some results examples below:
Figure 7.1. Project 1 2025

Figure 7.2. Project 2 2025

Figure 7.3. Project 3 2025

Figure 7. Examples of the Beta Distribution for the 4G, evaluated at the year 2025.

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Random value</strong></td>
<td><strong>Condition</strong></td>
<td><strong>Random value</strong></td>
</tr>
<tr>
<td><strong>Project 1</strong></td>
<td>79% Operative</td>
<td>90% Under construction</td>
</tr>
<tr>
<td><strong>Project 2</strong></td>
<td>52% Under construction</td>
<td>88% Under construction</td>
</tr>
<tr>
<td><strong>Project 3</strong></td>
<td>57% Operative</td>
<td>14% Operative</td>
</tr>
</tbody>
</table>

Table 3. Example of the simulation of the condition of the 4G Project at the year 2025.

- Robust cross-docking network.

Considering the scenarios created, the Heuristic Simulated Based Optimization is run. As a result, the following clusters are selected as a possible hub location.

<table>
<thead>
<tr>
<th>Cluster centroid</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Vega, Cundinamarca, Colombia</td>
</tr>
<tr>
<td>Medellín, Antioquia, Colombia</td>
</tr>
<tr>
<td>Flandes, Tolima, Colombia</td>
</tr>
<tr>
<td>Sincelejo, Sucre, Colombia</td>
</tr>
<tr>
<td>Chaparral, Tolima, Colombia</td>
</tr>
</tbody>
</table>

Table 4. Final hubs locations.
• Qualitative Analysis

The results for the qualitative analysis indicates that the ideal location for the cross-docking centers are towns with low land cost, as well as high offer in labor resource. Based on the assessment by the experts, the selected locations of the cross-docking centers have a high quality and reliability of roads, and in all the cases are related to the 4G projects that are currently in development. All the selected zones are feasible when evaluated by the security criteria.
This stage of the decision-making methodology is based on the present state of the evaluated criteria, and its expected behavior according to the experts’ knowledge.

- **Simulation model – Performance indicators**

  The result obtained from the developed methodology are compared with two cases mentioned above: case with no load consolidation (Case 1) and case of a local parcel services company in Colombia (Case 2).

  The implementation of the cross-docking network proposed generates a 92% decrease in the fixed cost versus the case one. The solution has a 2% over cost vs the case 2. However, our solution uses five hubs, it means two hubs more than the case 2, which explains the increase in the fixed cost. The amount of truck by size to be used are presented in the following table:

<table>
<thead>
<tr>
<th>Truck capacity / Case</th>
<th>Obtained solution</th>
<th>Local Company</th>
<th>No load consolidation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max. 4.5 Tons</td>
<td>6</td>
<td>6</td>
<td>959</td>
</tr>
<tr>
<td>Max. 8.5 Tons</td>
<td>19</td>
<td>13</td>
<td>21</td>
</tr>
<tr>
<td>Max. 17 Tons</td>
<td>14</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Max. 20 Tons</td>
<td>7</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Max. 22 Tons</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Max. 32 Tons</td>
<td>14</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Max. 35 Tons</td>
<td>23</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>Daily Fixed Cost (USD)</td>
<td>$ 8.898</td>
<td>$ 8.737</td>
<td>$ 103.176</td>
</tr>
</tbody>
</table>

Table 6. Results of the simulation based optimization heuristic- Used Truck Capacity.

The total traveled distance and the workforce time decrease 94% versus case 1 and 12% versus case 2. The variable cost is presented in the next table:
Table 7. Results of the simulation based optimization heuristic.

The transportation unit cost and the truck’s capacity is improved versus both cases, as it is shown in the graph below.

- Financial validation of the solution proposed.

The financial analysis is completed by considering the results of the Fix and Variable Costs of each situation evaluated. In this case situation represents the number of hubs open at the same time.
By evaluating the payback of the inversion of design of the cross-docking centers for each situation, the performance of the network is measured and its impact at a financial level is reviewed. Comparing the payback of the initial inversion that a cross-docking center demands for its construction using the average Bank of the Republic interest rate, with the payback of the savings against the previous state of the network for each situation, the financial viability of opening each cross-docking center is measured. The calculation of the payback for each situation is made by following the order that the optimization model recommends, and that no multiple cross-docking centers are open simultaneously.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Payback (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situation 2</td>
<td>0.5</td>
</tr>
<tr>
<td>Situation 3</td>
<td>3.0</td>
</tr>
<tr>
<td>Situation 4</td>
<td>3.6</td>
</tr>
<tr>
<td>Situation 5</td>
<td>4.4</td>
</tr>
<tr>
<td>Situation 6</td>
<td>8.1</td>
</tr>
</tbody>
</table>

The payback analysis indicates that opening the first cross-docking center has a high impact on the cost of the network as the present state does not consolidate packages and uses an inefficient model of distribution. Comparing the payback of each situation with the payback of the bank of the republic interest rate indicates that the migration from the network to a cross-docking model can be considered as an optimal investment from a financial criterion, as well it also indicates the periods of time that the company should wait between the construction of each cross-docking center to postpone the inversion and evaluate the state of the demand and 4G infrastructure with more certainty.


The presented results were obtained in a computer with an Intel (R) Core (TM) i5-3210M CPU @ 2.50GHz, 2501 Mhz, 2 Core (s), 4 Logical Processor (s). The deterministic model was executed with the current information of the road infrastructure in Colombia (not including the 4G projects) and 19 clusters, reducing the complexity of the problem in the Gusek software, obtaining results in 2.3 hours. While the HSBO model obtained results after 2 hours, using the data of both current and in development infrastructure, considering the following parameters:

a. Mutation rate: 40%

b. Initial population size: 100

c. Number of iterations without improvement: 1000

d. Number of scenarios: 200
These values were determined thanks to a sensitivity analysis starting from initial values of (a:5%, b: 30, c:100, d:50) and increasing the values of the parameters until there is not improvement in the answers vs the execution time. The executed sensitivity analyses are shown as follows:

**Figure 10.** Sensitivity analysis mutation rate percentage.

**Figure 11.** Sensitivity analysis mutation rate percentage.

**Figure 12.** Sensitivity analysis mutation rate percentage.
The most sensitive parameters was the number of iterations without improvement, where the diminution of the network cost was 17% once the parameters was steady. Moreover, the comparison between the HSBO model and the LP model showed that there was a decrease of 21% in the network cost using the HSBO model.

8. Conclusions and recommendations

In this study, a robust design of a cross-docking network was presented to deal with the inherent uncertainty of the Colombian infrastructure projects. For this purpose, the current and projected Colombian infrastructure were analyzed using different criteria. The proposed methodology was developed so that the identified criteria and relevant parameters can be adapted. Although the scope of this problem was focus on the methodology, it is recommended to focus in a deeper and more accurate acquisition of the input data, in future investigations. The more accurate the input data, the more applicable and executable the final solution is. This constraint about the use of historical data about infrastructure projects did not allow an improvement of the proposed analysis, using more sophisticated statistical techniques, however the simulation technique introduced in the genetic algorithm process allowed to achieve a robust solution dealing with the infrastructure uncertainty in Colombia. Also, the proposed methodology delivered a solution that guarantee a good performance in any final infrastructure case and considering the desired planning horizon.

The simulation technique allowed us to validate the proposed approach, comparing it with the current network used by local parcel company, finding better performance indicators. Although the problem was related to a parcel services company, the model can be adapted to other industries in future studies. Furthermore, this methodology can be reapplied in any transportation project, in any country and on a larger scale.

Future work can address the problem of vehicle routing, this can achieve better indicators and larger savings by reducing the number of used trucks, however this problem was not the focus of this study.
9. References


