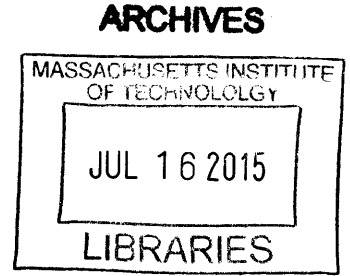An analytical model to increase air volumes and minimize the Net Achieved Rate
in air freight transportation
by

Per Niklas Blomberg
BSc in Economics - Lund University, 2014
BSc in Political Science - Uppsala University, 2010

Ramon Gras Alomà
MSc & BSc in Civil Engineering - UPC-BarcelonaTech, 2010

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 2015

© Niklas Blomberg, Ramon Gras. All rights reserved.

Signature of author:....   **Signature redacted**
Master of Engineering in Logistics Program, Engineering Systems Division
May 8, 2015

Signature of author:....   **Signature redacted**
Master of Engineering in Logistics Program, Engineering Systems Division
May 8, 2015

Certified by:........   **Signature redacted**
Dr. Jarrod Goentzel
Director, MIT Humanitarian Response Lab
Thesis Supervisor

Accepted by:..........................   **Signature redacted**
Dr. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor, Engineering Systems Division
Professor, Civil and Environmental Engineering

1

An analytical model to increase air volumes and minimize the Net Achieved Rate
in air freight transportation

by

Niklas Blomberg

and

Ramon Gras

Submitted to the Engineering Systems Division on May 8 2015
in Partial Fulfillment of the Requirements for
the Degree of
Master of Engineering in Logistics

Abstract

Airfreight forwarding companies must develop accurate forecasting tools to analyze the suitability and attractiveness of each incoming bidding process, to decide whether to participate in a tender and how to define the optimal commercial strategy. The "1:6" weight/volume ratio establishes that whenever the cargo ratio is different from 1:6 ($1m^3$:167 kg) forwarders must pay the higher rate: either volume or weight. Given the main constraints in terms of volumetric capacity and maximum weight per volumetric unit, the most profitable business opportunities consist of combining in the same load compatible products with different densities. The main target is to come close as possible to the desired 1:6 ratio, to minimize the average price per load. To remain competitive, airfreight forwarding firms must improve their consolidation techniques, to combine in the same load cargo with compatible densities. The availability of robust analytical resources will allow airfreight industry companies to improve their rate of success, in terms of enhancing both efficiency (by increasing air volumes and densities) and profitability (by minimizing the Net Achieved Rate). This thesis develops an analytical model based on meaningful metrics to provide airfreight forwarders with an accurate and solid forecasting tool to select the bids under consideration that best match with their current portfolio in terms of air volume usage for a given origin-destination lane. It also predicts breakeven rates to increase profitability by minimizing the Net Achieved Rate. Furthermore, the model provides a series of metrics and visualization tools to help air freight forwarding companies to improve their understanding of their current portfolio for a given origin-destination lane and define their commercial strategy with respect to air freight cargo tenders.

Thesis Supervisor: Dr. Jarrod Goentzel

Title: Director, MIT Humanitarian Response Lab

Acknowledgements

I would like to thank our contact persons at our sponsor company for their readiness to provide time and effort into the work with this project.

-Niklas Blomberg

I would like to thank my family for all their support and serving as an inspiration through their exemplary qualities.

-Ramon Gras

We would like to thank our advisor, Dr. Jarrod Goentzel, for stimulating us to unleash a creative approach when addressing the key moments of this thesis.

Table of Contents

5

List of Figures

List of Tables

1.- Introduction

The global air freight industry reached a value of $125.5 billion in 2015, after growing by 1.2% in 2014. According to market researchers, after experiencing some decline in 2012 and 2013, the global air freight industry is expected to slowly increase its value over the next five years. Analysts predict the performance of the sector to accelerate in the coming years, as they estimate a Compound Annual Growth Rate (CAGR) of 2.1% for the 2014-19 period. The higher business volumes occur in the international transportation field, accounting for 166,110 freight ton kilometers FTK in 2014, which represents 73.5% of the overall industry, as opposed to the domestic market, which accounts for the remaining 26.5% in the industry (Marketline Industry Profile, 2015).

A study by Qin, Zhang, Qi and Lim (2014) emphasizes that in today's global free market, third-party logistics providers (abbreviated 3PLs) are becoming increasingly important in the air freight industry. A third-party logistics provider is a firm that provides its customers with outsourced (or "third party") logistics services for part, or all of their supply chain management functions. Air freight forwarders are 3PL companies that serve as agents for the manufacturing exporters or shippers moving cargo to an overseas destination. These agents handle the import rules and regulations of foreign countries, the export regulations of governments, the methods of shipping, and the documents related to foreign trade. Export freight forwarders are licensed by the International Air Transport Association (IATA) to handle air freight. Freight forwarders assist shippers in preparing price quotations by advising on freight costs, port charges, consular fees, costs of special documentation, insurance costs, and their handling fees. Air freight forwarders recommend the packing methods that will protect the merchandise during transit or can arrange to have the merchandise packed at the port or containerized.

One of the main challenges that air freight forwarder companies face is to design strategies to optimize the efficiency of services, especially with respect to the combination of different products on the same aircraft covering the same route, a phenomenon known as consolidation. The main constraints affecting the freight transportation by air are the volumetric capacity and the maximum density allowed, as well as limitations for distribution strategies in the allocation of cargo along the aircraft.

In the supply chain industry, most cargo carriers do not have a direct relationship with the shippers, the end consumer. Cargo carriers usually rely on freight forwarders, who own the customer

relationship with the shippers and frequently dictate which contract is acceptable, after assessing the risks and potential benefits of the trade lane. Freight forwarders are engaged in the business of assembling, collecting, consolidating, shipping and distributing freight by aircraft. Air freight forwarders serve a dual role: to the shipper, they are indirect carriers which receive freight from various shippers under one tariff, usually consolidating the goods into a larger unit, which is then tendered to an airline; to the airlines, they are shippers.

This thesis addresses the question of how to deal with the consolidation problem in the air freight forwarding industry. In particular, this study focuses on developing an analytical model to rationalize decision-making criteria for the commercial strategy of an air freight forwarder, which we will refer as our sponsor company. The model aims to best address the bidding processes for buying and reselling air cargo space in order to improve air volume efficiency and maximize profitability of the loads, by minimizing the Net Achieved Rate for each origin-destination lane's consolidation process. To do so, we analyze two main datasets:

  o Historical dataset of 9 customers/shippers from different industries (high tech, fashion, consumer products, etc.) with respect to tenders, regardless of whether the sponsor company was awarded or not, in which case we were provided with the series of target rates that the customer submitted to the sponsor company for each round of the bidding.
  o A series of Requests for Information (RFI) for 9 customers/shippers, presenting business opportunities for different origin-destination pairs.

1.1.- Motivation

Air freight forwarder companies serve a dual role. To the manufacturing company (or shipper), air freight forwarders are indirect carriers, since they receive freight from various shippers under one tariff or rate, usually consolidating the goods into a larger unit, which is then tendered to an airline (or carrier). To the airline, air freight forwarders are shippers. Therefore, the main challenge that air freight forwarders face when trying to maximize profits is being forced to buy air cargo space in advance to the airlines/carriers, before the bidding processes have been held, without having certainty about the profitability of the loads. This uncertainty results in numerous cases of freight forwarders participating in bidding processes that are not beneficial and may result in negative Return on Investment (ROI) ratios. For example, transporting light goods purchased in

advance, if not combined in the same plane with other heavier products with higher added value, can lead to losses to the company.

This is what we can call the "consolidator's bidding dilemma": by following an aggressive commercial strategy based on proposing relatively low rates to the shipper, the freight forwarder boosts the possibility of being awarded the load. But this aggressive policy may endanger its profitability, and eventually cause losses to the company. Moreover, not being able to combine products with compatible densities and volumetric characteristics may cause the air freight forwarder to lose business opportunities, since some products that may appear not to be profitable per se, are likely to be profitable after a consolidation process when combined with other loads that present compatible features such as density or total volume.

Therefore, to be successful, air freight forwarders need to develop and implement analytical methods to understand the nature of their current portfolio for each focus origin-destination lane. Being able to perform analysis for each particular origin-destination lane will allow freight forwarders to design visualization tools that provide a better understanding of their current situation lane by lane, so that they can select the most attractive incoming bids for each route and propose competitive rates to present the most profitable and efficient solution after consolidation.

1.2.- Scope of Research

The fundamental question of our sponsor company is:

"How can air freight forwarders use an extensive database with historical data from past tenders involving awarded bids and target rates to select the strategic bids that best combine with current portfolio in terms of air cargo space usage and to define the rates to minimize the Net Achieved Rate and increase overall profitability of loads for each origin-destination lane?"

For this purpose, the sponsor company has at its disposal a vast amount of data from past tenders, both awarded and target rates for each round of the bidding processes. In addition, the sponsor company provided us with a series of datasets presenting incoming bids under consideration, including routes, gross weight and volumetric data. The key challenge was to extract value from these datasets by developing a methodology to both build a database that integrates all

the key variables and to design an analytical model capable of solving the consolidation problem and providing benchmarks for proposed rates, metrics and visualization tools to guide decision-making for the commercial strategy.

The scope of the project was intentionally narrowed by the sponsor company to focus on developing a framework to guide the commercial strategy, rather than embracing the extra challenges that derive from considering operational constraints such as variability in size of airplanes, pallet size, integration of intermodal networks, delays, etc.

1.3.- The air freight industry

The air freight industry has three main agents: shippers, carriers, and third party logistics (3PLs) firms. Shippers have products or goods that need to be transported from an origin to a destination and delivered to a consignee; carriers (airlines) are the companies in charge of providing equipment and aircrafts to transport goods for hire; shippers outsource their transportation management to 3PL companies such as air freight forwarders, which assume the role of the shipper in shipper-carrier relationships. As a result, air freight forwarders serve a dual role as both indirect carriers and shippers. This thesis assumes that the rates/tariffs that air freight forwarders pay to carriers as a deterministic constraint, and focuses on analyzing the rates/tariffs that freight forwarders charge to shipper companies.

Global tender management process

There are three kinds of businesses in air freight transportation: cargo-only transportation operated by air freight forwarders, cargo-only integrators such as FedEx and UPS, and combination carriers ("belly cargo"). Ever since a series of deregulation processes erupted in the 1980s, the industry witnessed an extraordinary growth of other cargo-only carriers, with major increases occurring during the early 1990s and the early 2000s, whereas the unit costs of FedEx and UPS remained relatively stable. All other all-cargo carriers achieved significant reductions in unit costs. Since 1990, their average unit costs have nearly halved, decreasing from around 50 cents per ATM to nearly 25 cents per ATM (Donatelli 2012).

In recent years, there has been an increasing interest in increasing profit margins by optimizing air volumes and density efficiency such as focusing on business opportunities in combination carriers / belly cargo in passenger flights. According to a study developed by Boeing, freighter aircraft currently carry a little over half of freight traffic (around 56% by freight tonne kilometres) and this percentage is expected to remain fairly steady through to 2033. Although freighters carry slightly more than half of the cargo traffic, they account for well under half of cargo capacity (CAPA - Centre for Aviation, Boeing, IATA, 2014). The possibility of combining passenger baggage and freight from shippers significantly increases overall profitability for each trip, offering interesting opportunities for consolidation.

Our study focuses on increasing efficiency with respect to the relationship between air freight forwarders and shippers. The client or shipper typically first releases and RFI (Request for Information) in order to gage prospective suppliers' global network and capabilities, financial stability, technical capabilities, etc. Air freight forwarders (and/or carriers) then compile responses and submit them to the client. After analyzing the different proposals, clients/shippers release an RFQ (Request for Quotation) or "tender". Suppliers then receive the documents, analyze the requirements and submit any clarifying questions to the client.

After sharing basic information regarding the conditions of the shipment, suppliers (air freight forwarders) pre-fill certain fixed components using historical rate databases (or commercially approved local charge databases), and the lead pricing team circulates the internal pricing template to global pricing teams. Lead pricing team receives back pricing information, and consolidates all responses into an internal pricing master file. The results are commercially analyzed, strategically adjusted and submitted to the client. Clients analyze responses, and make a decision on suppliers to award cargo, or to invite to subsequent bidding rounds – typically requesting pricing reductions round-over-round.

Clients launch subsequent rounds of the RFQ, often providing feedback on rate levels as compared to the competition (anonymously). Final negotiations are often done in a face-to-face meeting with a client, where final pricing decisions can be made. Having the ability to consolidate loads from different tenders for each origin-destination lane offers the possibility of charging twice for the same air space, what can potentially boost overall profitability.

1.4.- Sponsor company profile: a global air freight forwarder

Our sponsor company is a global 3PL enterprise who serves as an air freight forwarder, and aims to improve their understanding of the consolidation problem. To do so, this thesis develops an analytical model that allows them to rationalize decision-making when approaching tender processes, in order to increase air volume usage and maximize profitability of loads. To this end, the sponsor company provided us with a series of data from their past bids, whether awarded or not, for 9 customers from different industries, as well as illustrative examples of RFIs from different clients.

One of the main challenges air freight forwarders face is dealing with technical/operational limitations with respect to physical characteristics of cargo that affect air freight transportation. In general, we can assert that extremely bulky or low-density products are not suitable for air transport, either because they cannot be accommodated on an aircraft or the resulting cost per pound would be uneconomically high (Airports Council International - North America, 2013).

Although the reasons for a shipper to transport a good by air are complex, most industry professionals and members of the academic community agree upon a number of factors that are believed to have the broadest and most significant influence on the "air-eligibility" of specific commodities. Air freight accounts for about 1.5 percent of total freight by weight transported worldwide, but some 30 percent of total freight value. The types of goods that have a greater propensity to travel by air rather than much cheaper surface modes of transport can be summarized as those which either have a high value-per-pound; or the products which have a specific reason, such as the perishability of the product, for them to be transported with greater urgency than the speed allowed for maritime freight transportation.

The all-cargo airline industry is constituted of a small group of airline companies that focus solely on moving cargo by air, rather than carrying passengers. Air transport is one of the five modes of transportation (highway, rail, maritime, and pipeline being the others) used for the global movement of goods. While all modes but pipeline can transport the same commodities, maritime and air transport are the only two modes able to support intercontinental freight movement. Maritime transport offers low-cost movement of goods, whereas airfreight offers the benefit of speed, reliability, and security. Changes in world air cargo traffic are strongly linked with changes in the world gross domestic product (GDP); therefore, as the world economy expands, so does the demand for air transport. According to the Boeing 2012 annual forecast study, the worldwide

demand for the movement of air cargo grew annually at an average rate of 7.1% between 1987 and 1997; however, this growth slowed after September 11, 2001, to an annual growth rate of 4.1%. After the terrorist attacks that happened in the U.S. on September 11, 2001, the price of fuel experienced a significant increase, boosting the cost of air shipments, an effect that ultimately caused companies to migrate toward less expensive road, rail, and maritime transport (Boeing, 2012).

Uncertainty in revenue management planning

With respect to the passenger airline industry, which is quite stable, revenue management planning in the air cargo business is much more difficult to predict. Cargo airlines suffer from a lack of detailed historical booking data, while passenger airlines have at their disposal a vast amount of historical booking data which they use to estimate demand and determine tariffs and prices on various routes. As a result of the meaningful differences between the supply and demand of passenger and air cargo airlines, revenue management for cargo is more complex than for the passenger business.

When addressing predictability regarding capacity supply, the main issues that arise are: heterogeneous production platforms, large number of routing possibilities, considerable number of restrictions and uncertainty with respect to available capacity, multi-dimensional capacity, and multi-segment flights. On the other side, market demand issues include stowage loss, unequal trade lanes, short booking periods, volatile business, continuous show-up rates, market structure, and data shortcomings. For cargo shipments, the amount of cargo is volatile and uncertain until departure time; that is, the weight or volume may fluctuate, taking up more or less room on the aircraft, unlike passenger airlines in which a seat is a seat. Another issue is market structure. Air cargo airlines typically only provide capacity to a limited number of customers, such as freight forwarders, who make most of the bookings. Therefore, the loss of one booking can have a large impact on the revenue for that flight.

As previously mentioned, air freight can be a viable option for high value, high density and low volume cargo. However, Air freight is only cost efficient up to a certain point and not all high value and high density products are viable to load on an airplane. Whenever the combination of

value added, density and volume cargo is not enough favorable, the only chance for assuring profitability of loads is consolidation.

In particular, two characteristics of the air freight sector make improvements in consolidation processes a strategic asset of the first order:

- o The existence of an upper bound in density, the 1:6 weight/volume ratio
- o The overwhelming dominance of relatively light cargo (lighter than 1:6) with respect to denser cargo

Being able to master and excel in the art of consolidation would allow air freight forwarders to approach tender processes with an accurate estimation of the potential level of consolidation that can be applied to a given load. In this way, they could design a more aggressive strategy of tariff/rate setting that will allow them to increase their current portfolio while increasing the global level of benefits. These techniques are especially relevant with respect to relatively light cargo that is not high value added, such as some perishables (fruits, flowers) as opposed to textiles or high tech consumer goods.

1.5.- The consolidation system: combination of cargo to minimize the Chargeable Weight

Since the Black-McKellar Air Mail Act, which re-introduced competitive bidding as a means for commercial airlines to procure air freight contracts, was approved in June, 1934, the emphasis in air transportation has been on strict competition between carriers in the establishment of routes and rates with provision for awards to the lowest bidders (Harding, 1940). The main problem that airfreight forwarders face when trying to maximize profits is being forced to buy air cargo space in advance, before the bidding processes have been held, without having total certainty about the profitability of the loads. This uncertainty results in numerous cases of freight forwarders participating in bidding processes that are not beneficial and may result in negative Return on Investment (ROI) ratios. For example, transporting light goods purchased in advance, if not combined in the same plane with other heavier products with higher added value, can lead to losses to the company.

Airfreight forwarding companies face two main challenges to increasing both revenue and profits:

1.- Determining whether or not participate in a tender process and at which rates, and

2.- Combining in the same lot cargo loads composed of different products that result in an optimal average weight/volume ratio of 1:6.

The next paragraphs will present two essential concepts in air freight transportation: the Chargeable Weight, and the consolidation of cargo. Both concepts will be critical in determining the consolidation potential of a bid under consideration with the current business or portfolio for a given origin-destination lane, as well as the influence of such bid in the Net Achieved Rate for the whole lane. A posteriori we will review in Section 2: Literature Review the state of the art studies regarding the different approaches that have been designed to face the consolidation problem to increase air volume usage to minimize the average price per load and maximize profits for a certain origin-destination lane.

Chargeable Weight: Two main limitations affect airfreight transportation: volumetric capacity and maximum weight allowed per volumetric unit. In order to ensure the viability of the flight while guaranteeing appropriate safety conditions, a maximum charge density per unit volume is set. The "1:6" weight/volume ratio establishes that whenever the cargo ratio is different from "1m3:167 Kg" forwarders must pay the Chargeable Weight, which is the highest rate: either for the actual gross weight or the space necessary (Volumetric Weight) for the transport of the consignment.

Consolidation of cargo: Since products show a significant variety of densities and volumes, and do not adjust to the 1:6 ratio, the most profitable business opportunities consist of combining in the same load compatible products with different weight in order to come close as possible to the desired ratio, to minimize the average price per load. This process enables airfreight forwarders to maximize the occupied space and charge for the transportation of two products at the same time while paying just one trip.

1.6.- Contributions

The main contributions achieved by this thesis are:

o Built the database: structure the architecture of the database to make it as simple, complete and usable as possible; identifying gaps in tracking information, suggesting software procedures to improve data traceability, and implement an easy, scalable and repeatable process to update the database including the incoming bids.

o Developed metrics to analyze and depict current portfolio / current business for any given origin – destination lane with respect to the level of consolidation and profitability.

o Designed an analytical process to describe the most attractive business opportunities and define benchmarks to establish rates to increase profitability while being competitive.

o Constructed a series of visualization tools that describe the results of the analytical model and depict the main characteristics of the incoming bids under consideration to provide senior management with accessible, synthesized information to face decision-making in the commercial strategy of the air freight forwarder.

We analyzed two years' worth of data provided by the sponsor company, and based our approach upon the data we were given, taking it as a representative depiction of the market.

The methodology we used to build the analytical model is based on integrating three levels of analytical procedures:

1. The database contained in a spreadsheet updatable in real time that integrates datasets from past tenders, whether awarded or not, and data from incoming bids / RFQs.

2. The analytical model embedded in SQL server that calculates a series of metrics for each origin – destination lane based upon the spreadsheet. By refreshing the link between the two at any given moment the analytical model calculates the proper metrics in real time.

3. The visualization tools designed with Tableau, that depict the main metrics calculated using the analytical model embedded in SQL server through a series of interactive graphs.

1.7.- Thesis Outline

This thesis aims to provide air freight forwarder companies with an analytical tool to identify the most attractive business opportunities and propose a series of metrics to address the tender processes held by shippers from the different industries by focusing on consolidation potential and profitability.

The methodology used to build database is based upon integrating a series of spreadsheets presenting datasets from past tenders and incoming bids with the analytical model embedded in SQL server that calculates a series of metrics that address the consolidation potential for an incoming bid with respect to the current portfolio for a given origin-destination lane that includes all customers operating in such lane. Such metrics are depicted using visualization tools that we created with Tableau.

Those procedures contribute to standardize data gathering and add value in decision-making rationale criteria with respect to the commercial strategy of air freight forwarders, in order to guide high level strategic decisions as well as bidding processes when selling air cargo space to shippers. By having a better understanding of both their current business and the business opportunities that arise, managers from air freight forwarders and logistics professionals can rationalize decision when addressing air freight transportation problems with a rigorous approach: better understanding will lead to better decisions. The resulting analytical model brings air freight forwarders a number of metrics illustrated with visualization tools that guide:

o Selection of the most attractive bids under consideration, the ones that best match the current portfolio.
o Air volume usage and density optimization.
o Estimation of breakeven and suggested rates for each bid under consideration after consolidation with the current portfolio for each one of the 9 origin-destination lanes (see Figure 1.7.1) to maximize profitability by minimizing the Net Achieved Rate.
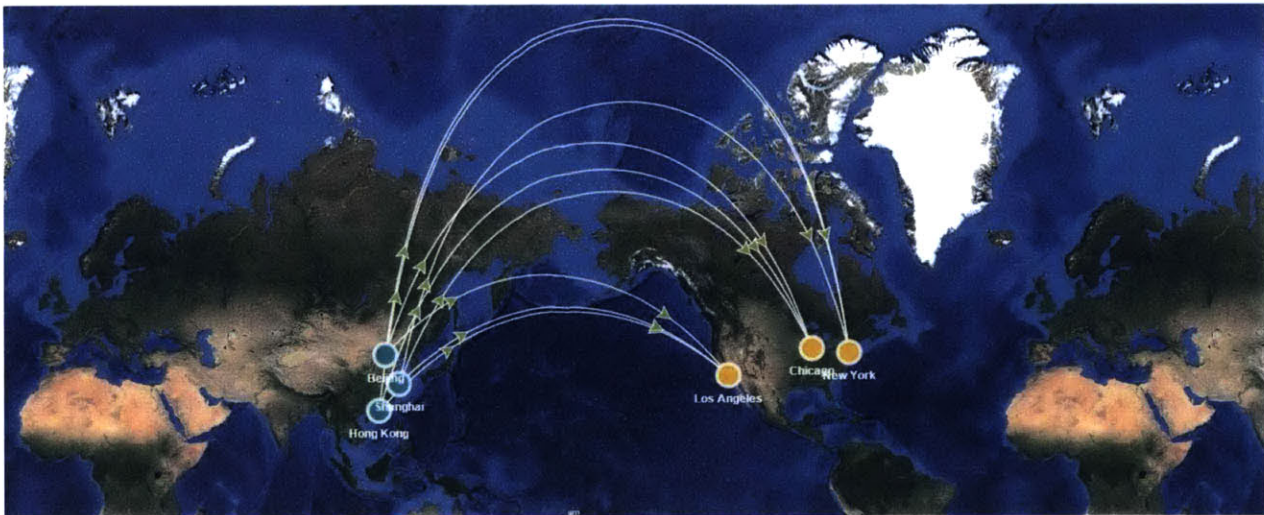
Figure 1.7.1. The 9 origin-destination focus routes considered in the study

19

2.- Literature Review

2.1.- Introduction

According to Vinod and Narayan (2008), air freight accounted in 2008 for US$55bn annually, which represents 12% of the total revenue of the airline industry. Despite this significant turnover, several factors such as higher fuel prices, increased competition and excessive regulation contribute to the downward pressure on profitability and declining yields. For these reasons, the cargo industry is facing decreasing margins. Lack of both transparency in the tendering system and focus on premium products contributes to this phenomenon. Unlike the passenger airline business, where future published fares of an airline are transparent to competitors, future cargo rates of competitors are not available in the airfreight industry. Thus, an analytical tool able to estimate the rates accurately is an essential tool for companies active in the sector.

The authors of the study assert that "rate consistency requires a consistent and repeatable process, either with a set of simple tools or advanced decision support that can duplicate the rate creation process along with all the assumptions and overrides that were applied when either a published rate sheet or a customer-specific rate sheet is generated" (Vinod and Narayan, 2008, p317). This thesis aims to design a robust, consistent and repeatable process based on an analytical model of the consolidation procedures and composed of selected metrics and a set of visualization tools that help guide the commercial strategy of the company.

Given the main constraints in terms of volumetric capacity and maximum weight per volumetric unit, the most profitable business opportunities consist of combining in the same load compatible products with different densities in order to come close as possible to the desired 1:6 ratio, so that they minimize the average price per load. In this thesis, we offer a review of the research done on airfreight commercial strategies regarding the consolidation technique. In the next section we discuss the theoretical and heuristic models that address the different consolidation approaches. In successive sections we discuss the influence of the consolidation approach on the profitability of the loads, and we present the emerging research area of aligning the consolidation approach with both the operational level and the strategic level. Our main contribution will be to contribute to rationalize decision-making regarding the commercial strategy of an air freight forwarder company, by developing a set of procedures to:

1.- Build a database that involves all the relevant information for past tenders, current portfolio and bids under consideration for each origin-destination lane

2.- Define metrics depict current portfolio / current business for any given origin – destination lane

3.- Design an analytical model to select which of the incoming bids best combines with the current portfolio.

4.- The model will as well depict through visualization tools the space management efficiency of the consolidation, and will provide an estimation of the profitability for a series of rates for a given rate under consideration.

2.2.- Heuristic models to approach consolidation

A study by Huang and Chi (2006) establishes the basic conceptual framework for further research on using Lagrangian relaxation-based heuristics as the backbone to solve the consolidation problem. By using a heuristic model as opposed to an exact-solution algorithm the authors greatly alleviate the computation load and provide a suitable solution for problems with practical size. However, the model showed some flaws when addressing large-scale problems, such as failing to define a tight lower bound that would strengthen the solution algorithm, or refining the parameter setting system. In addition, the authors did not develop in detail a systematic analysis of capacity constraints.

Several research projects have explored this methodology to refine the design of the constraints of the mathematical problem, whether related to the variability of the size of aircraft, discounts, limitations relating to the consolidation process, or the number of flights available for a certain period of time. Huang and Chang (2009) develop a solution algorithm for the air cargo revenue problem, based on approximating the expected revenue function in the dynamic programming model while considering the stochastic volume and weight of shipments. Their major contribution is to empirically check that whenever the volume is the bottleneck of the capacity, the Joint Approximation Heuristic (JHC) control significantly out-performs the hypothetico-deductive (HD) model control.

The main contribution of Wong, Leung; and Hui (2009) is to formulate a shipment planning model that includes realistic constraints, develops a solution methodology that utilizes the

21

characteristics of the shipment planning environment, and examines the managerial implications of the model constraints on the overall shipment cost. Despite their advances in defining a mixed 0-1 model, they do not delve into the consequences of establishing long-term strategic alliances.

Not all studies have followed a heuristic approach. Li, Bookbinderand and Ehedhli (2012) explore the airfreight forwarder consolidation problem, and devise two solution methodologies to address large problem sizes. The first is based on Lagrangian relaxation, where the problem is decomposed into a set of knapsack problems, and a set of network flow problems. The second is a local branching heuristic that combines branching ideas and local search. Their main contribution to the field opened by Huang and Chi is that, in both cases, the model incorporates the effects of economies of scale, by addressing the effects of quantity discounts. By showing consistent experimental results in providing robust heuristics, such a methodology opens a promising path to follow.

Cargo loading at an operational level: One of the limitations that was not addressed in depth until recently is the operability of consolidation processes with respect to the size of cargo containers. Qin, Zhang, Qi and Lim (2012) design a solution for the freight allocation and consolidation problem that involves shipping and loading items into containers from different sizes. Their model outperformed CPLEX software when solving medium and large size intances.

For a better understanding of operational constraints, it is useful to read the research led by Xiang, Ding-you and Ying-gui in 2014. The authors address the optimization of high density and concentrated-weight freights loading, taking into account the mechanical constraints such as implementing an even distribution of the freight's weight and unconcentrated loading on the floor of the automobile or aircraft.

2.3.- Profitability of loads

One of the most relevant decisions that the air freight forwarders face during the tender process is whether to pursue a business opportunity based on its estimated profitability after consolidation. Han, Tang, and Huang (2010) develop a Markov model for single-leg air cargo revenue management under a bid-price policy. Their analysis focuses on estimating whether to accept or reject the booking. Their main constraint is to accept only cargo when revenue exceeds the

opportunity cost, that they calculate (Figure 2.3.1) based on bid prices. Their proposed solutions are derived by maximizing a reward function of a Markov chain.
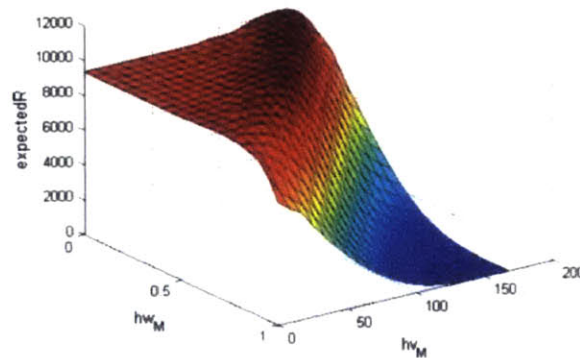
Figure 2.3.1.- Surfaces of expected revenue with respect to bid prices. Source: Han, Tang & Huang (2010)

Another variable influencing decision making in the design of the commercial strategy of an air freight forwarder is capacity management. Xiao and Yang (2010) refine a mathematical formulation to model the inventory control problem with two-dimensional capacities. Unlike existing heuristic approaches, they derive an analytical form of the optimal solution for general two-dimensional Revenue Management problems. Their main innovation is to introduce elements of marginal analysis for the different load combinations and their respective rates. When assessing the overall profitability for a given origin-destination lane, the cost structure is another critical factor to bear in mind. Onghena, Meersman, and Van de Voorde (2014) evaluate the cost structure of the integrated air freight business by means of a translog cost function. Their results suggest that integrators exhibit strong scale and density economies in the short and the long term, and conclude that the concentration in the integrated air freight industry will continue.

2.4.- Planning: Route design

To assess the influence of consolidation on both flight planning and route design, Chan, Chow, So and Chan's study (2012) focuses on multi-agent-based framework to facilitate process automation for the air cargo industry. Their research project revolves around enhancing two labor-intensive flight planning processes, namely cargo consolidation and equalization, and integrating them into the planning strategy. Deepening on route design and intermodal transportation alternatives Patrick Burnson asks to what extent the consolidation affect the

23

evolution of the airfreight industry, noting the trend towards concentration (Burnson, 2013) that benefits large industry players, to the detriment of companies with less footprint in the market. Malchow and Kadafani develop a discrete distribution of maritime route assignment that can be useful when addressing airfreight project efficiency (Malchow and Kadafani, 2004).

2.5.- Strategic approach: leveraging economies of scale

Strategic decision making can be supported in various ways. Burnson emphasizes the effects of leveraging economies of scale when considering mergers and acquisitions in the airfreight industry (Burnson, 2012), while Bowen and Leinbach (2003) focus on the role that large global forwarders have played in bringing advanced logistics services in the South East Asia region, as well as on the impact of the so-called economies of scale and scope.

2.6.- Conclusions

The constraints that affect the air freight industry now force relatively low profit margins. The fact that airfreight forwarders must purchase air cargo space before bidding processes occur, requires freight forwarding companies to develop a rigorous methodology to carefully design systematic cargo consolidation processes, to improve process efficiency (increase air volume) and ensure profits to the greatest possible number of tenders.

Though several useful models have been developed, multiple important elements have not yet been adequately addressed. The most relevant research gaps:

- o Integration of models to create a comprehensive model that takes into account several variables, from physical properties of the cargo, to the inherent profitability of the load, an estimation of the optimal rate for a bid under consideration, and including marginal economic analysis regarding revenue, market share, marginal costs and profitability.
- o The consequences of long-term alliances.
- o Lack of focus on premium products.
- o The variability of the bidding process, which deserves statistical analysis.

This thesis addresses the first problem and proposes an analytical model to describe the main metrics that define the current portfolio regarding volume usage efficiency and profitability. In addition, it provides an integrated solution to select the bids under consideration that best combine with the current portfolio, and visualization tools to depict an estimation of overall profitability for any given rate applicable to the incoming bids.

In summary, with recent advances in developing the theoretical framework and the basis for heuristic models to improve the consolidation analysis and design in the airfreight industry, firms have an opportunity for significant progress in their understanding of the commercial strategy. We propose the use of advanced tools and a comprehensive heuristic model able to define a consolidation strategy that will enable airfreight forwarders to have better resources ahead of the decision-making process, which will result in increased air volumes and better Net Achieved Rates, entailing a significant increase in profits.

3.- Objective

The main objective of this thesis is to design an analytical model that provides air freight forwarders with a series of visualization tools to guide decision-making when addressing the tender processes of buying and reselling air cargo space from carriers to shippers. The model aims to assess the attractiveness of incoming bids with respect to their current portfolio or current business for each origin-destination lane, by depicting air freight consolidation opportunities and suggesting a series of references for rates to face the Requests for Quotation processes. In this context, the most attractive business opportunities will be defined by their potential level of consolidation and their ability to improve air volume usage and density efficiency with respect to the 1:6 weight/volume ratio, i.e., by minimizing the Net Achieved Rate.

3.1.- Create a framework to build a dynamic database

- o Select the critical variables and units that affect potential consolidation and profitability
- o Identify the meaningful relations between the critical variables that will define potential consolidation and profitability
- o Define a framework to build the dynamic database that integrates
- o Establish a structure to develop a repeatable, scalable, robust process to gather data from RFQs in order to guarantee traceability of inputs and business opportunities, whether awarded or not
- o Facilitate the integration of the datasets in a database management software system that integrates both the datasets with information from incoming bids and the metric designed to assess attractiveness of business opportunities

3.2.- Design metrics to estimate consolidation potential, profitability and breakeven rates

- o Design an analytical model embedded in a database management software system
- o Populate the database with the datasets built with the spreadsheet
- o Design a series of critical metrics to depict attractiveness of business opportunities from different customers for each origin-destination lane
- o Depict the current portfolio in real time

o Use statistical analysis to analyze cargo properties and create simulation tools to populate the database in case the data provided by the sponsor company is incomplete and limits the potential of the analytical model

o Establish the basis for further research on marginal analysis with respect to consolidation potential and profitability

3.3.- Design visualization tools to depict the potential and attractiveness of bids under consideration/current bids

o Design visualization tools with specialized software to depict both the current portfolio and consolidation and profitability potential of incoming bids

o Integrate the visualization tools with the analytical model embedded in a database management software

o Provide the visualization tools with interactive applications to focus on certain aspects of the analysis when assessing attractiveness of bids under consideration

4.- Methodology

This section explores how we interpreted and used the datasets in order to:

1.- Define a series of metrics composed of useful parameters to describe the consolidation potential and profitability of a particular lane, and design a series of visualization tools that will illustrate and guide decision-making regarding the commercial strategy of an air freight forwarder, by providing the senior managers with meaningful graphs that depict both the current portfolio and the consolidation potential and profitability for each bid under consideration after consolidation.

2.- Build an analytical model using SQL server software that integrates both the database that will be periodically updated with data from incoming bids by the professionals of the air freight forwarder, and the simulation tool that is embedded in the model based on the metrics formerly defined and depicted using the visualization tools created with Tableau.

We applied the methodology developed to the 9 target lanes we are focusing on: the routes between three Chinese cities (origin) and three American cities (destination), as seen in Table 4.0.1.

Table 4.0.1. Origin and destination cities under study

Origin (China)		Destination (United States of America)	
City	IATA Airport Code	City	IATA Airport Code
Beijing	PEK	New York	JFK
Hong Kong	HKG	Chicago	ORD
Shanghai	PVG	Los Angeles	LAX

The sponsor company brought us two main datasets: limited physical information regarding a series of recent Requests for Information from nine customers (RFI01 to RFI09), and Air Rates datasets for nine customers (C1 to C9) with extensive information about bids, whether awarded or not, from the last two years.

The only two RFI customers that had useful information for this thesis were:

1.- RFI01, retail/lifestyle mostly footwear, and

2.- RFI03, retail/lifestyle, sports and outdoor

28

Since the purpose of this thesis is to design a framework to implement an analytical model, the sponsor company provided us with Air Rates information regarding only 9 customers, selected from different industries in order to reproduce on a small scale a representative sample of the air freight transportation industry. The industries represented by the 9 Air Rates customer datasets are listed in Table 4.0.2.

Table 4.0.2. Air Rates customers and industries

Air Rates Customer ID	Industry
C1	Lifestyle/clothing selling both in retail stores and non-branded stores
C2	Retail - clothing
C3	Industrial
C4	Industrial
C5	Technology - hardware
C6	Retail - clothing
C7	Technology - hardware
C8	Toy industry
C9	Lifestyle/clothing selling both in retail stores and non-branded stores

We used the Air Rates spreadsheet as the reference dataset, considering the awarded bids as the current portfolio/current business, and the non-awarded bids as potential bids under consideration/incoming bids/current bids. The underlying assumption is based upon the fact that, as the sponsor company informed us, most customers repeat similar tender RFQ on a yearly basis, so that non awarded bids from past years can presumably be expected to be released recurrently in successive years.

The core objective was to select the most meaningful inputs from the sponsor company Air Rates datasets, identify any relevant flaws or incomplete columns, come up with reasonable assumptions to populate any significant flaws if needed by using data from RFI customers as a reference, construct the database using an intuitive and clear interface that allows for repeatable and scalable processes and design the metrics and the visualization tools that describe not only the current portfolio but also the consolidation potential and profitability for each bid under consideration, for any given origin-destination lane.

The main limitations we found with the main dataset, the Air Rates, were:

1.- scarcity of data, especially regarding gross weight, and

2.- lack of data with respect to density of each load.

These two variables are strictly necessary to address the consolidation problem in air freight forwarding. Thus, in order for us to be able to show the potential and usability of the analytical model and the visualization tools, we decided to populate the datasets by using reasonable and realistic assumptions. The type of data that is recreated using simulation tools is information susceptible of being collected by an air freight forwarder company. Hence, gathering such information in a regular basis in practice should not suppose any obstacle. Since the purpose of this thesis is focused on developing an analytical model, the fact that some of the information we use do not strictly correspond to real data will not affect the rigor of the proposal.

In making the decision to complete the database with ex novo created values for two variables, Gross Weight and Density, we assume the consequences in terms of potential and limitations of the model. On the one hand, we lost precision regarding the internal correlations between certain variables, such as cargo density and profitability. On the other hand, we gain the ability to display the final result of our analysis by using visualization tools, as well as presenting the potential of the model and suggest lines of research for the future. If we had taken the decision not to complete the databases, in certain specific cases, this thesis would have suffered severe limitations on the ability to display the analytical model results using graphs and visualization tools. Moreover, since analyzing the internal correlations of certain variables is not the primary object of this thesis, we decided to proceed to populate the dataset using simulation tools that replicate statistically solvent and significant trends we identified on data from both the RFI and the historical tenders. The assumptions and criteria used to populate certain values of the database are explained in detail in the corresponding sections.

4.1.- Consolidation opportunities: methodology, nomenclature and assumptions

In this thesis, we were interested in evaluating the suitability of pursuing each bid under consideration with respect to the resulting combination with the current portfolio for a given origin-destination lane, and to determine a series of metrics to suggest references for proposing rates during the tender process to increase air volume and profitability per lane. These metrics are

intended to guide decision-making whenever an air freight forwarder is facing tender processes to sell air cargo space to shippers from different industries and consolidate the resulting cargo mix. To do so, we defined the nomenclature to be used consistently through the research project, as well as the constraints that define the boundaries of the consolidation problem we wanted to optimize and depict via visualization tools.

The essential variables that we need to design the analytical model are:

γ_w = gross weight, the total weight in kg for a certain bid

υ = volume, the total volume for a certain bid in CBM (cubic meters, m^3)

υ_w = volumetric weight, the total volume for a certain bid, normalized using the 1:6 WV Ratio

$\delta = \frac{\gamma_w}{\upsilon}$ density of cargo, as gross weight per volumetric unit, in $\frac{kg}{m^3}$

Origin = Origin Gateway for a certain bid

Destination = Destination Gateway for a certain bid

o-d = origin-destination pair, for each lane

Shipper = the company who ships cargo to the consignee. The shipper in the context of this thesis is the client of the air freight forwarder

Carrier = the airline, the legal entity that is in the business of transporting goods for hire

RFI = Request For Information, the initial announcement of a tender process held by a shipper/manufacturing company in order to gage prospective suppliers' global network and capabilities, financial stability, technical capabilities etc.

RFQ = Request For Quotation, the tender, a negotiating approach whereby the buyer asks for a price quotation from a potential seller/supplier from specific quantities of goods or transportation services that a buyer needs over a certain time and at a fixed price, and the buyer agrees to purchase such services exclusively from the seller during that time

NET/NET Rate = Gateway to gateway transportation service rate the carrier charges to the air freight forwarder

31

FSC = Fuel Surcharge rate the carrier charges to the air freight forwarder

SSC = Security Surcharge rate the carrier charges to the air freight forwarder

β_i rate the air freight forwarder charges to the shipper for a certain bid i

ϱ_i = NET/NET Rate + FSC + SSC, rate we pay to the carrier/airline for a certain bid i

τ_r = Target Rate for a given o-d lane, reference provided by the shipper

Current Business = current portfolio of awarded bids for a certain o-d pair

Current Bid = incoming bid under consideration for a given lane

NAR = Net Achieved Rate, total

DIM Factor = 1:X Weight/Volume Ratio for a certain bid, being $X = \dfrac{1000}{\frac{\gamma_w}{v}} = \dfrac{1000 \cdot v}{\gamma_w}$

ω_v = 1:6 Weight/Volume Ratio for air freight, equivalent to the maximum density (upper bound, threshold) permitted in a certain shipment on average: $167 \dfrac{kg}{m^3}$

According to the nomenclature described above, we also defined two significant variables that affect the consolidation processes:

Chargeable Weight = Max(Gross Weight, Volumetric Weight)

σ = chargeable weight for the shipper

α = chargeable weight for the carrier/airline

Net Achieved Rate = NAR $= \dfrac{\sum_{k=1}^{n} \alpha_i \cdot \varrho_i}{\sum_{k=1}^{n} \sigma_i}$

The NAR operates therefore as an estimation of the level of consolidation of a certain o-d lane, since it provides us with an approximation to the rate we pay to carrier after consolidation. The smaller the NAR, the better the level of consolidation. However, since we considered this metric to describe only part of the goodness of the consolidation potential, we defined other metrics such as the consolidation delta Δ to describe the level of potential consolidation for a current bid with the current business in further chapters.

The main assumptions we considered to define the boundaries and constraints of the consolidation problem are:

o The air cargo the air freight forwarders buy to the carriers/airlines was purchased in advance, and the capacity is always used at 100%. The sponsor company wanted to focus on improving their understanding of incoming business opportunities, and therefore we agreed that all issues related to carrier capacity management were out of scope.

o The chargeable rate to be paid to the carrier/airline can be defined as a constraint, since we estimate it as the sum of NET/NET Rate, FSC and SSC.

o The analytical model focuses on the Request for Quotation processes: what current bid is the most attractive for our current business to increase air the average density usage? What is the rate we should charge to the shipper to increase profits for a given o-d lane?

o The chargeable weight rate β_i to be charged to the shipper is the main variable to analyze.

o The target rates τ_i provided by the shipper during the different rounds of the RFQ/tender processes are reliable references of a lower bound / threshold.

o The probability of being awarded with a certain shipment of cargo increases as we lower the proposed rate β_i for a current bid. However, given the information comprised on the datasets we were provided with by the sponsor company, we do not have enough information in our database to estimate with enough accuracy probabilities and confidence levels through marginal analysis on sensitivity.

o We consolidate all the cargo present in a given o-d lane, considering 100% efficiency at an operational level. We do not consider inefficiencies derived from inadequate operational strategies.

o All RFQ and units are normalized per year, to facilitate comparison between data and an homogeneous treatment of the datasets within the analytical model.

o To design the framework for the analytical model we will exclusively focus on analyzing 9 customers/shippers from different industries and on 9 routes between 3 Chinese cities and 3 American cities defined in Table 4.0.1.

Profits are defined for each o-d lane, as a result of the consolidation of cargo from all customers that operate in a given lane, as:

$$\Pi = \sum_{k=1}^{n} \sigma_i \cdot \beta_i - \sum_{k=1}^{n} \alpha_i \cdot \varrho_i$$

Selecting the most attractive incoming bid to maximize profitability, therefore, will depend primarily on three variables for each piece of potential business or current bid: the rate β_i we charge to the shipper, the consolidation potential Δ of the bid under consideration, and the density δ of the current bid with respect to the density of the current business, as shown in this equation:

$$\Pi = \Pi(\beta_i, \Delta, \delta)$$

After selecting the essential variables that will describe the current business as well as the estimated consolidation and profitability for the current bid after consolidation with the current business for a given o-d lane, we built a dynamic database to integrate all the datasets from different customers and be able to perform segmentation analysis for different purposes: per lane, per customer, per country, etc.

The sponsor company provided us with the records of data for customers forth above in classic spreadsheet format. However, air freight forwarder companies require data management tools that permit to analyze information in a recurrent, automatic and systematic way through calculations that operate over the incoming RFQ datasets, to synthesize and interpret as quickly as possible the results of the analysis and knowingly participate in the bidding processes.

The flexibility that characterizes spreadsheets has unfortunately a counterpart: while it is easy to create formulas, reference cells, copy and paste data, and link worksheets and spreadsheets together, as the work gets more complex, spreadsheets become more difficult to change and manage. While spreadsheets are ideal for creating one time analysis, they become problematic as the data grows and evolves over time. As new rows and columns get added, summary ranges and formulas may need to be modified or new ones created, data and formulas are not consistently updated, and these mistakes lead to bad results and decisions. The challenges of spreadsheets are due to the difficulty maintaining them accurately over time and scaling the volume.

The need to constantly handle new inputs over time and to integrate and combine different variables from different datasets to create complex metrics using a scalable framework that is a characteristic necessity for the air freight industry decided us to the need to find an alternative to traditional spreadsheets. Therefore, we chose to build the analytical model using management software database. The main advantages of management programs database on the spreadsheet are data structure and normalization through multiple tables, scalability, data and referential integrity, queries and reports and automation through macros and VBA modules.

The tool selected to build such database was the SQL server software, because it allows users to perform through queries calculations and macros and to simultaneously submit summary reports that integrate exclusively selected variables and parameters for a particular purpose. Such synthetic reports will be the basis to build the visualization tools that express in a concise and intuitive way a synthesis of the results of the calculations and parameters which are the most representative with respect to decision-making of the commercial strategy.

4.2.- Analyze Database: 2 sets of data

The air rate data set contains information regarding the participation of our sponsor company in the bidding processes on lanes with different customers. A lane consists of an origin and a destination, and an amount of cargo that a shipper want sent during a certain time period. The bidding process is a process where a shipper announces a lane that freight forwarders can bid on. The bids consist of a rate, the price, which the shipper will have to pay to the freight forwarder in order to have the cargo moved. The rate is expressed in US Dollars per chargeable kilogram shipped. The amount of chargeable kilograms is the greatest value of the gross weight and the volumetric weight, the gross weight is just the mass while the volumetric weight is found by multiplying the volume of the shipment expressed in cubic meters by 1000/6, or roughly 167. This product is then the volumetric weight in kilograms. The conversion rate is derived from the ratio of 1 metric ton per 6 cubic meters, which is the industry standard for determining the chargeable weight of air cargo.

This data set contains data on bidding on lanes with 9 different customers. The data was received in two separate Excel files and the data is separated by customer so that each customer is

on a separate tab. The tabs are not standardized meaning that the columns on the tabs are different, in different order, and columns including similar information may have different names. However, the sheets have between 60 and 120 columns, and as mentioned above the names of the columns are not the same.

In order to create a coherent data set we needed to decide what data we wanted and determine which columns contained that information. An alternative approach would be to create a complete dataset with all the information from spreadsheets. Due to the large number of columns and the non-standardized names it would be time prohibitive to go through every column for each customer and determine which of them are the same. Our assessment is furthermore that it would not add much value to our analysis; most columns contain information that is not relevant for our core analysis. For example for some customers there is a lot of information tracking cost and services to and from the airport, our focus is on the actual air transport so we only need the information related to that.

To determine what data we needed we started with the very basic information on the lanes like origin, destination, the chargeable weight per year, number of shipments, and the relevant time period. The next type of data we needed was the data related to the bidding process. There are two types of data related to the bidding process: bids and targets. Bids are the price suggested by our sponsor to the customer, usually there are several rounds of bidding so there might be several bids for one lane. In our consolidated data sheet we ended up with columns to represent four different rounds since this was the highest number of rounds for any of the customers. Targets (τ) are prices provided by the customer, which they think the bidding freight forwarders should aim for, just like bids there might be different targets τ for each round. Sometimes the targets are given as a rate in dollars per kg and sometimes it is expressed as how far off a bid was, for example a bid might be said to be 0-25% above the target. The underlying information is obviously the same, but if it is to be used in a coherent way it needs to be converted, preferably so that both values are available: the actual target rate and how far off the bid was. A related relevant piece of data is if our sponsor won the lane, and if so, how many percent of it they won. This can be used to analyze the business they ended up with, and contrasting it to the business that was available.

Another necessary part of the data is information on the cost for our sponsor on each lane. The dataset contains several different types of costs that oscillates among customers, but in dialogue

with our sponsor company we will use three specific columns as the operational cost: the Net/Net rate, the fuel surcharge, and the security surcharge. The Net/Net rate is the basic rate that the carrier charges the freight forwarder, while the fuel and security surcharges, as the names strongly imply, are additional fees that are levied on top of the basic rate. Our sponsor consider these three combined to be the operational cost and that their operating margin is the price they charge to the shipper, minus the operational costs they pay to the carrier. The operating margin then needs to cover additional costs and overhead, as well as net profit. For the purpose of our thesis we will use these costs in the same way. In addition to being in line with how the sponsor sees it, it also has the clear benefit of clarity in the data. The three costs columns are present for all the customers in a very consistent way. One issue with the operational costs is that for some lanes there are several different costs which are dependent on which weight bracket a shipment on lane will end up in. The different brackets mean there can be several different operational costs contained in a single row.

To be able to estimate the costs for lanes with different costs we made the assumption that every shipment within that lane will be the same size. This is probably not entirely realistic, but according to our sponsor it is the general practice in the industry, so it may not be too far of in most cases. This assumption allows us determine how big each shipment will be by just dividing the total chargeable kilograms per year by the number of shipments and using this average to determine which weight bracket is most appropriate to use when determining the cost. In practice we did this by splitting single rows into multiple rows where every row only has one weight bracket. Row in this context is referring to a row in a spreadsheet, in most cases one row corresponds to one lane, but here we use the term row since we are splitting it so that one lane is actually represented by several rows. After that we only kept the ones where the average shipment size corresponded to the weight bracket. This does in absolute terms that some data is lost, namely the costs for other weight brackets, but in practice it means we make more data available for the analysis, because if we cannot decide which cost to use for a certain lane we cannot use that lane at all.

The final core piece of information needed by us is data on the gross weight and volumetric weight of the lane. The purpose of this data is to evaluate the potential and level of consolidation that is possible and obtained. There is however no information of this type in the dataset.

In order to be able to do any form of work on how to handle consolidation we therefore assigned weight-to-volume ratios for each lane by using the lognormal distribution described

elsewhere. To be precise we generated the denominator of the ratio assuming the numerator was 1. The ratios were then used together with the chargeable weight to calculate the gross and volumetric weights. This was done by normalizing the denominators to the 1:6 ratio by dividing them by 6. If this quotient was greater than 1 the chargeable weight was divided by it to get the gross weight and the volumetric weight would be equal to the chargeable weight. If on the other hand the quotient was less than one it would be multiplied with the chargeable weight to get the volumetric weight, and the gross weight would be equal to the chargeable weight. All of these columns discussed above were then copied into one tab in a new excel spreadsheet so that the relevant data from all the nine customers were collected in one table.

The next step after cleaning and consolidating the data was to use the data to calculate the metrics we developed. We considered a few different ways to do this. An easy way would be to use Excel for calculations. The benefit of Excel is mainly simplicity. It has a lot of functions already built in and it is very easy to make one set of calculations, then using those calculations for another step of calculations by using cell references. The main problem with this approach is scalability and performance. In our dataset we only had nine customers, but that alone added up to over 25 000 rows of data. The way Excel works means that to complete a calculation a column has to be added where the formula is inputted and then duplicated to all the rows in the column. Already with 25 000 rows this ended up being a somewhat sluggish process and on several occasions Excel simply became unresponsive. It is easy to see how this problem would be greatly exacerbated if all customers were added, or a longer time period was covered. In addition the process of adding columns is essentially adding the data, this not only increases the size of the file and makes performance, but also introduces risks of data corruption since manipulations are done directly in then spreadsheet where the data is stored. A final issue was that is Excel does not handle relationship between different pieces of information very well, so if at some future point the system should be extended to include for example RFI data that could be linked with bid data through a common ID that is very unpractical to do in Excel. Due to the performance issues, the risk of data corruption, and the problems to manage relationships we decided that Excel was not sufficient.

We also tried using Microsoft Access which allowed better performance with more data and the possibility of running queries instead of creating new columns. We did however encounter some limitations with Access as well. One function we absolutely needed to have was one that could look at two columns in the same row and select the highest value of the two. This would be used to

determine the chargeable weight by comparing the gross and volumetric weights. We would also like to be able to generate aggregates of a set of rows, and then use those aggregates in relation to other individual rows. The purpose would be to determine the current business and then consider how adding another bid would affect it. We could not find any good ways to do this in Access so decided to switch to Microsoft SQL Server instead. Of the three SQL server is definitely the most powerful and scalable option, the drawback is that it is probably also the one that it takes the longest to learn. However since it is so powerful it is possible to set it so that it just takes the inputted data, performs the necessary calculations and then outputs it directly to Tableau. Meaning that most user will not have to know SQL, just how to import the data.

After deciding to use SQL we imported the cleaned data from Excel as a new table in a new SQL database.

4.3.- Data scarcity and dataset population by statistical simulation

The lack of information in the Air Rates dataset regarding the volume of cargo for each tender prevents the calculation of the density of volume δ to be transported, and therefore, any estimate of potential consolidation delta Δ. Hence, we chose to simulate the density values υ_w (DIM Factor) of incoming bids using as a reference the only values available: the densities of the RFI the sponsor company provided us with. Specifically, only two out of the nine RFI customers (RFI01 and RFI03) presented data on gross weight γ_w and volumetric weight υ_w that allowed for the calculation of the density of the cargo to be transported. Being customers from different industries (lifestyle/fashion and industrial) and being the results of the analysis similar enough as a whole we decided to take as a reference for the distribution of densities the aggregate datasets on density data from both clients RFI01 and RFI03.

The main criterion we used to realistically simulate the density values is the statistical analysis via goodness of fit tests. A posteriori we randomly assigned the resulting values of the selected probability distributions to populate the column with the mission values.

The other major limitation has been as a result of having very few data on the gross weight (γ_w), only 6,162 out of 25,926 cases. Similarly, we decided to populate the gross weight cells with randomized data based on the resulting distribution of the goodness of fit tests applied to the data

recorded with respect to gross weight in order to significantly strengthen the datasets in a realistic and reasonable way to make the most out of the visualization tools.

The fact that we randomly assigned the values of the resulting distributions naturally ignores the possible internal correlations between certain variables, but this fact is beyond the scope of this thesis and does not alter at all the main objective to pursue: the construction of an analytical model using as a reference datasets that in the future can systematically be gathered and stored. We performed the Goodness of Fit tests for different statistical distributions (Normal, Lognormal, Chi-Squared, Weibull, Gamma) that can potentially resemble the distribution of such variable to be able to simulate in a realistic way values for volumetric density (DIM Factor), in order to use this randomized set of δ values to populate the historical Air Rates data.

The criteria used to determine the goodness of fit of the samples to a series of theoretical distributions are the non-parametric tests known as Kolmogorov-Smirnov and Anderson-Darling. Although the statistical analysis of goodness of fit test of the sample datasets presented different degrees of adjustment to each of the distributions, the Kolmogorov-Smirnov presented in all the cases significantly higher goodness of fit levels than those of Anderson-Darling. The reason is that although Anderson-Darling is based on a similar approach to Kolmogorov-Smirnov, the former uses a more comprehensive measure of difference tan the latter, so that Anderson-Darling is more sensitive to the tails of the distribution than Kolmogorov-Smirnov.

Consequently, we decided to choose the theoretical distribution that best fit the sample taking as reference the goodness of fit test Kolmogorov-Smirnov, which is the only test which presented in all cases enough positive cases with confidence levels above 20%.

Regarding hypothesis testing as per Kolmogorov-Smirnov, the null and the alternative hypotheses are:

H_0: the data follows the specified distribution;

H_A : the data does not follow the specified distribution.

The hypothesis regarding the distributional form will be rejected at the chosen significance level α if the test statistic, D, is greater than the critical value obtained from a table. The fixed values of α (0.01, 0.05 etc.) are used to evaluate the null hypothesis H_0 at various significance levels.

Before merging the data from both RFI01 and RFI03, we analyzed them separately and we obtained significantly similar results: in both cases we obtained the best adjustment by using the lognormal distribution (see Figure 4.3.1). In the case of the RFI01, the lognormal distribution was adjusted with parameters σ=0.28459 μ=2.1363 as described in Tables 4.3.1 to 4.3.3 and Figures 4.3.2 to 4.3.5. With respect to RFI03, the lognormal distribution was adjusted with parameters σ=0.27317 and μ=2.3846, as described in Tables 4.3.4 to 4.3.6 and Figures 4.3.6 to 4.3.9.

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\Pi}} e^{\frac{-(\ln(x)-\mu)^2}{2\sigma^2}}$$

Figure 4.3.1a. Theoretical lognormal distribution chosen to simulate and randomize density values to populate the incoming bids datasets

However, as seen in Figures 4.3.7 and 4.3.8, we identified an abnormal spike in the RFI03 dataset around the DIM Factor value 6: up until 248 cases presented a generic default reference "6" value for the 1:6 WV Ratio that did not correspond to the real density of shipments. Thus, we ignored those default pieces of information from our analysis in order to avoid bias in the results.

So, after isolating the value of "6" present by default in 248 cases and merging both data records RFI01 and RFI03, as described in Tables 4.3.7 to 4.3.9 and Figures 4.3.10 to 4.3.13, the resulting distribution that provides the best goodness of fit was lognormal with parameters σ=0.33116 μ=2.2633, shown on Figure 4.3.1b.

$$f(x; 2.2633, 0.33116) = \frac{1}{x \cdot 0.33116\ \sqrt{2\Pi}} \cdot e^{\frac{-(\ln(x)-2.2633)^2}{2 \cdot 0.33116^{\ 2}}}$$

Figure 4.3.1b. Lognormal distribution with adjusted parameters chosen to simulate and randomize density values to populate the incoming bids datasets

Therefore, we selected the lognormal function defined by Figure 4.3.1b as the result of performing the goodness of fit tests on merged datasets RFI01 and RFI03b to simulate randomized numbers to populate the missing DIM Factor/density of cargo values on the Air Rates database. By doing so, we were able to have enough pieces of information regarding Chargeable Weights α and σ as to perform analysis on the level of consolidation and profitability.

Table 4.3.1. Descriptive Statistics of the density/DIM Factor datasets - RFI client 01

Statistic	Value
Sample Size	972
Range	14
Mean	9.536
Variance	6.6623
Std. Deviation	2.5811
Coef. of Variation	0.27067
Std. Error	0.66645
Skewness	0.91154
Excess Kurtosis	0.88642

Percentile	Value
Min	4
5%	6
10%	7
25% (Q1)	8
50% (Median)	9
75% (Q3)	11
90%	13
95%	15
Max	18

Table 4.3.2. Adjusted Parameters of the statistical distributions provided by the Goodness of Fit Tests applied to the density/DIM Factor of the RFI client 01

#	Distribution	Parameters
1	Chi-Squared	n=9
2	Chi-Squared (2P)	n=6 g=3.7925
3	Gamma	a=13.649 b=0.69864
4	Gamma (3P)	a=7.5349 b=0.92276 g=2.5831
5	Lognormal	s=0.26176 m=2.2204
6	**Lognormal (3P)**	**s=0.28459 m=2.1363 g=0.71642**
7	Normal	s=2.5811 m=9.536
8	Weibull	a=3.7566 b=10.519
9	Weibull (3P)	a=2.35 b=6.5151 g=3.7621
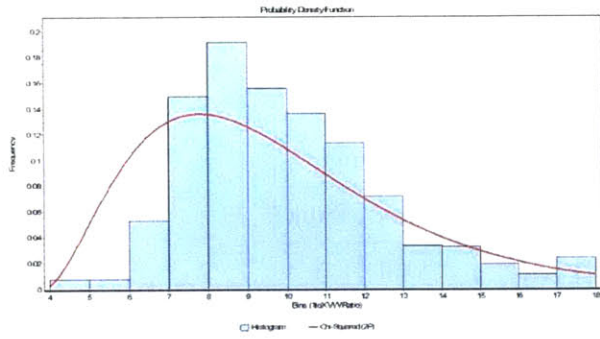
Chi-Squared
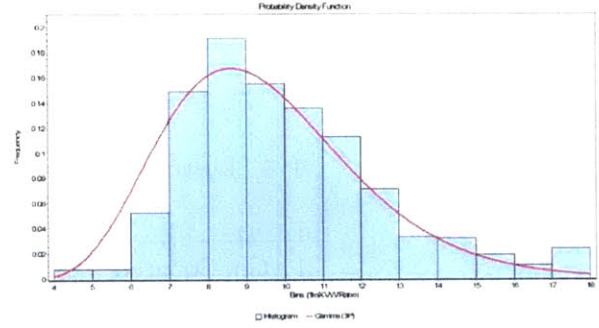
Gamma 3P
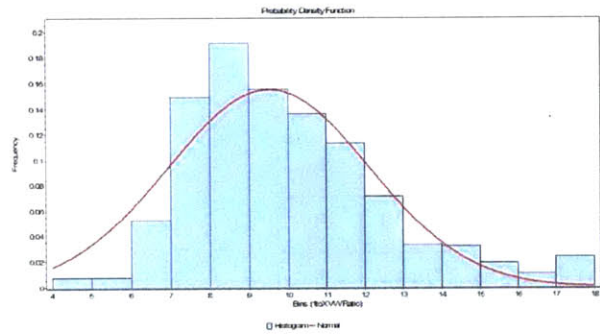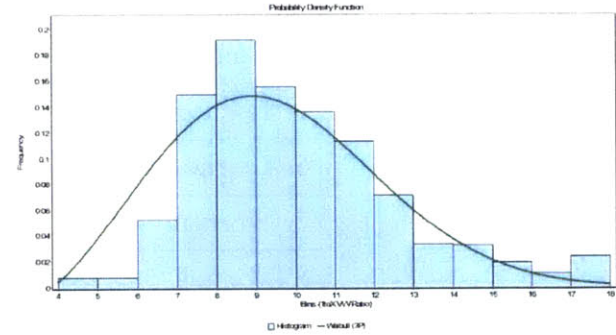
Normal

Weibull
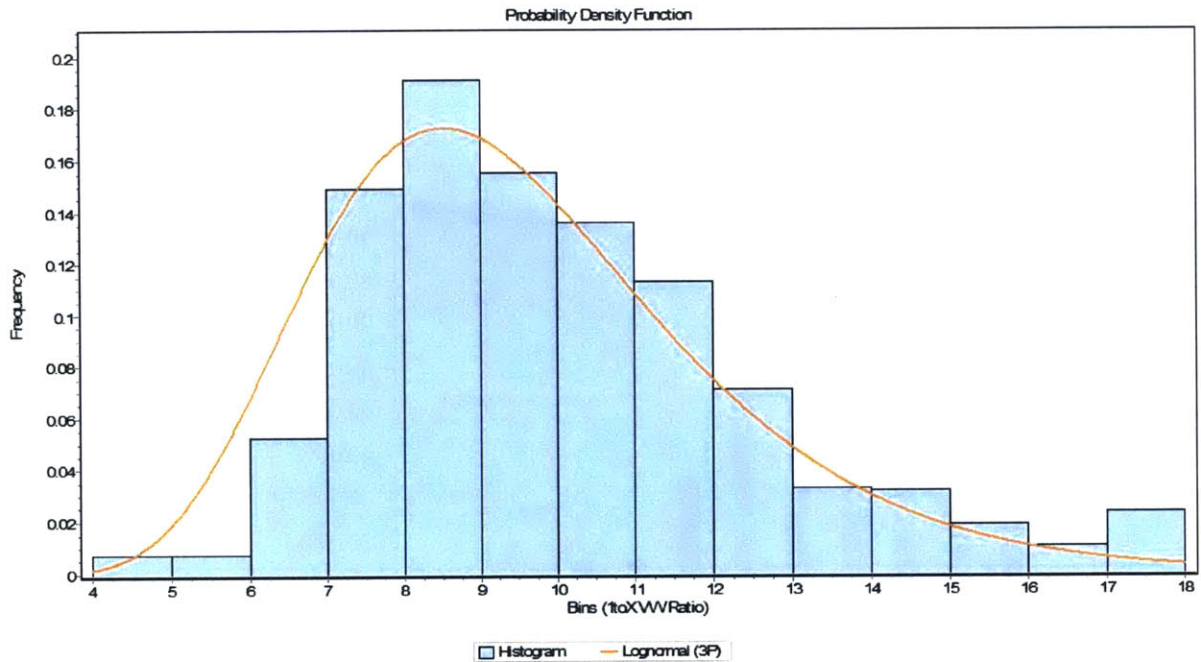
Lognormal distribution

Figure 4.3.2. Results of the Goodness of Fit Tests applied to the density/DIM Factor of the RFI client 01

Table 4.3.3. Summary ranking of goodness of fit tests for the density/DIM Factor dataset - RFI client 01

#	Distribution	Kolmogorov Smirnov		Anderson Darling	
		Statistic	Rank	Statistic	Rank
1	Chi-Squared	0.29601	9	2.0004	1
2	Chi-Squared (2P)	0.15074	8	2.0093	2
3	Gamma	0.11239	5	3.977	4
4	Gamma (3P)	0.11064	2	4.2914	8
5	Lognormal	0.11235	4	4.2459	7
6	**Lognormal (3P)**	**0.10915**	**1**	**4.224**	**6**
7	Normal	0.14501	7	4.6899	9
8	Weibull	0.13601	6	3.9444	3
9	Weibull (3P)	0.11221	3	4.1193	5
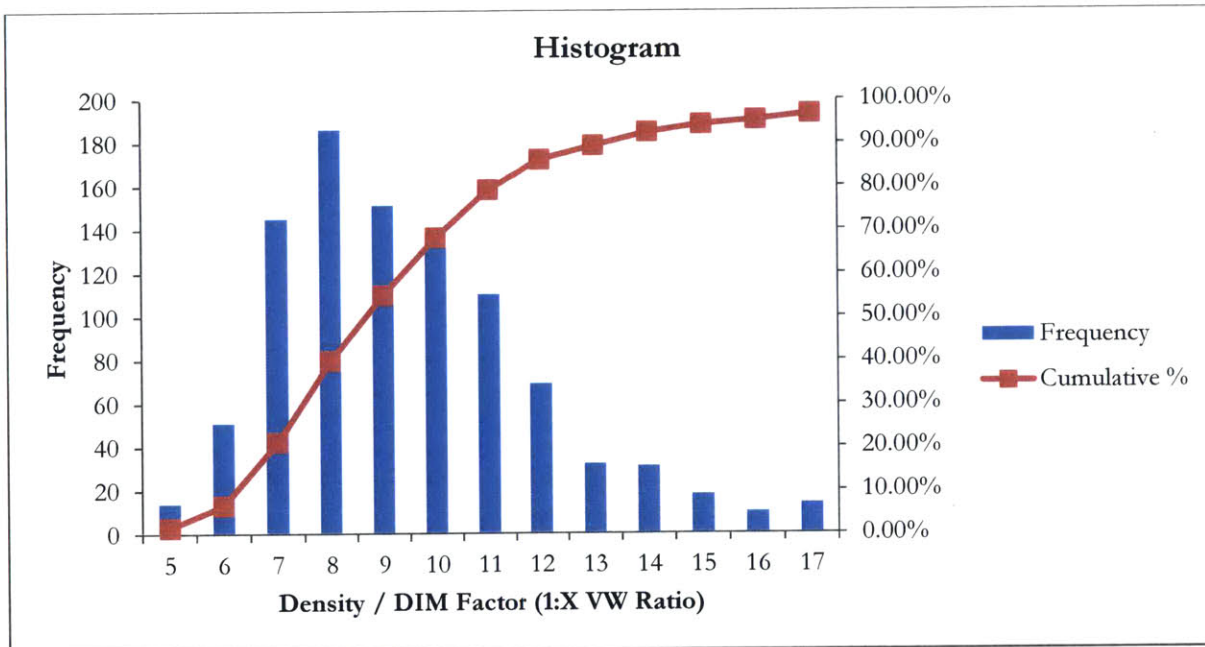
Figure 4.3.3. Histogram of the sample dataset for the density/DIM Factor - RFI client 01

Figure 4.3.4. Comparison between RFI01 and Expected Lognormal - Frequency Distribution for the density/DIM Factor - RFI client 01

Figure 4.3.5. Comparison between cumulative distributions for real and expected Frequency

Lognormal - Distribution for the density/DIM Factor - RFI client 01

45

Table 4.3.4. Descriptive Statistics of the density/DIM Factor datasets - RFI client 03

Statistic	Value
Sample Size	2552
Range	19
Mean	9.4469
Variance	9.9128
Std. Deviation	3.1485
Coef. of Variation	0.33328
Std. Error	0.50416
Skewness	0.80227
Excess Kurtosis	0.51702

Percentile	Value
Min	1
5%	6
10%	6
25% (Q1)	7
50% (Median)	9
75% (Q3)	11.5
90%	13.5
95%	15.5
Max	20

Table 4.3.5. Adjusted Paremeters of the statistical distributions provided by the Goodness of Fit Tests applied to the density/DIM Factor of the RFI client 03

| #	Distribution	Parameters
1	Chi-Squared	n=9
2	Chi-Squared (2P)	n=8 g=0.91957
3	Gamma	a=9.0029 b=1.0493
4	Gamma (3P)	a=9.2366 b=1.017 g=0.05279
5	Lognormal	s=0.33353 m=2.1913
6	**Lognormal (3P)**	**s=0.27317 m=2.3846 g=-1.8224**
7	Normal	s=3.1485 m=9.4469
8	Weibull	a=3.1435 b=10.548
9	Weibull (3P)	a=2.8522 b=9.5703 g=0.91702

Chi-Squared
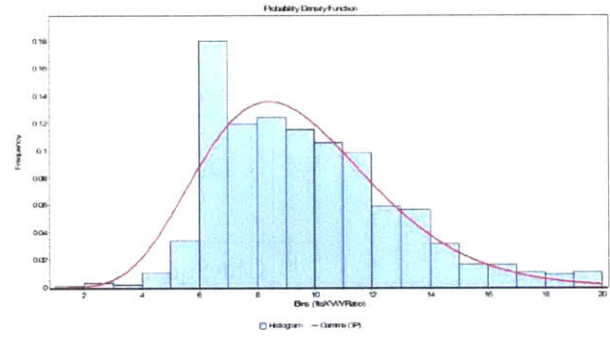
Gamma 3P
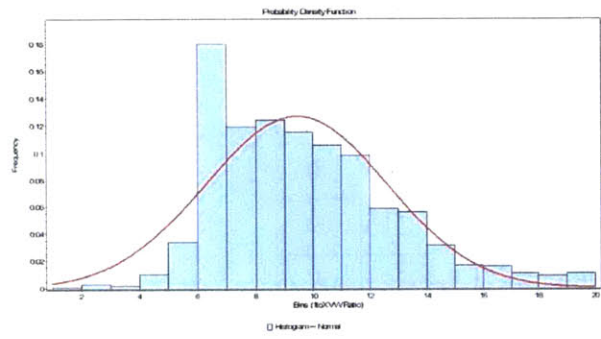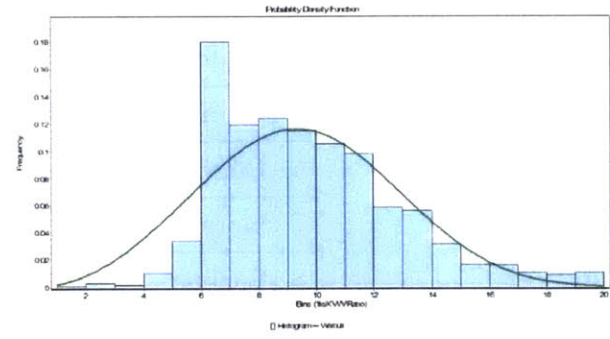
Normal

Weibull
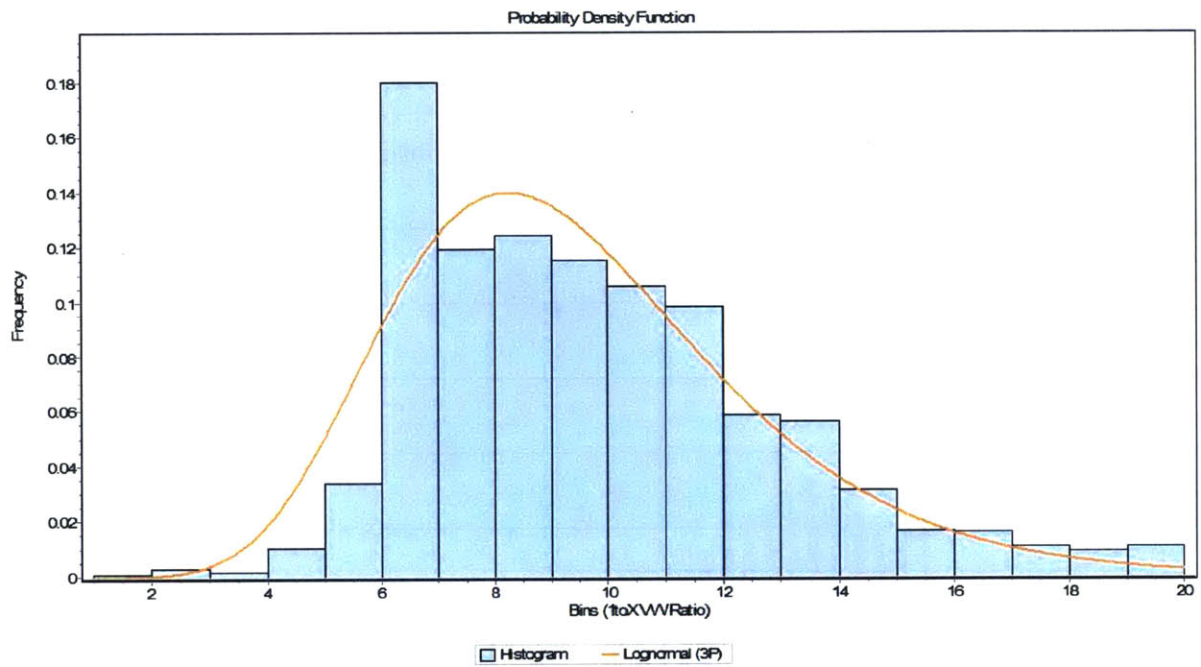
Lognormal distribution

Figure 4.3.6. Results of the Goodness of Fit Tests applied to the density/DIM Factor of the RFI client 03

Table 4.3.6. Summary ranking of goodness of fit tests for the density/DIM Factor dataset - RFI client 03

#	Distribution	Kolmogorov Smirnov		Anderson Darling	
		Statistic	Rank	Statistic	Rank
1	Chi-Squared	0.21093	9	4.8962	1
2	Chi-Squared (2P)	0.2019	8	6.6565	2
3	Gamma	0.07568	4	11.636	6
4	Gamma (3P)	0.07036	3	12.34	7
5	Lognormal	0.06634	2	13.886	9
6	**Lognormal (3P)**	**0.06603**	**1**	**12.535**	**8**
7	Normal	0.09087	5	10.831	5
8	Weibull	0.10696	7	9.4922	3
9	Weibull (3P)	0.10253	6	10.11	4

Comparison between RFI03 and Expected Lognormal - Frequency Distribution
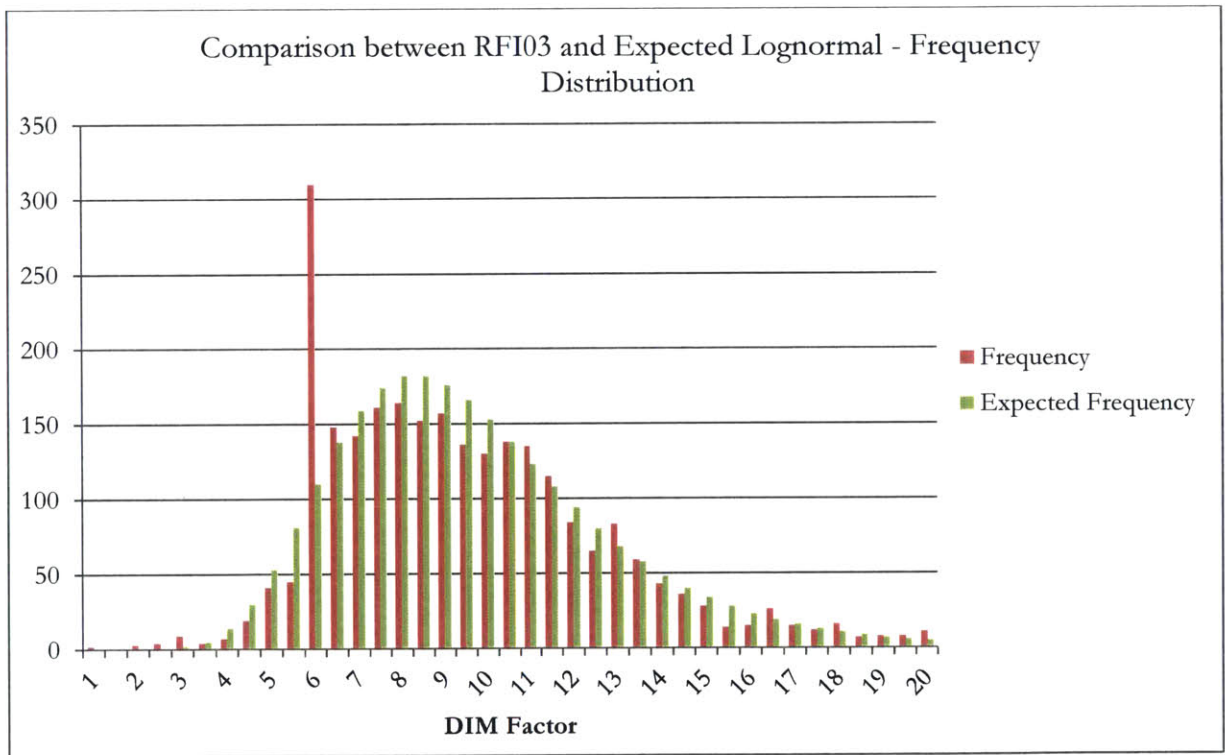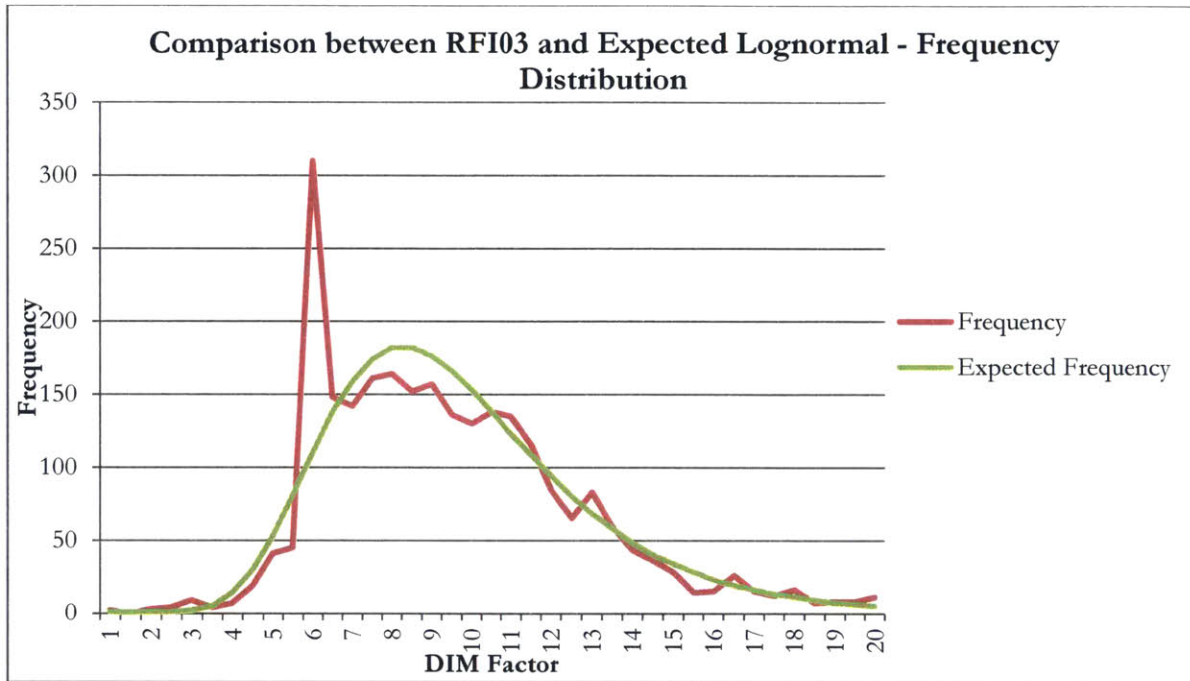
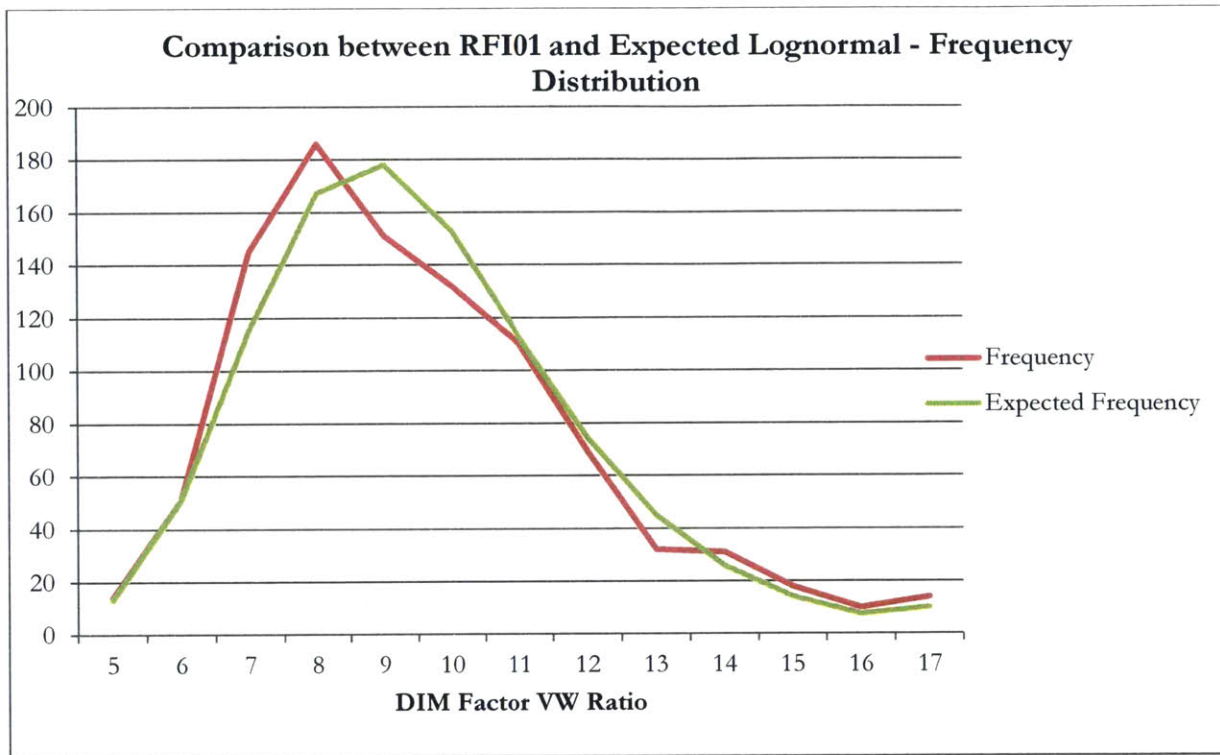Figures 4.3.7 and 4.3.8. Comparison between RFI03 and Expected Lognormal - Frequency Distribution for the density/DIM Factor - RFI client 03
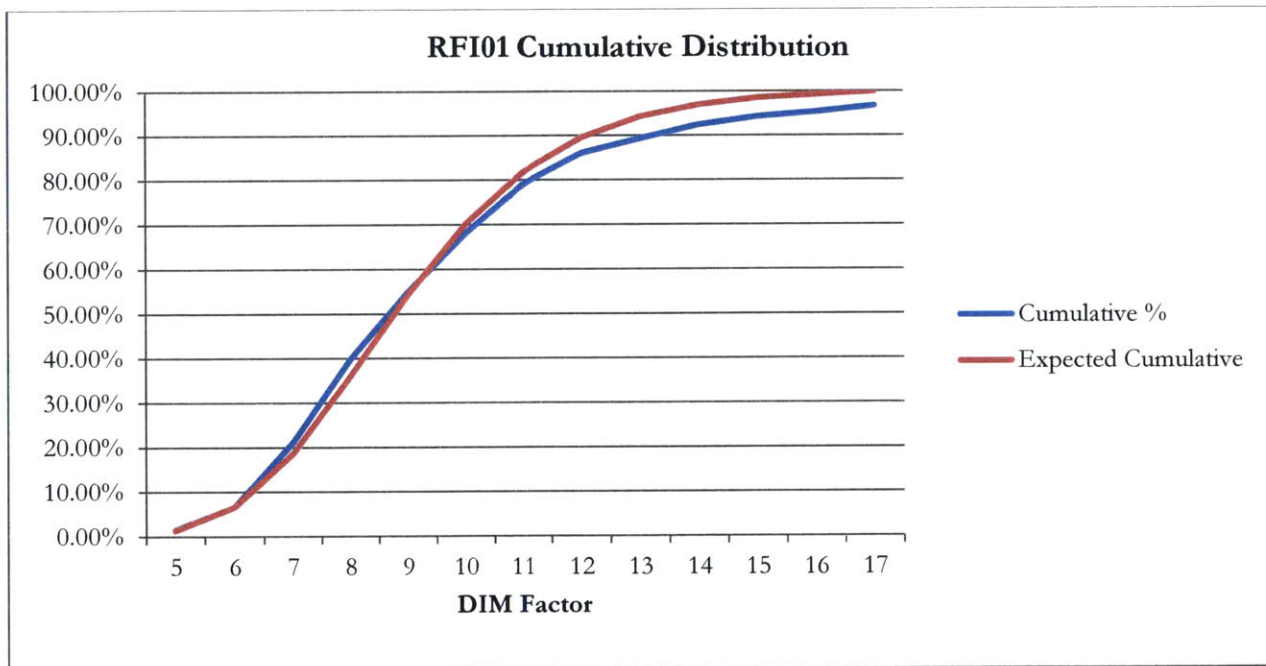

Figure 4.3.9. Comparison between cumulative distributions for real and expected Frequency

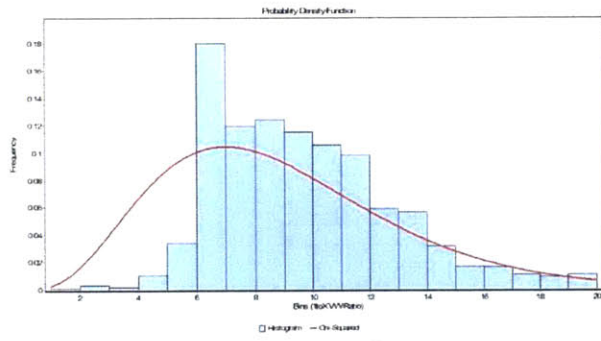Lognormal - Distribution for the density/DIM Factor - RFI client 03

49

Table 4.3.7. Descriptive Statistics of the density/DIM Factor datasets - RFI client 03b

Statistic	Value
Sample Size	2304
Range	19
Mean	9.9944
Variance	9.752
Std. Deviation	3.1228
Coef. of Variation	0.31246
Std. Error	0.69828
Skewness	0.69495
Excess Kurtosis	0.55381

Percentile	Value
Min	1
5%	6
10%	7
25% (Q1)	8
50% (Median)	10
75% (Q3)	12
90%	14
95%	16
Max	20

Table 4.3.8. Adjusted Parameters of the statistical distributions provided by the Goodness of Fit Tests applied to the density/DIM Factor of the RFI client 03b

| #	Distribution	Parameters
1	Chi-Squared	n=9
2	Chi-Squared (2P)	n=9 g=0.8466
3	Gamma	a=10.243 b=0.97574
4	Gamma (3P)	a=14.382 b=0.81392 g=-1.7114
5	**Lognormal**	**s=0.32404 m=2.2523**
6	Lognormal (3P)	s=0.20436 m=2.6888 g=-5.0312
7	Normal	s=3.1228 m=9.9944
8	Weibull	a=3.349 b=11.113
9	Weibull (3P)	a=3.0864 b=10.255 g=0.80967
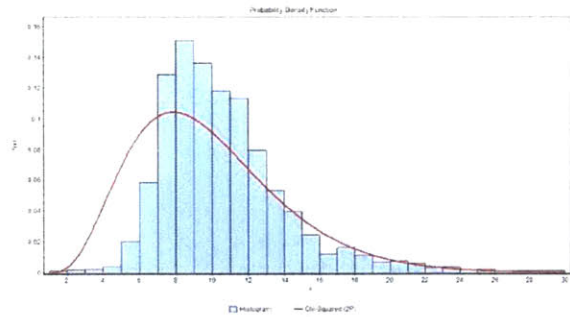
Chi-Squared
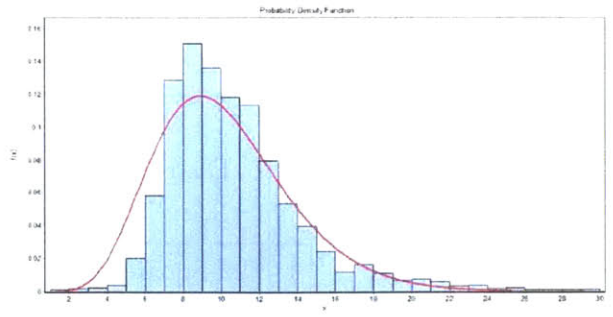
Gamma 3P
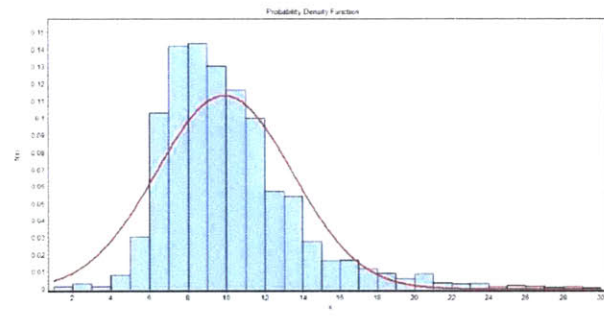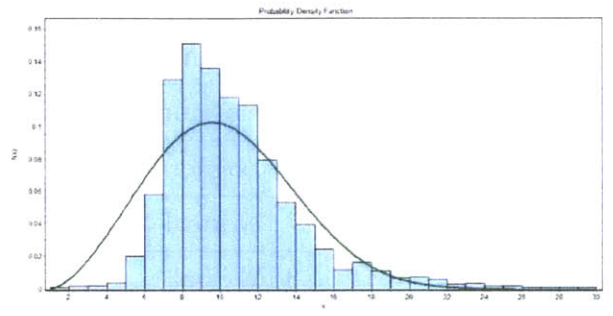
Normal

Weibull
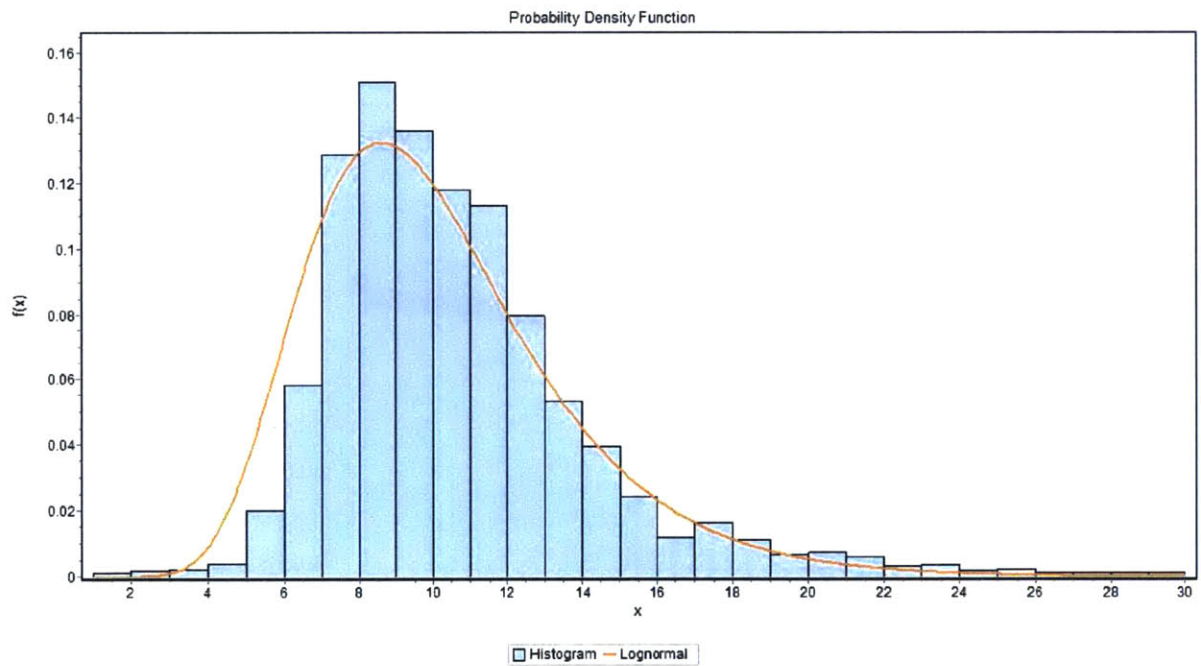
Lognormal distribution

Figure 4.3.10. Results of the Goodness of Fit Tests applied to the density/DIM Factor of the RFI client 03b

Table 4.3.9. Summary ranking of goodness of fit tests for the density/DIM Factor dataset - RFI client 03

#	Distribution	Kolmogorov Smirnov		Anderson Darling	
		Statistic	Rank	Statistic	Rank
1	Chi-Squared	0.26566	9	2.7806	1
2	Chi-Squared (2P)	0.17831	8	3.1602	2
3	Gamma	0.08171	2	7.174	8
4	Gamma (3P)	0.08897	4	6.6199	6
5	**Lognormal**	**0.07547**	**1**	**8.4941**	**9**
6	Lognormal (3P)	0.08487	3	6.68	7
7	Normal	0.11786	7	5.5526	5
8	Weibull	0.10421	6	4.9873	3
9	Weibull (3P)	0.09967	5	5.4446	4
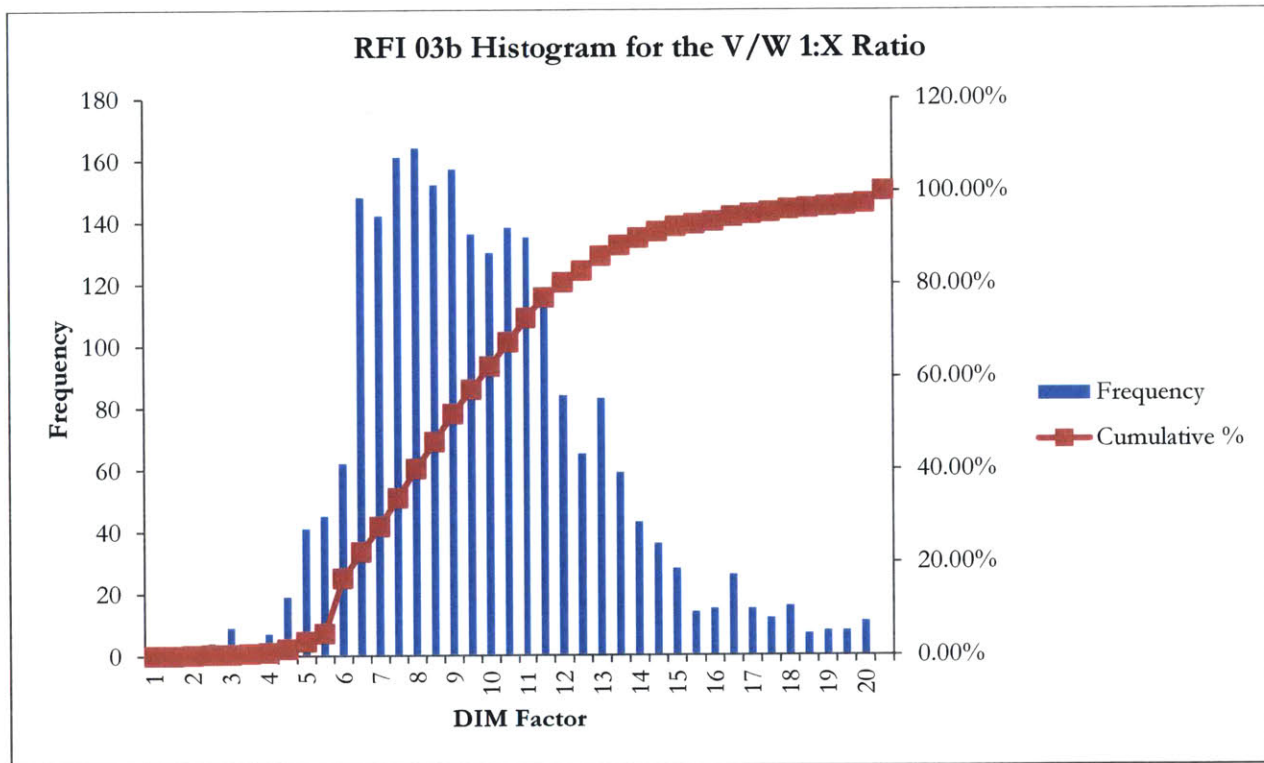
Figure 4.3.11. Histogram of the sample dataset for the density/DIM Factor - RFI client 03b
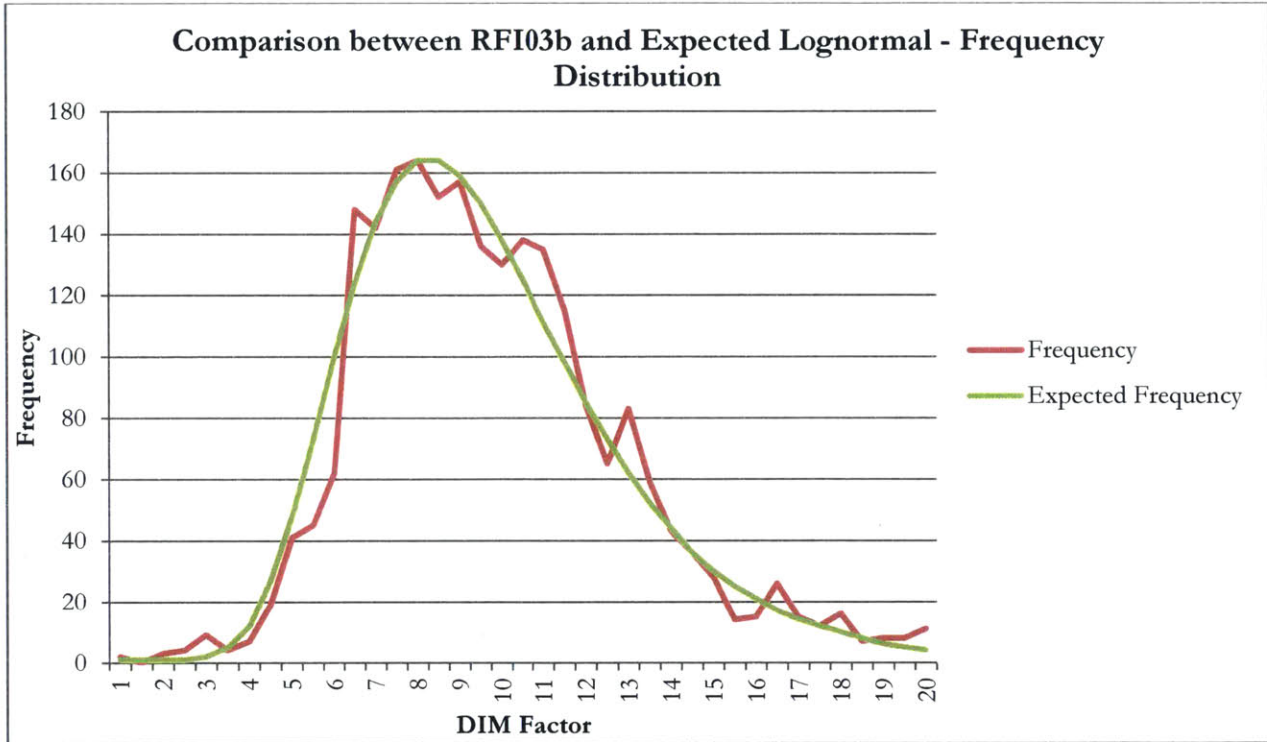
Figure 4.3.12. Comparison between RFI01 and Expected Lognormal - Frequency Distribution for the density/DIM Factor - RFI client 03b
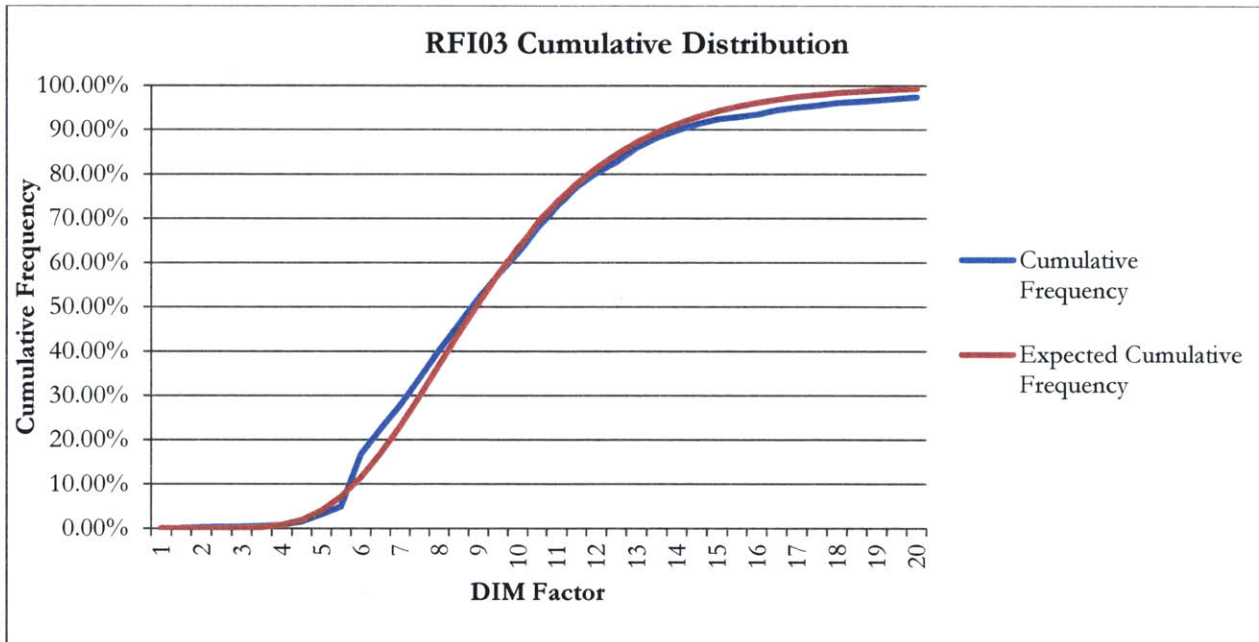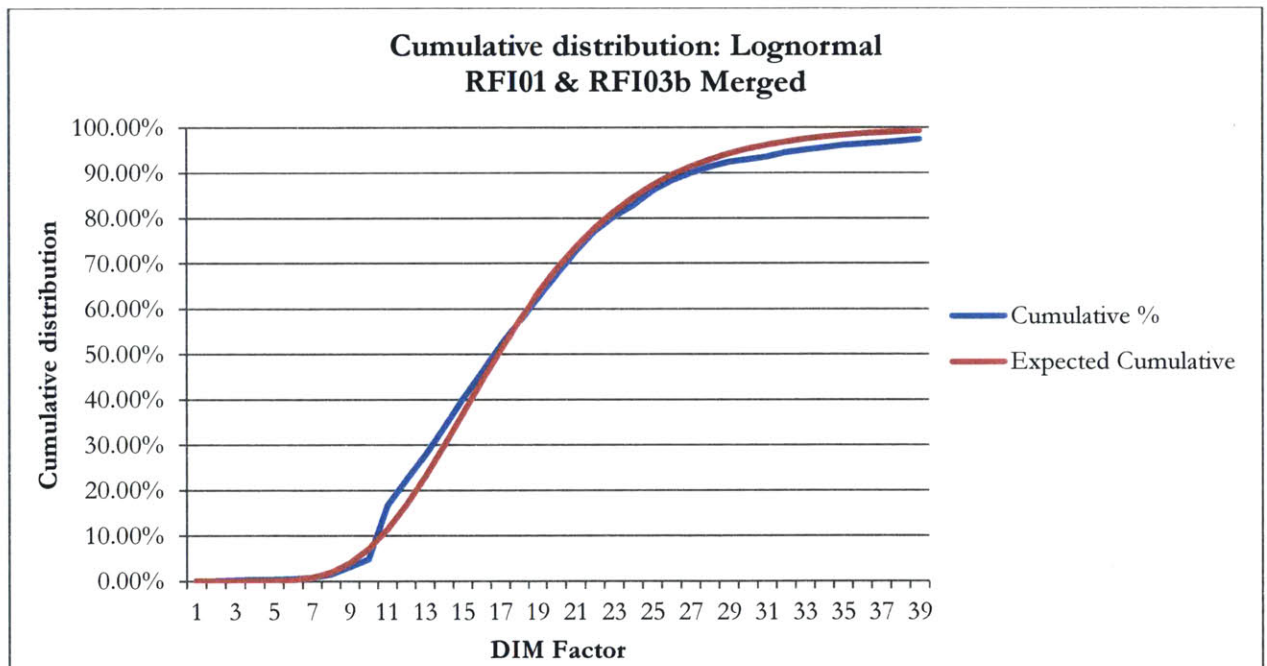

Figure 4.3.13. Comparison between cumulative distributions for real and expected Frequency Lognormal - Distribution for the density/DIM Factor - RFI client 03b

Table 4.3.10. Descriptive Statistics of the density/DIM Factor datasets - RFI clients 01 & 03b

Statistic	Value		Percentile	Value
Sample Size	3276		Min	1
Range	29		5%	6
Mean	10.157		10%	7
Variance	12.552		25% (Q1)	8
Std. Deviation	3.5429		50% (Median)	9
Coef. of Variation	0.34881		75% (Q3)	12
Std. Error	0.64685		90%	14
Skewness	1.4522		95%	17
Excess Kurtosis	3.482		Max	30

Table 4.3.11. Adjusted Parameters of the statistical distributions provided by the Goodness of Fit Tests applied to the density/DIM Factor of the RFI clients 01 & 03b

#	Distribution	Parameters
1	Chi-Squared	n=10
2	Chi-Squared (2P)	n=9 g=0.88763
3	Gamma	a=8.2192 b=1.2358
4	Gamma (3P)	a=8.8716 b=1.1219 g=0.20418
5	**Lognormal**	**s=0.33116 m=2.2633**
6	Lognormal (3P)	s=0.28786 m=2.398 g=-1.3215
7	Normal	s=3.5429 m=10.157
8	Weibull	a=2.8874 b=11.354
9	Weibull (3P)	a=2.653 b=10.349 g=0.93682
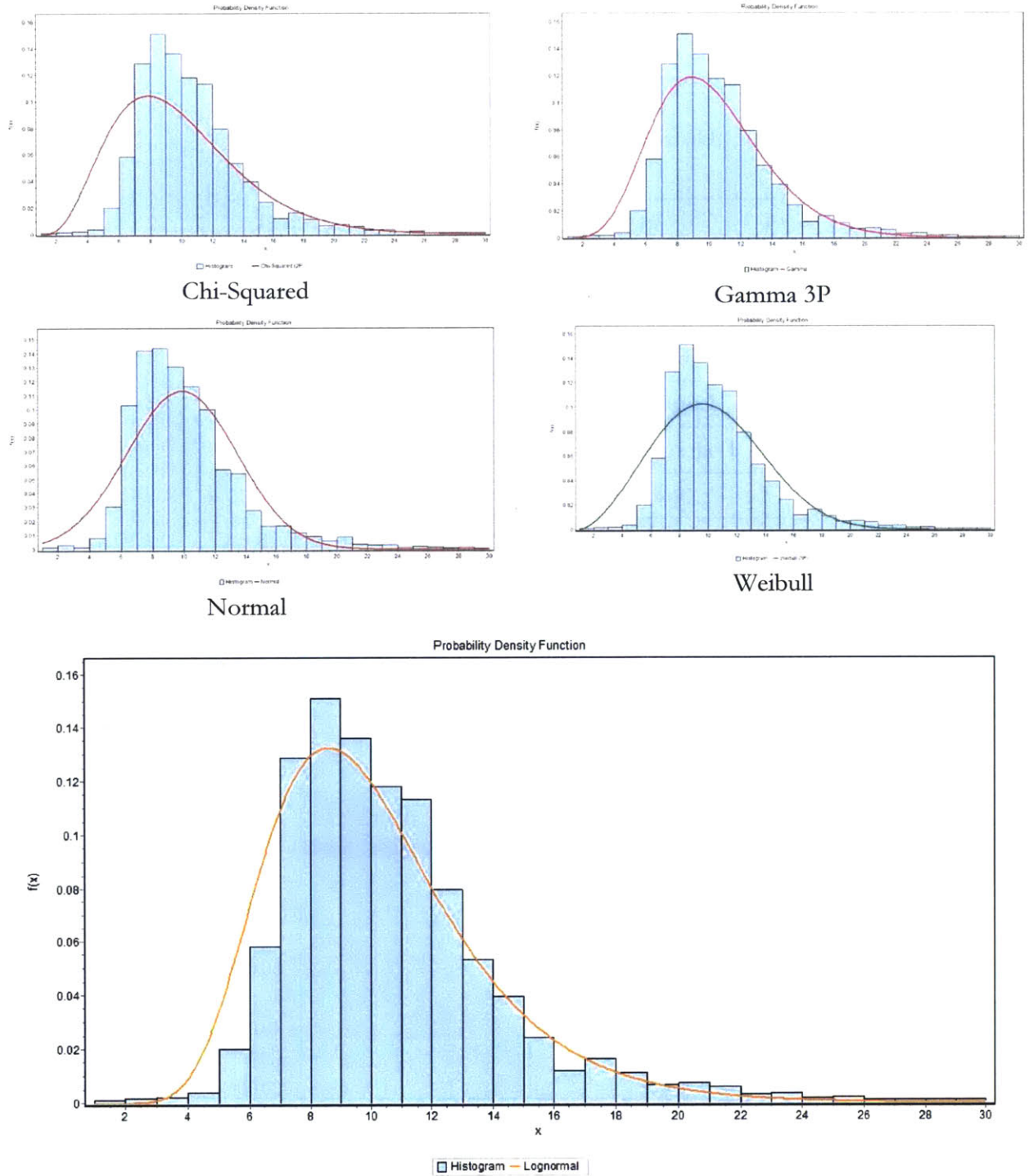
Figure 4.3.14. Results of the Goodness of Fit Tests of the density/DIM Factor of the RFI clients 01 & 03b

Table 4.3.12. Summary ranking of goodness of fit tests of density/DIM Factor dataset - RFI clients 01 & 03b

#	Distribution	Kolmogorov Smirnov		Anderson Darling	
		Statistic	Rank	Statistic	Rank
1	Chi-Squared	0.18883	9	17.487	1
2	Chi-Squared (2P)	0.18566	8	19.812	2
3	Gamma	0.10233	4	27.27	4
4	Gamma (3P)	0.09904	3	30.675	6
5	**Lognormal**	**0.08323**	**1**	**26.876**	**3**
6	Lognormal (3P)	0.08904	2	28.001	5
7	Normal	0.13848	7	40.302	9
8	Weibull	0.13398	6	33.679	8
9	Weibull (3P)	0.12928	5	33.046	7
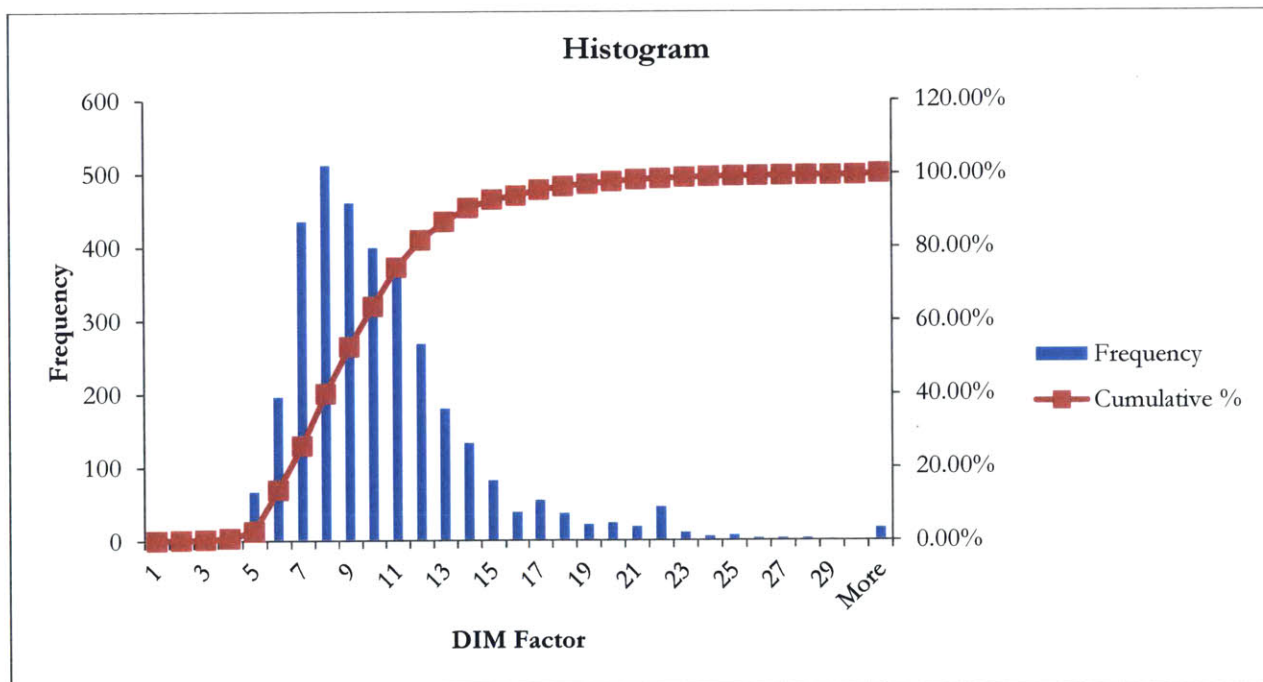
Figure 4.3.15 Histogram of the sample dataset for the density/DIM Factor - RFI clients 01 & 03b

Gross Weight simulation

The significant scarcity of data regarding the gross weight γ_w of air cargo for each bid substantially limits the usability and potential of both the analytical model and the visualization tools. The fact that the gross weight γ_w is an essential variable that affects both the consolidation processes and the profitability through the calculation of the Chargeable Weight for the Shipper σ and the Net Achieved Rate (NAR) drastically reduces the number of tenders under study.

Consequently, we chose to populate the spreadsheet cells concerning the gross weight of those bids that had all other variables available, via simulation of randomized numbers according to the distribution that bests fits the known gross weight values. As a matter of fact we were able to merge gross weight data from all nine bid customers to perform the goodness of fits tests and run the simulation.

The results of tests of goodness of fit presented as expected higher levels of goodness of fit regarding the Kolmogorov-Smirnov test rather than the Anderson-Darling test, since the latter emphasizes error for both tails. Thus, we opted for the distribution which presents a better adjustment of the sample values of the Gross Weight (merged Air Rates Customers C1 to C9) when applying the Kolmogorov-Smirnov goodness of fit test: the Gamma Distribution, and even better the Generalized Gamma Distribution.

$$\text{Generalized Gamma Distribution } f(x; a, d, p) = \frac{\frac{p}{a^d} \cdot x^{d-1}}{\Gamma(d/p)} e^{-(\frac{x}{a})^p},$$

with parameters a=0.38684 b=1906.8 g=100.0

Therefore, we selected the Generalized Gamma Distribution defined by the function shown above as the result of performing the goodness of fit tests on merged datasets C1 to C9 to simulate randomized numbers to populate the missing gross weight values on the Air Rates database. By doing so, we were able to have enough pieces of information regarding Chargeable Weights for both carrier/airline (α) and shippers (σ) as to perform analysis on the level of consolidation and profitability.

Table 4.3.13. Descriptive Statistics of goodness of fit tests of Gross Weight γ_w datasets - Bid clients 01 to 09

Statistic	Value		Percentile	Value
Sample Size	4196		Min	100
Range	4900		5%	100
Mean	1022.8		10%	100
Variance	1.5348E+6		25% (Q1)	100
Std. Deviation	1238.9		50% (Median)	500
Coef. of Variation	1.2112		75% (Q3)	1500
Std. Error	175.2		90%	3100
Skewness	1.4582		95%	3800
Excess Kurtosis	1.1562		Max	5000

Table 4.3.14. Adjusted Parameters of the statistical distributions provided by the goodness of fit tests

of Gross Weight γ_w datasets - Bid clients 01 to 09

#	Distribution	Parameters
1	Exponential	l=9.7770E-4
2	Exponential (2P)	l=0.00108 g=100.0
3	Gamma	a=0.68162 b=1500.6
4	**Gamma (3P)**	**a=0.38684 b=1906.8 g=100.0**
5	Lognormal	s=1.347 m=6.1091
6	Lognormal (3P)	s=6.6379 m=9.9011 g=100.0
7	Normal	s=1238.9 m=1022.8
8	Weibull	a=0.79519 b=890.55
9	Weibull (3P)	a=0.55448 b=628.87 g=100.0
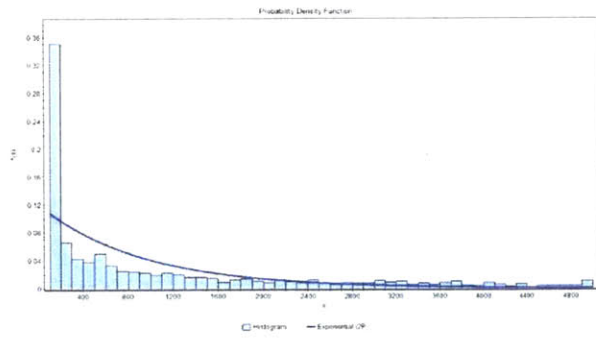
Exponential
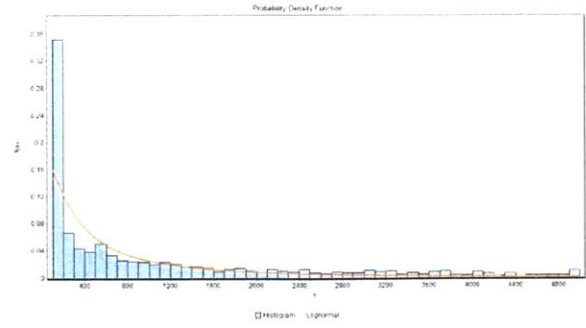
Lognormal
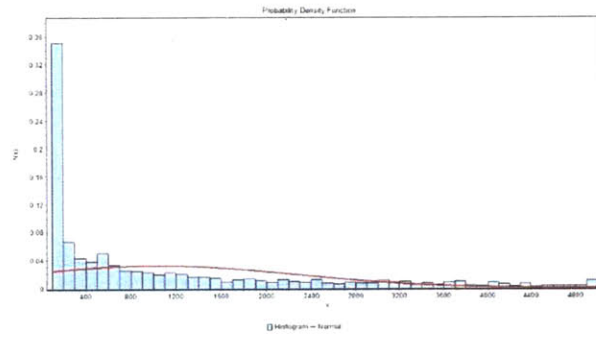
Normal
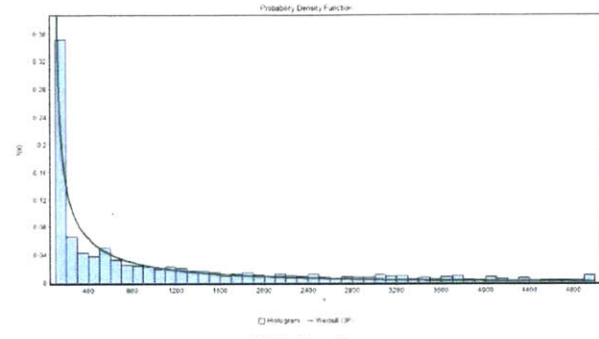
Weibull
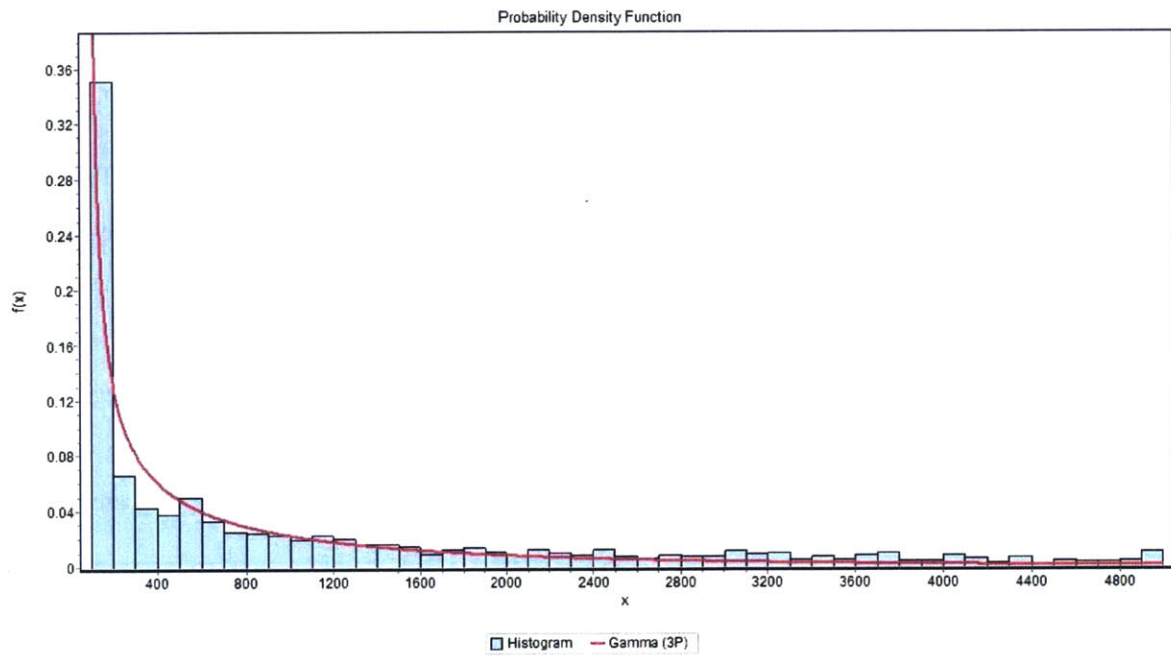
Gamma

Figure 4.3.16. Results of the Goodness of Fit Tests of the Gross Weight γ_w datasets - Bid clients 01 to 09

Table 4.3.15. Summary ranking of goodness of fit tests of Gross Weight γ_w datasets – Bids clients 01 to 09

#	Distribution	Kolmogorov Smirnov		Anderson Darling	
		Statistic	Rank	Statistic	Rank
1	Exponential	0.25814	6	38.48	7
2	Exponential (2P)	0.35129	8	44.62	8
3	Gamma	0.18154	2	34.715	3
4	**Gamma (3P)**	**0.09613**	**1**	**51.623**	**9**
5	Lognormal	0.21917	4	37.768	6
6	Lognormal (3P)	0.58376	9	20.447	1
7	Normal	0.22817	5	35.126	4
8	Weibull	0.19013	3	35.656	5
9	Weibull (3P)	0.35129	7	33.378	2

4.4.- Performance Metrics: current situation and business opportunities

In this section, a framework for performance metrics is presented (see Table 4.4.1) considering the two major objectives: provide a reference for better understanding of both the current business and the current bids with respect to air volume and density usage and profitability after consolidation. These metrics operate at the strategic level for the commercial strategy for a given lane as a whole, and are able to be combined in further research with metrics which apply specifically at an operational level, such as profitability per shipment, parameters to define the allocation/distribution of pallets along the aircraft, consolidation delta Δ per shipment, etc.

Measures are grouped in two main categories: consolidation performance metrics and profitability performance metrics. Consolidation performance metrics relate to physical properties which define compatibility between different loads to be transported via aircraft, whereas profitability performance metrics reflect measurements and references to calculate the level of absolute and percentage profits and provide with a series of different reference rates in relation to different scenarios to guide decision-making in the successive rounds of negotiations of the bidding process.

Table 4.4.1. Consolidation and profitability Performance Metrics

Metric	Definition	Formula & Units
	CONSOLIDATION PERFORMANCE METRICS	
C%	Consolidation percentage per lane	$C\% = 1 - \dfrac{\sum_{k=1}^{n} \alpha_i}{\sum_{k=1}^{n} \sigma_i}$ [%]
Δ	Consolidation Delta per lane	$\Delta = C\%(f) - C\%(i)$ [%]
	PROFITABILITY PERFORMANCE METRICS	
Γ	Level of aggressiveness in bidding	$\Gamma = \beta / \varrho$ [ø]
Π_λ	Profits per lane per year	[$/year]
$\Pi_\%$	Profitability per lane per year	[%]
β_0	Break-even rate for the whole lane	$\beta_0 = \{\beta \mid \Pi_\lambda = 0\}$ [$/kg]
β_β	Break-even rate for the bid	$\beta_\beta = \varrho$
β_π	Break-even rate for constant profits	$\beta_0 = \{\beta \mid \Pi_\%(year\,n+1) = \Pi_\%(year\,n)\}$ [$/kg]
β_λ	Constant profitability per lane rate	$\beta_\lambda = \{\beta \mid \Pi_\%(year\,n+1) = \Pi_\%(year\,n)\}$ [$/kg]

Consolidation percentage per lane (C%) is one minus the ratio between the aggregate carrier chargeable kg (α) and the aggregate shipper chargeable kg (σ), and describes the percentage of consolidated volumetric weight that is charged twice to the shippers while combining such loads in the same air space purchased in advance to the carrier.

Consolidation Delta per lane (Δ) is the incremental level of consolidation we acquire or reach by being awarded with the current bid.

The level of aggressiveness in bidding (Γ), defined as the ratio between the rate we charge to the shipper and the rate we pay to the carrier/airline, is a reference in assessing the tenders. The more aggressive the approach to address the bidding processes, the more likely the air freight forwarder will be awarded with the load. Nonetheless, air freight forwarders and 3PL companies in general need to establish a balance in the level of aggressiveness with which the bid is approached, since an excessive level of aggressiveness can damage profitability.

Profits per lane per year (Π_λ) describes the total amount of annual benefit by route.

Profitability per lane per year $\Pi_\%$ describes the percentage of annual benefit over revenue by origin-destination route.

Initial consolidation percentage per lane (C%(i)) is the consolidation percentage for the current business before being awarded with any bid under consideration.

Final consolidation percentage per lane (C%(f)) is the C% after being awarded with the current bid.

Break-even rate for the whole lane (β_0) is the rate that generates no global profit for a given o-d lane. This metric can be used by air freight forwarders as a lower bound or threshold not to cross in any case when addressing tenders.

Break-even rate for the bid (β_β) is the rate that guarantees break-even for each particular bid. The rate we charge to the shipper is the same as the rate we pay to the carrier/airline. In other words, it follows a neutral level of aggressiveness ($\Gamma = 1$) in bidding.

Break-even rate for constant profits (β_π) is a rate that guarantees the same profits in absolute numbers overall per lane. It can be used as a lower bound, threshold or minimum value to determine β.

Constant profitability per lane rate (β_λ) guarantees the same level of profitability as the current business. Higher rates, if awarded, would bring not only more profits in absolute terms to the air freight forwarder, but also a level of profitability higher than the current one for a given lane.

Those three metrics are created to define reference values for guiding decision-making when addressing tenders and estimating the most profitable, yet still competitive, rate β. Such metrics will define three thresholds or milestones that will create four regions of values for β rates. For each bid under consideration we define three thresholds for rates:

Threshold 1: Break-even for the whole portfolio (zero profits overall)

Threshold 2: Break-even for the bid (same profits overall)

Threshold 3: Constant profitability threshold (we keep the same % of Π)

Thus, the four profitability regions can be defined as:

Area A: $\beta > \beta_\lambda$ (desirable: more revenue, more profits, higher profitability)

Area B: $\beta_\pi > \beta < \beta_\lambda$ (acceptable: more revenue, more profits, less profitability in %)

Area C: $\beta_0 > \beta < \beta_\pi$ (more revenue, less profits)

Area D: $\beta < \beta_0$ (losses to the company)

Those profitability regions graphically and conceptually depicted in Figure 4.4.1 will provide the senior management with solid references to address the tenders and assess the However, further research involving the implementation of an optimization model based on Lagrangian-relaxation heuristics at the operational level would also allow for performing sensitivity analysis with respect to the level of aggressiveness. If on top of the current analytical model we overlap a probabilistic study with regard to the chances of winning a tender in terms of the rate proposed, we would be able to refine more the sensitivity analysis when choosing a rate that provides the greatest benefits, and in turn suppose a competitive value.
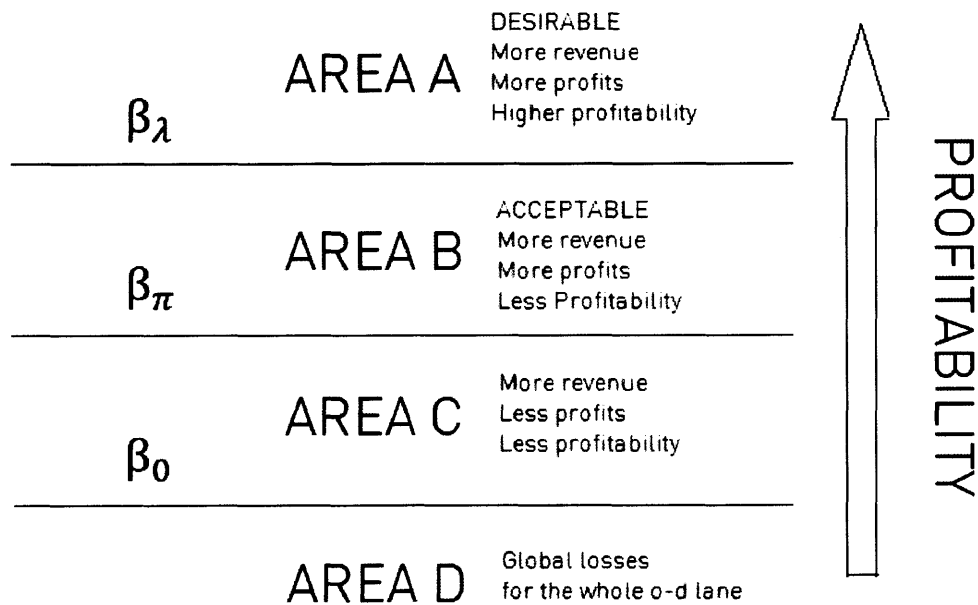
Figure 4.4.1. Profitability regions defined by the Profitability performance metrics

How to build the queries

With the data imported into SQL Server we then wrote a query to take the original data and calculate the metrics we wanted. Due to some limitations in SQL Server this had to be done with a series of derived tables. Essentially a derived table is a table that is created within a query and only exists within that query. The derived tables allowed us to calculate sums and other aggregate values, and then reference those values later in the query. In this section we explain selected parts of the query we wrote in order to create an understanding for the process and the underlying structure, we do not however think it is useful to explain every single line of code so we have strived to only explain the key parts.

```
SELECT Origin, Destination, Origin+'-'+Destination AS [OriDes], Customer, id, [Total cost], [min bid], [kg_per_year], [gross weight], [Vol weight],
(SELECT dbo.InlineMax(SUM([gross weight]), SUM([Vol weight]))/SUM([kg_per_year]) FROM Rates) AS [Consolidation],
[kg_per_year] * (SELECT dbo.InlineMax(SUM([gross weight]), SUM([Vol weight]))/SUM([kg_per_year]) FROM Rates) [carrierchkg],
[kg_per_year] * [min bid] AS [revenue],
[Total cost] * [kg_per_year] * (SELECT dbo.InlineMax(SUM([gross weight]), SUM([Vol weight]))/SUM([kg_per_year]) FROM Rates) AS [cost],
([vol weight]*(6)/1000) AS [m3]
FROM Rates
WHERE
(Origin = 'PVG' OR Origin = 'BJS' OR Origin = 'PEK' OR Origin = 'HKG')
AND (Destination = 'ORD' OR Destination = 'JFK' OR Destination = 'LAX')
AND [Min bid] > 0 AND [awarded kg] > 0
```

The first derived table calculates a number of aggregated values for the business that was awarded to the company. This will be used as the current business and then other bids that were not awarded will be compared to these aggregates to see how they would impact the business. The values calculated in this derived table is first the percentage of non consolidated kg which is found by first taking whichever is larger of the sum of the gross weight and the sum of the volumetric weight, or in our nomenclature α (the carrier chargeable kg). α is then divided by the sum of the chargeable kg from the shippers (σ), this ratio is subtracted from one to obtain the consolidation percentage. The ratio of non-consolidated kg is then multiplied with σ for individual piece of

business to get the separate α. The final calculation in this table is to multiply α with ϱ (the rate charged by the airline) to get the cost for every row.

```sql
SELECT
OriDes, Origin, Destination,
COUNT(id) AS [#],
SUM(revenue) AS [revenue],
SUM(cost) AS [cost],
SUM(revenue)-SUM(cost) AS [profit],
(SUM(revenue)-SUM(cost))/SUM(revenue) AS [margin],
SUM(kg_per_year) AS [shipperchkg],
SUM(carrierchkg) AS [carrierchkg],
SUM([gross weight]) AS [totgross],
SUM([Vol weight]) AS [totvol],
1-dbo.InlineMax(SUM([gross weight]), SUM([Vol weight]))/SUM([kg_per_year]) AS [cons],
SUM(cost)/SUM(kg_per_year) AS [NAR],
(1000/(167*(SUM([gross weight])/SUM([vol weight])))) AS [1:],
(SUM(revenue)-SUM(cost))/SUM(m3) AS [profit/m3],
SUM(kg_per_year * [total cost])/SUM(kg_per_year) AS [WeightedAvgRate], SUM([vol weight]) AS
[VW], SUM([gross weight]) AS [GW],
(CASE WHEN SUM([vol weight])-SUM([gross weight])>0 THEN SUM([vol weight])-SUM([gross
weight])
WHEN SUM([vol weight])-SUM([gross weight])<=0 THEN 0 END) AS [freeGW],
(CASE WHEN SUM([gross weight])-SUM([vol weight])>0 THEN SUM([gross weight])-SUM([vol
weight])
WHEN  SUM([gross weight])-SUM([vol weight])<=0 THEN 0 END) AS [freeVW],
SUM([gross weight])/(dbo.InlineMax(SUM([gross weight]), SUM([Vol weight]))) AS [grosseff],
SUM([vol weight])/(dbo.InlineMax(SUM([gross weight]), SUM([Vol weight]))) AS [voleff]
```

The next derived table calculates the aggregated values for the portfolio of current business. A key metric calculated here is the profit, which is found by using the calculated values for revenue and cost for each row from the previous derived table and subtracting the sum of costs from the sum of revenues for each origin destination pair. Other key values that are calculated here and will be used in the next step is the amount of unused volume and unused gross weight. Per definition one of these two values will be zero (most likely the amount of unused volume). For the amount of unused gross weight this is done with a logical statement that either returns the difference between volumetric kg and gross weight if the difference is non-negative, or zero if the difference is negative. The amount of unused volumetric kg is calculated the same way, only inversed.

```
--This is a derived table that creates the column [OriDes] which allows it to be joined to the
portfolio aggregates below,
--it also needs contain the other columns needed for the query above.
FROM (SELECT Origin, Destination, Origin+'-'+Destination AS [OriDes], Customer, id, [Total cost],
[min bid], [kg_per_year], [gross weight], [Vol weight],
(SELECT dbo.InlineMax(SUM([gross weight]), SUM([Vol weight]))/SUM([kg_per_year]) FROM
Rates) AS [Consolidation],
[kg_per_year] * (SELECT dbo.InlineMax(SUM([gross weight]), SUM([Vol
weight]))/SUM([kg_per_year]) FROM Rates) [carrierchkg],
[kg_per_year] * [min bid] AS [revenue],
[Total cost] * [kg_per_year] * (SELECT dbo.InlineMax(SUM([gross weight]), SUM([Vol
weight]))/SUM([kg_per_year]) FROM Rates) AS [cost],
([vol weight]*(6)/1000) AS [m3]
FROM Rates
WHERE
(Origin = 'PVG' OR Origin = 'BJS' OR Origin = 'PEK' OR Origin = 'HKG')
AND (Destination = 'ORD' OR Destination = 'JFK' OR Destination = 'LAX')
AND [Total cost] > 0 AND [awarded kg] = 0) B
```

The third derived table creates the column [OriDes], which is just the origin destination pair as one value, for every row that contains a bid under consideration. The reason for this is to allow a join between a piece of business under consideration on a certain origin-destination pair and compare that bid to the current business for that origin-destination pair.

The final and longest part of the query is the part that actually creates the output result by using the derived tables. It calculates the change in consolidation, [consDelta], for each bid under consideration by adding the values from that bid to the aggregated values from the portfolio for the current business and thereby getting a new consolidation value. The actual consolidation is then subtracted from the new value to get the change in consolidation in percentage points.

Another key value calculated is how much of the bid that can be consolidated or in other words ride for free. This is done by finding the extra volume and gross weight, where extra is referring to how much more volumetric/gross weight there is compared to the other one. Again, per definition one of the two has to be zero. For example, if a shipment consists of 8 volumetric kg and 10 gross kg there is two extra gross kg that can potentially be consolidated if there is spare gross kg capacity on the route.

```
SELECT B.OriDes, Origin_Airport.CityName AS [OriCity], Destination_Airport.CityName AS
[DesCity],
Origin_Airport.Lat AS [OriLat], Origin_Airport.Lon AS OriLon, Destination_Airport.Lat AS DesLat,
Destination_Airport.Long AS [DesLon],
B.customer, B.id, B.[Total Cost], B.[kg_per_year] AS [shipperchkg], ROUND(B.[gross weight],0)
AS [gross weight],
ROUND(B.[Vol weight],0) AS [vol weight],
ROUND(1000/(167*(B.[gross weight]/B.[Vol weight])),2) AS [lane 1:],
ROUND(1000/(167*((AP.GW+B.[gross weight])/(AP.VW+B.[Vol weight]))),2) AS [portfolio 1:],
(1-(dbo.InlineMax(AP.totgross+B.[gross weight], AP.totvol+[Vol
weight])/(AP.shipperchkg+B.[kg_per_year]))-AP.cons) AS [consDelta],
ROUND((CASE WHEN B.[gross weight]>B.[Vol weight] THEN B.[gross weight]-B.[Vol weight]
WHEN B.[gross weight]<=B.[Vol weight] THEN 0 END),0) AS [extragross],
ROUND((CASE WHEN B.[gross weight]<B.[Vol weight] THEN B.[Vol weight]-B.[gross weight]
WHEN B.[gross weight]>=B.[Vol weight] THEN 0 END),0) AS [extravol],
```

In the next step the extra kg is compared to the excess capacity on the route. The extra vol/gross kg that is covered by matching excess capacity can be consolidated and will be listed as general extra kg. This consolidated cargo will essentially ride for free and not be taken into account when the cost of the bid is calculated.

```
--This gives the extra field by using the extravol and extragross
ROUND((
(CASE WHEN
(CASE WHEN B.[gross weight]>B.[Vol weight] THEN B.[gross weight]-B.[Vol weight] WHEN
B.[gross weight]<=B.[Vol weight] THEN 0 END)
<AP.[freeGW]
THEN (CASE WHEN B.[gross weight]>B.[Vol weight] THEN B.[gross weight]-B.[Vol weight] WHEN
B.[gross weight]<=B.[Vol weight] THEN 0 END)
WHEN (CASE WHEN B.[gross weight]>B.[Vol weight] THEN B.[gross weight]-B.[Vol weight]
WHEN B.[gross weight]<=B.[Vol weight] THEN 0 END)
>=AP.[freeGW] THEN AP.[freeGW] END)
+ (CASE WHEN
(CASE WHEN B.[Vol weight]>B.[gross weight] THEN B.[Vol weight]-B.[gross weight] WHEN B.[Vol
weight]<=B.[gross weight] THEN 0 END)
<AP.[freeVW]
THEN (CASE WHEN B.[Vol weight]>B.[gross weight] THEN B.[Vol weight]-B.[gross weight] WHEN
B.[Vol weight]<=B.[gross weight] THEN 0 END)
WHEN (CASE WHEN B.[Vol weight]>B.[gross weight] THEN B.[Vol weight]-B.[gross weight]
WHEN B.[Vol weight]<=B.[gross weight] THEN 0 END)
>=AP.[freeVW] THEN AP.[freeVW] END)
),0)
AS [extra],
```

Finally we calculate the β (rate) needed to maintain the same profit margin on the route. This was done by solving the following equation for β:

$$\mu = profit\ margin;\ \varepsilon = kg\ that\ are\ consolidated$$

$$\mu = 1 - \frac{\beta\sigma - (\sigma - \varepsilon)\rho}{\beta\sigma} \Leftrightarrow \beta = \frac{(\sigma - \varepsilon)\rho}{(1-\mu)\sigma}$$

When the solved equation is then written in SQL code it looks like this:

```
--cPM
ROUND((((B.[kg_per_year]
-(
(CASE WHEN
(CASE WHEN B.[gross weight]>B.[Vol weight] THEN B.[gross weight]-B.[Vol weight] WHEN
B.[gross weight]<=B.[Vol weight] THEN 0 END)
<AP.[freeGW]
THEN (CASE WHEN B.[gross weight]>B.[Vol weight] THEN B.[gross weight]-B.[Vol weight] WHEN
B.[gross weight]<=B.[Vol weight] THEN 0 END)
WHEN (CASE WHEN B.[gross weight]>B.[Vol weight] THEN B.[gross weight]-B.[Vol weight]
WHEN B.[gross weight]<=B.[Vol weight] THEN 0 END)
>=AP.[freeGW] THEN AP.[freeGW] END)
+
(CASE WHEN
(CASE WHEN B.[Vol weight]>B.[gross weight] THEN B.[Vol weight]-B.[gross weight] WHEN B.[Vol
weight]<=B.[gross weight] THEN 0 END)
<AP.[freeVW]
THEN (CASE WHEN B.[Vol weight]>B.[gross weight] THEN B.[Vol weight]-B.[gross weight] WHEN
B.[Vol weight]<=B.[gross weight] THEN 0 END)
WHEN (CASE WHEN B.[Vol weight]>B.[gross weight] THEN B.[Vol weight]-B.[gross weight]
WHEN B.[Vol weight]<=B.[gross weight] THEN 0 END)
>=AP.[freeVW] THEN AP.[freeVW] END)
))*B.[Total Cost])/((1-AP.[margin])*B.[kg_per_year]),2) AS [cPM]
```

4.5.- Designing the visualization tools

We considered how to assess the attractiveness of a bid under consideration, and present it in an intuitive and concise way through displaying all the key variables and metrics involved, such as size, density and consolidation potential within the same interactive graph. With the query complete we used it to create a View, which is something that in the database basically behaves like a table, but

68

instead of containing raw data it is based on a query. The difference between a view and a derived table is that the view is saved as an entity in the database and it can therefore be accessed by external programs, unlike the derived table that can only be referenced within the query where it is created.

After creating the View, connecting it to Tableau to create the visualization is a very straightforward process. Tableau has built in support for accessing SQL Server and reads the View like it would any other table. So the data from the query can then be accessed in Tableau. This is a live connection to the database, so if something is added to the database it will show up in Tableau upon refresh.

In Tableau we created a two different scatter plots that show a number of different qualities for each bid under consideration. The first one has the change in consolidation on the vertical axis and on the horizontal axis it shows what the density ratio for the portfolio will be with the bid added. The size of the bid in chargeable weight is shown in the graph by the size of the dot and each route has its own shape. We also use color to indicate how dense or light a piece of business is, green is for dense business and red for light. The darker the color the denser/lighter it is. A second scatter plot shows the change in consolidation on the vertical axis and the size of the bid on the horizontal. Colors are used to distinguish between the different routes.

We also created a dashboard that has either of the scatter plots as the main window. Then it has two maps, one for origins and one for destinations, these maps are used as filters to be able to click the map in order to select which routes to look at. The dashboard also contains a number of sliders and selectors to be able to filter the content of the scatter plots further.

5.- Results

5.1.- Depiction of data available

The graph and table below (Table 5.1.1 and Figure 5.1.1) show the information available in the Air Rates data set. In total it contained over 25 000 rows. 98% of these rows contained information about the cost (ϱ) for the lane. Information on the amount of chargeable weight for the lane is not as readily available, 76% of the rows lack this information. The numbers are similar if we look at target data. Target data (τ) is data on what the shipper suggested that the freight forwarder should bid in order to be successful. However if count how many rows that have both chargeable weight and target information we only find that in 12% of the rows.

Table 5.1.1. Data available in datasets

Total rows	25926
Total cost > 0	25537
Kg > 0	6162
Max target > 0	6485
Target and kg > 0	2983
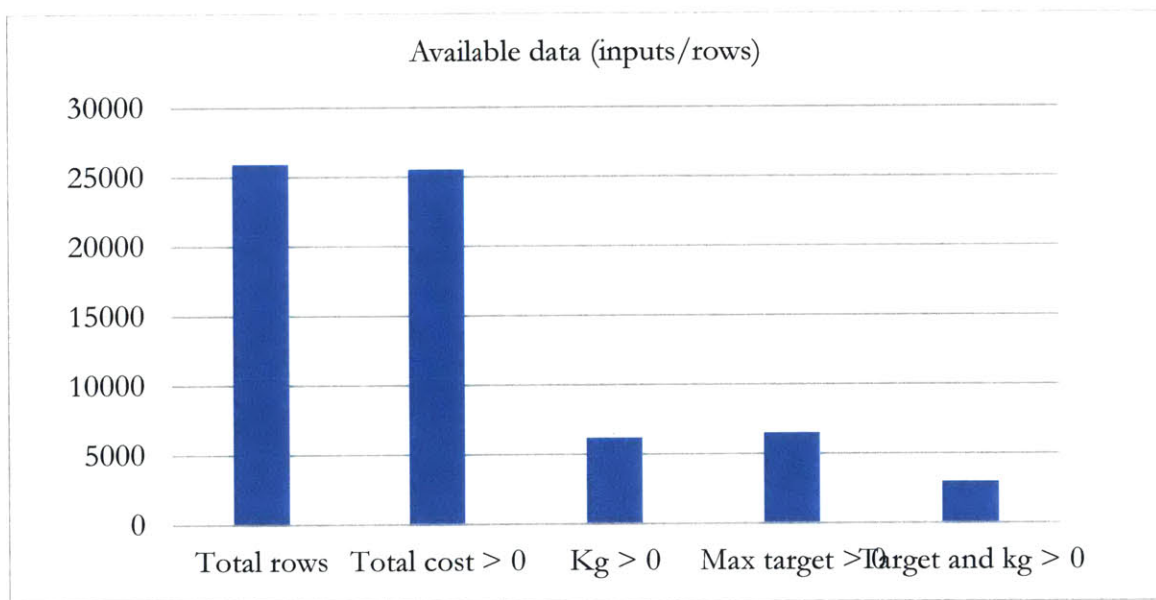
Figure 5.1.1. Data available in datasets

70

If we restrict the data to the nine routes that are within the scope given by the sponsor we see a similar story. Almost all of the bids have information on ϱ, but drastically fewer contain the other two pieces of information we are counting here. When we narrow down the scope to the nine lanes we are focusing on and only count the ones that have information on chargeable weight only 342 out of over 25 000 rows, 1.3%, remain.

Figure 5.1.2. Data available in sponsor company datasets regarding the 9 focus routes

Table 5.1.2. Data available in sponsor company datasets regarding the 9 focus routes

9 routes	1478
9 Total cost > 0	1462
9 Kg > 0	342
9 Max target > 0	224
9 Target and kg > 0	172

With respect to RFI01 and RFI03, the lack of data on airline/carrier rates ϱ and target rates prevents us from developing further analysis on profitability after consolidation for a given current bid. Nevertheless, a study of data on physical properties using Tableau visualization allows us to discriminate intuitively what markets present major business opportunities. For example, as shown in Figure 5.1.3, for customer RFI01 the major destination gateway airports in terms of potential gross weight of goods transported correspond to those of Vietnam, Hong Kong, Indonesia and China. Also, according to Figure 5.1.4, the main destination gateways for aggregate gross weight are located in the United States, Germany and to a lesser extent Russia and Canada.

Map based on Longitude (generated) and Latitude (generated). Size shows sum of Sum of Gross Weight (KGS). The marks are labeled by sum of Sum of Gross Weight (KGS). Details are shown for Origin Country.
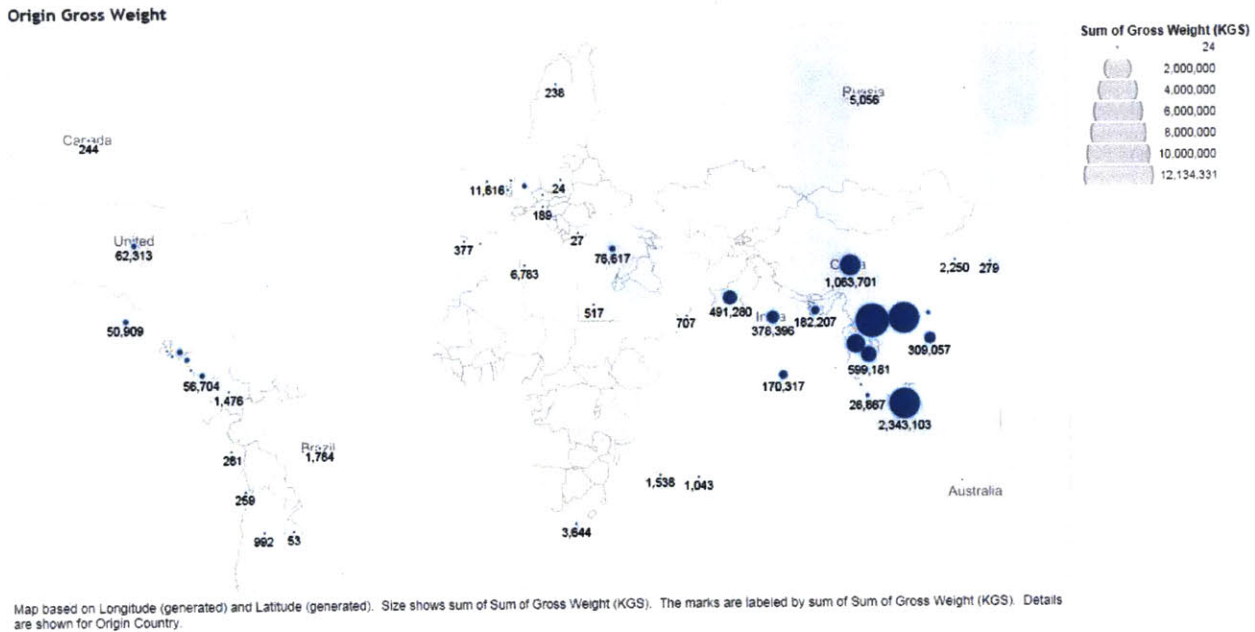
Figure 5.1.3. Distribution of business opportunities by origin with respect to gross weight - RFI01

As for density of cargo transported by air, we can perform market segmentation by country, since we can observe significant marked differences with respect to the average values for density of shipment δ. For the customer RFI01, for example, the average density per cubic meter of products in Argentina is of 426 kg per CBM. Products transported with origin airport gateway in China, however, have an average density of 107.6 kg per CBM, which represents a 9.34 DIM factor, which is relatively light with respect to the 167 kg per CBM reference value for 1:6.

Destination GW of shipments
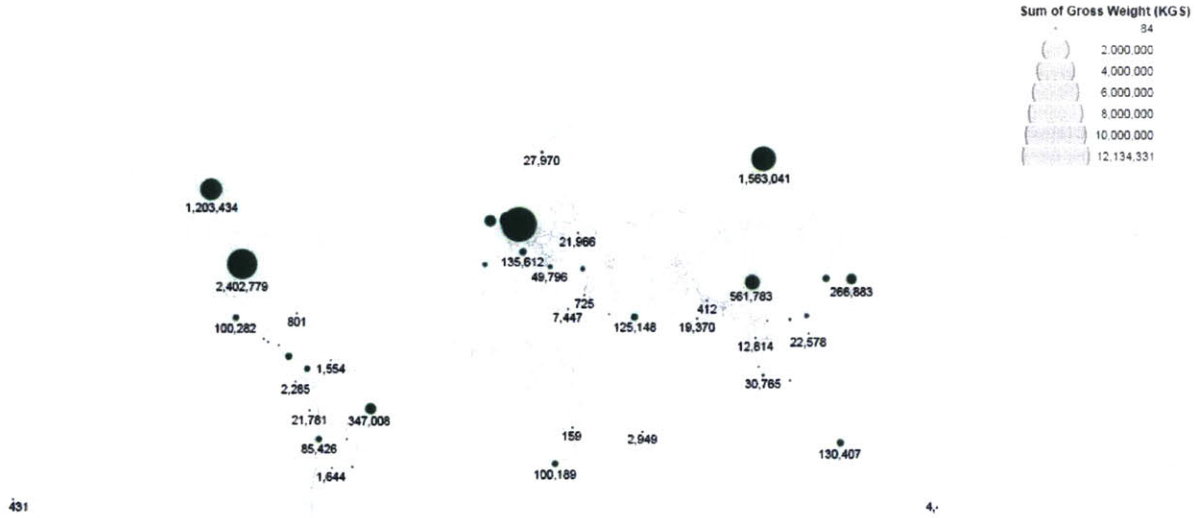
Sum of Gross Weight (KGS)
- · 84
- 2,000,000
- 4,000,000
- 6,000,000
- 8,000,000
- 10,000,000
- 12,134,331

27,970

1,563,041

1,203,434

21,966

135,612
49,796

2,402,779

561,783   266,883

725
7,447   125,148 19,370   412

100,282   801

12,814   22,578

• 1,554

2,285   30,785

21,781   347,008

159   2,949

85,426   130,407

1,644   100,189

431   4,·

Map based on Longitude (generated) and Latitude (generated). Size shows sum of Sum of Gross Weight (KGS). The marks are labeled by sum of Sum of Gross Weight (KGS) Details are shown for Dest Country.

Figure 5.1.4. Distribution of business opportunities by destination with respect to gross weight - RFI01

Origin Density of shipments

Average of WV
27.4 ──────── 511.4

59.3   105.8

177.4
75.9

148.2
153.2   102.2   107.6   118.7

108.8   128.9
544.4   142.3 130   100.5 104.9
139.5
123.4
258.9   233

84.5

143.8   117.6

153.3

134.1   107.1   137.3

11   152.5

100.3
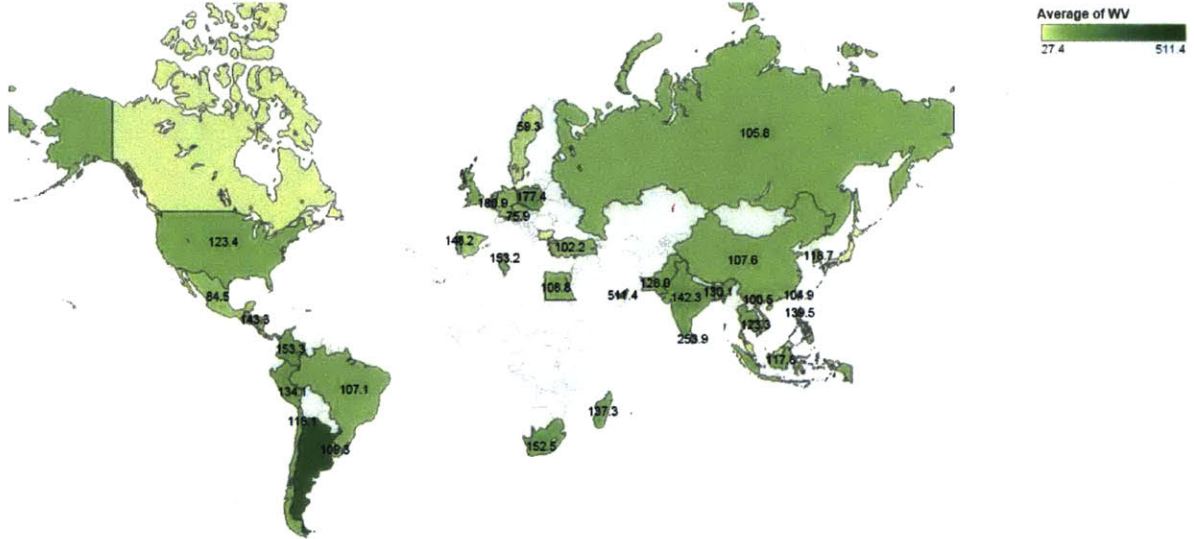
Map based on Longitude (generated) and Latitude (generated). Color shows sum of Average of WV. The marks are labeled by sum of Average of WV. Details are shown for Origin Country.

Figure 5.1.5. Origin density δ of shipments

5.2.- Creation of the Database

Figure 5.2.1 shows a query that returns the most relevant data from the database. It contains 13 different data columns and holds all the data that is necessary for creating the metrics we use in the visualization. So it is displaying the data that needs to be collected and inputted in order for the calculating query and hence the visualization to work. A few of the weight to volume measurements are technically redundant, since only two of them are needed to calculate the other two. A list of the columns in this query is found in Table 5.2.1.

Figure 5.2.1. Screenshot of the database build upon the Air Rates datasets populated with datasets from RFI01 and RFI03.

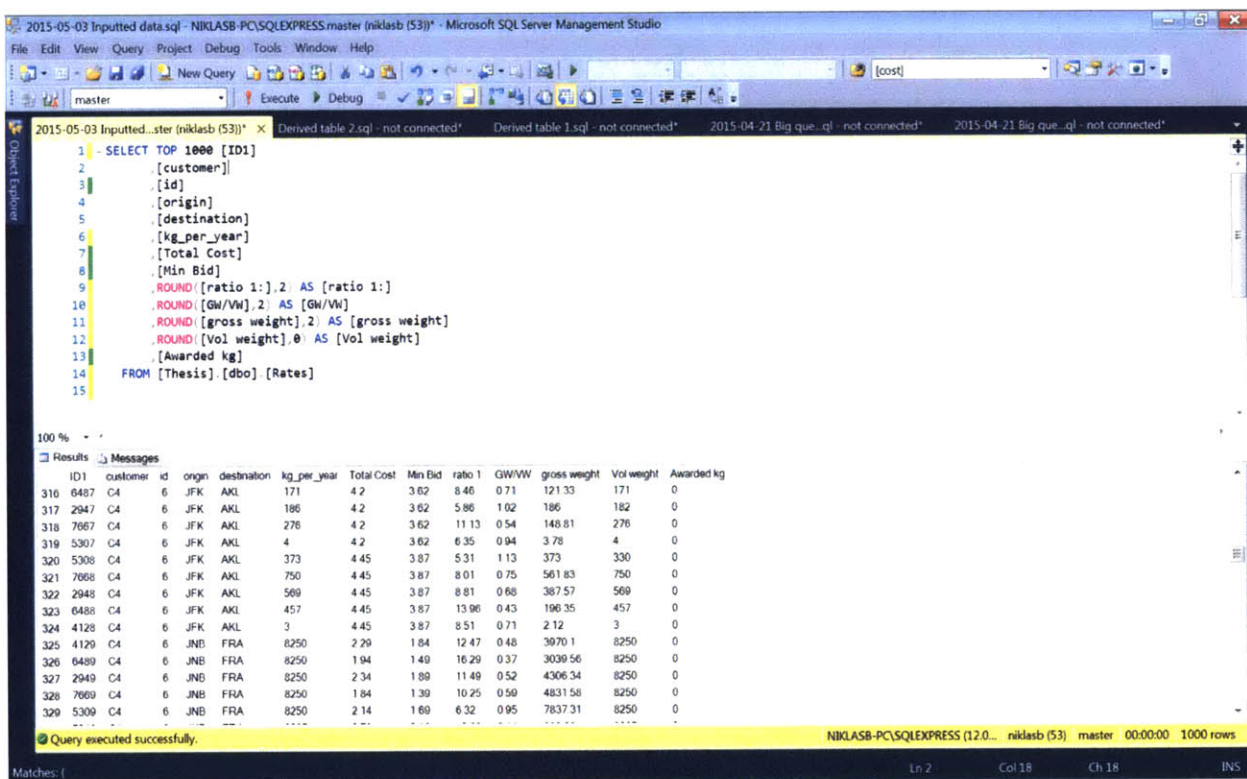Table 5.2.1. Key data fields with relevant variables to be used as inputs in the analytical model embedded in SQL to calculate meaningful metrics

Code	Description
ID1	Database identifier
customer	Customer code
id	Lane id code
origin	Origin airport
destination	Destination airport
kg_per_year	Annual chargeable kg
Total Cost	Cost per kg
Min Bid	Lowest bid submitted by sponsor
ratio 1:	Cubic meters per metric ton
GW/VW	Gross weight divided by volumetric weight
gross weight	Annual gross weight kg
Vol weight	Annual volumetric kg
Awarded kg	Annual awarded chargable kg

5.3.- Data simulation results

The results of the goodness of fit tests to analyze the suitability of a series of statistical distributions to simulate the values for the density to populate the DIM Factor values of the database was to select the Lognormal distribution with parameters σ=0.33116 μ=2.2633 as shown in Figures 5.3.1 to 5.3.3.

$$f(x; \mu = 2.2633, \sigma = 0.33116) = \frac{1}{x \cdot 0.33116 \ \sqrt{2\Pi}} \cdot e^{\frac{-(\ln(x)-2.2633)^2}{20.33116\ ^2}}$$

Figure 5.3.1. Lognormal distribution chosen to simulate and randomize density values to populate the incoming bids datasets
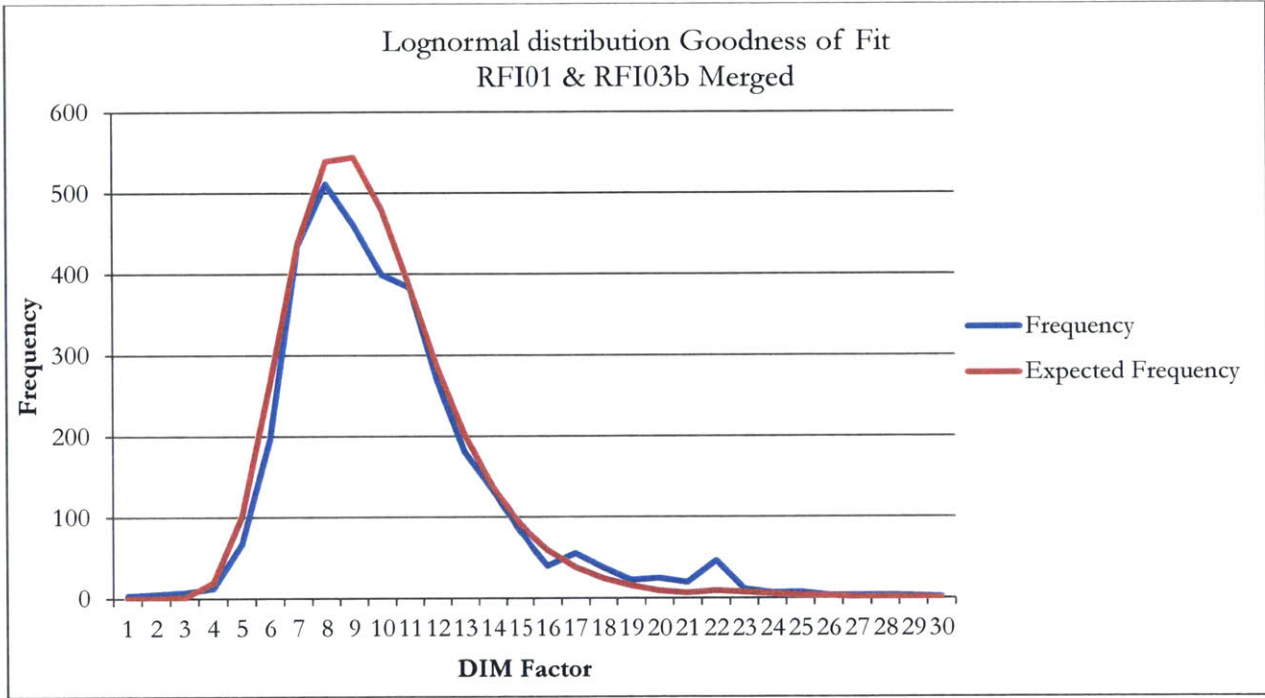
Figure 5.3.2. Comparison between RFI01 and Expected Lognormal - Frequency Distribution for the density/DIM Factor - RFI clients 01 & 03b
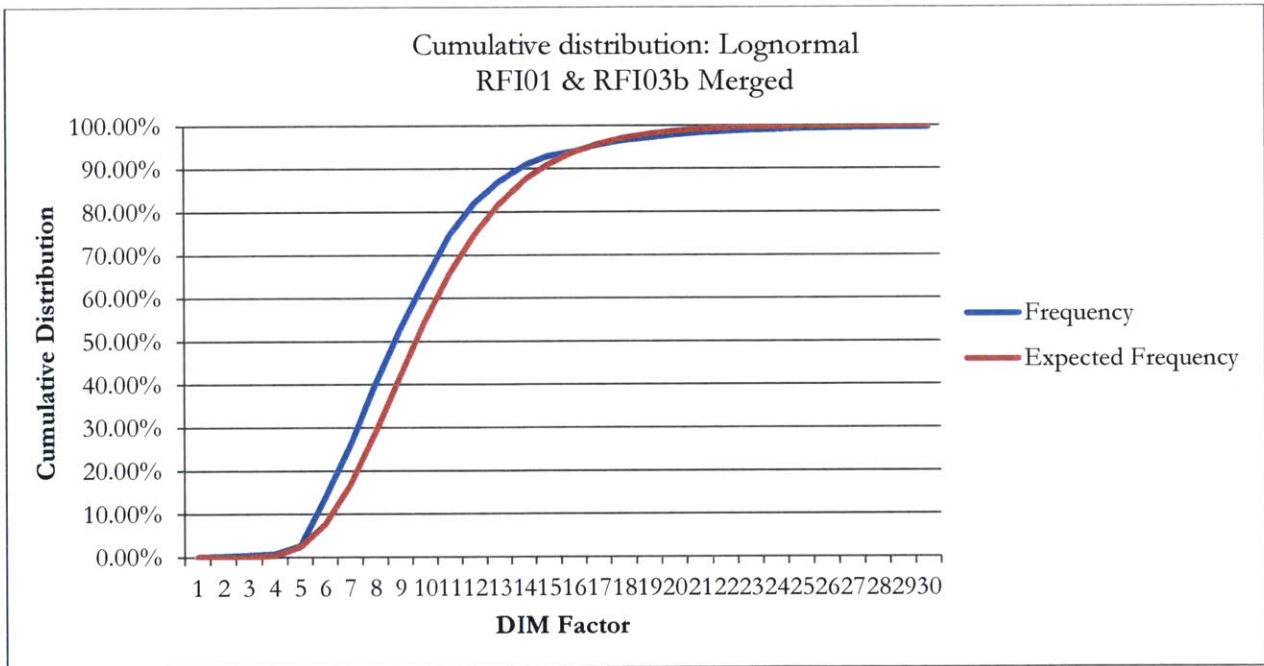
Figure 5.3.3. Comparison between cumulative distributions for real and expected Frequency Lognormal - Distribution for the density/DIM Factor - RFI clients 01 & 03b

As a result of the goodness of fit tests to analyze the suitability of a series of statistical distributions to simulate the values for the gross weight to populate the database we opted for the distribution which presents a better adjustment of the sample values of the Gross Weight (merged customers 01 to 09) when applying the Kolmogorov-Smirnov goodness of fit test: the Generalized Gamma Distribution with parameters a=0.38684 b=1906.8 g=100.0.

$$\text{Generalized Gamma Distribution} \quad f(x; a, d, p) = \frac{\frac{p}{a^d} \cdot x^{d-1}}{\Gamma(d/p)} e^{-\left(\frac{x}{a}\right)^p}$$

Figure 5.3.4. Generalized Gamma distribution chosen to simulate and randomize gross weight values to populate the incoming bids datasets
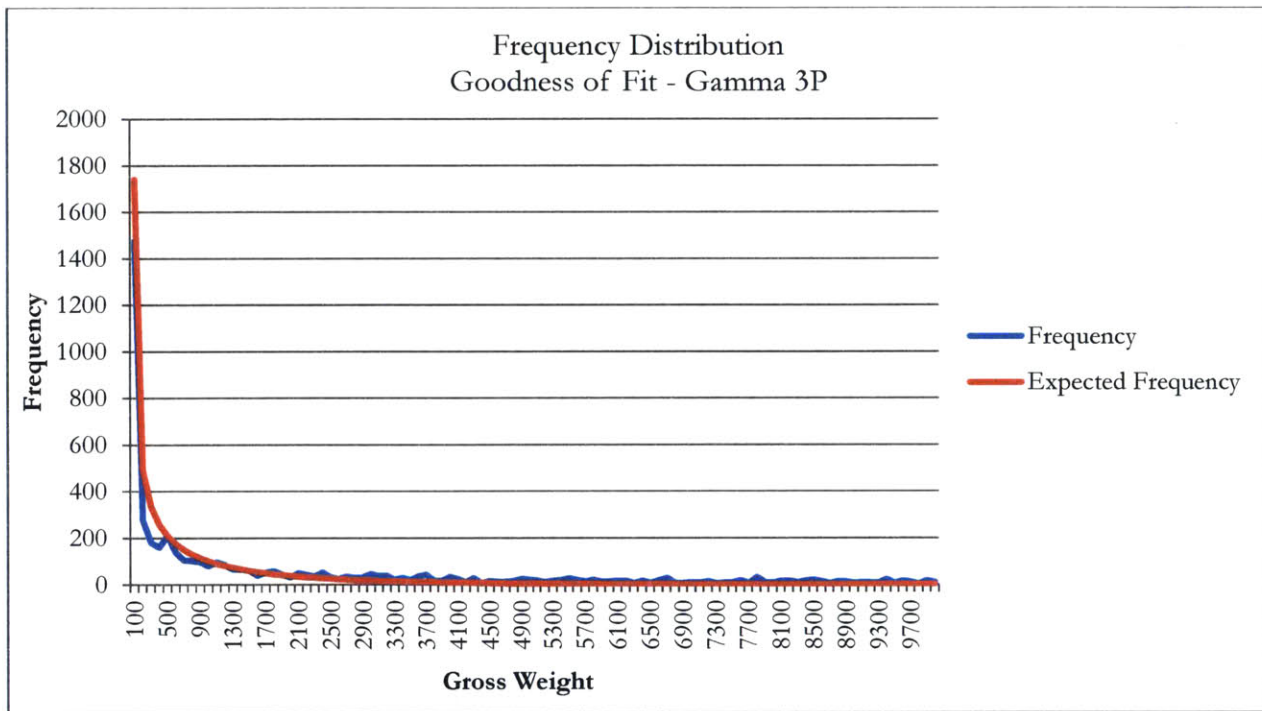
Figure 5.3.5. Comparison between real and expected Lognormal - frequency distribution for the Gross Weight γ_w datasets – Bids clients 01 to 09

5.4.- Analytical Model - forecasting tool

The result of the calculating query is shown below (figure 5.4.1), the columns shown there are the ones that go into the view that is used to generate the visualization in Tableau. All the values in this output table are derived from the data inputted in the previous section. Figure 5.4.b outlines all the columns that are accessible to Tableau through this query.
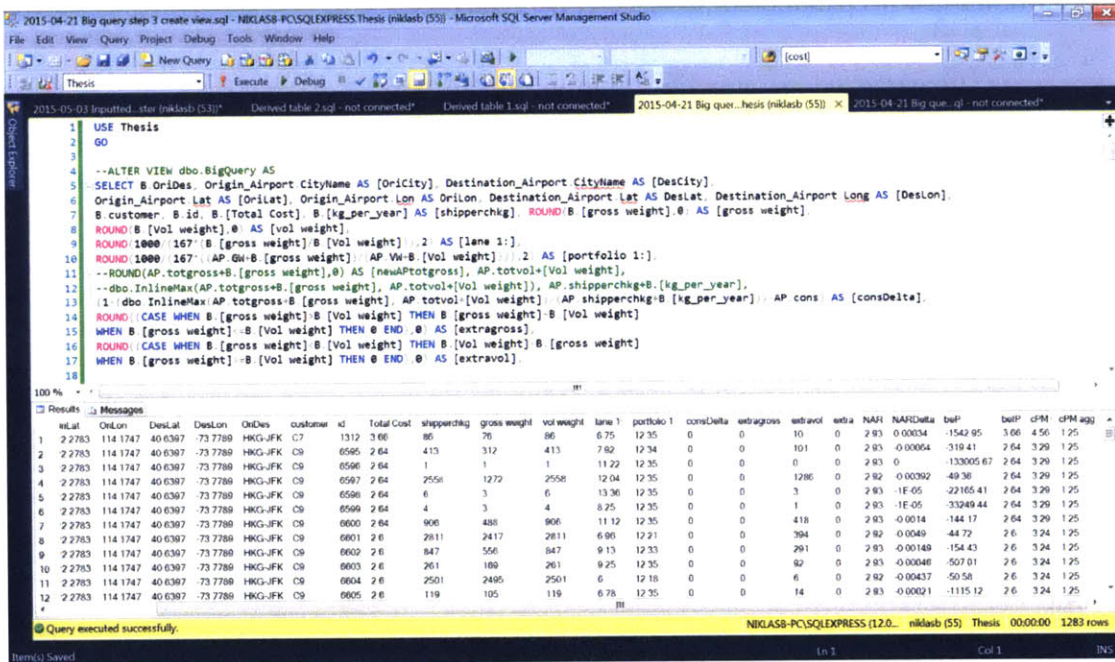
Figure 5.4.1. Result table for the main query executed in SQL

Table 5.4.1. Result columns from the query presenting meaningful variables and Performance Metrics

OriCity	Origin City
DesCity	Destination City
OriLat	Origin Latitude
OriLon	Origin Longitude
DesLat	Destination Latitude
DesLon	Destination Longitude
OriDes	Origin-Destination pair
customer	Customer ID
id	Lane ID
Total Cost	Cost $/kg

78

shipperchkg	Annual chargeable kg charged to the shipper
gross weight	Annual gross weight kg
vol weight	Annual volumetric kg
lane 1:	Cubic meters per metric ton for this bid
portfolio 1:	Cubic meters per metric ton for the portfolio if this bid is added
consDelta	Change in consolidation ratio of the portfolio if this bid is added
extragross	Number gross kg that can be consolidated
extravol	Number of volumetric kg that can be consolidated
extra	Number of shipper chargeable kg that can be consolidated
NAR	Net Achived Rate for portfolio if this bid is added
NARDelta	Change in the NAR of the portfolio if this bid is added
beP	Rate in $ at which the portfolio will have 0 profits
beIP	Rate in $ at which the portfolio profit wont change if this bid is added
cPM	Rate in $ at which the operating profit margin for the portfolio wont change if this bid is added
cPM agg	Estimation of aggressiveness potetnial: cPM divided by Total Cost. Lower number means more potential to be aggressive.

5.5.- Visualization tools

The Figure 5.5.1 presents the first dashboard with scatter plot showing the change in consolidation on the vertical axis and the resulting portfolio density ratio if the bid is added to the portfolio of current business. Also note that Hong Kong and Los Angeles are selected on the maps, so only bids on this route is shown in the plot.

Figure 5.5.1. Visualization tool 01: Consolidation Delta for consolidated DIM Factor

On the second dashboard (Figure 5.5.2) the scatter plot is showing change in consolidation and the size of the different pieces of business. All cities are selected on the map so all routes are showing on the plot, each route is represented by a color.
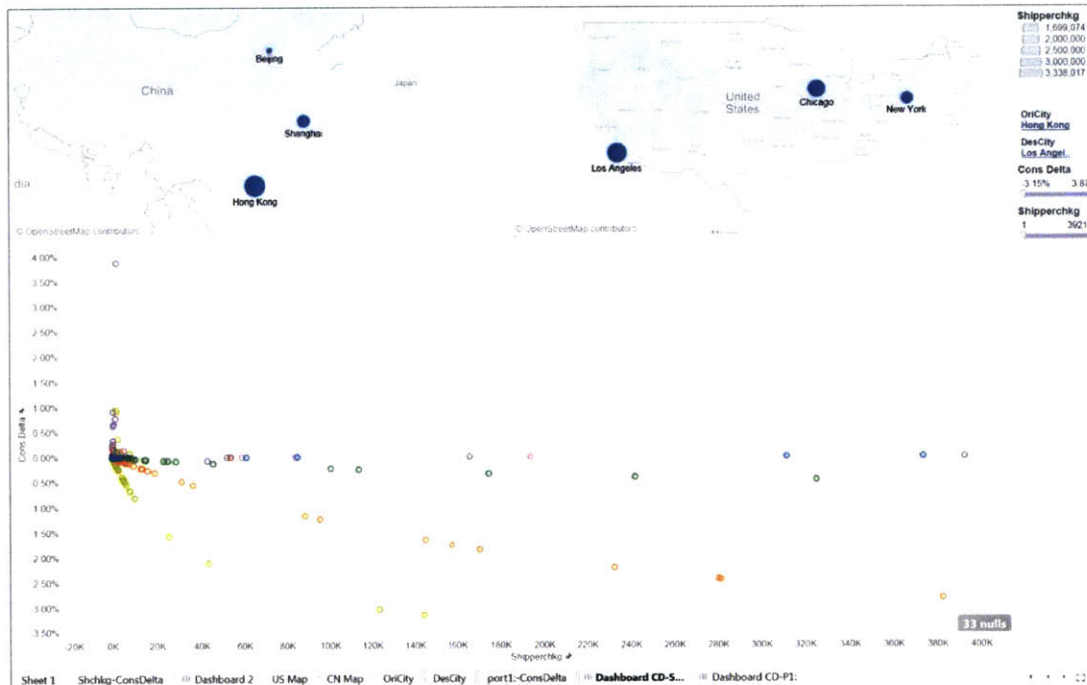

Figure 5.5.2. Visualization tool 02: Consolidation Delta for Chargeable Weight σ

6.- Discussion

6.1.- Dataset analysis

A key takeaway from working with the data provided is that there does not seem to be a standardized approach for collecting it. The sheets for the different customers are similar, but sufficiently different to make the standardization of the data very time consuming. We got the impression that the data is not collected with the purpose of supporting a centralized decision-making process, but rather to track the workflow of the account manager for the customer. This impression comes from the large number columns of which only a few would seem to be useful on a strategic level.

Another issue with the data is the large number of columns that are mostly empty. We suspect that the high amount of data fields are having a detrimental effect on the data input discipline, if most fields are left empty it is less clear which ones are the most important and thus should be pursued with extra effort.

6.2.- How to build a database

As mentioned above the air rate data set has a somewhat bewildering range of fields, this can be contrasted with the quite limited number of key inputs that are needed for the metrics we are using. We would suggest that the commercial level implement a data input system that is the same across the business and only contains a limited number of fields that are critical at the commercial meeting. A standardized input system for the key data will significantly simplify the consolidation of the data. A standardized system would also communicate which information is seen as important for higher levels. The account managers can then maintain their own data sets for their needs, similar to the ones kept now.

The standardized input system could be implemented in different ways; it could be done directly in SQL, through a web-based system, or simply by submitting Excel files. Which method that would work best would depend on the current IT infrastructure and capabilities of the staff. The fastest way to implement would probably be to collect Excel files, since it would require minimum training and system development. Main drawback would be that the central data processing might be time consuming if there is a very large number of customers.

6.3.- Data simulation discussion

The results of the goodness of fit tests applied to the DIM Factor of the current bids for the RFI01 and RFI03b datasets helped us understand the nature of the distributions that best describe such variable. When analyzing DIM Factor value distribution for RFI datasets (footwear and sportswear industries) we recurrently obtained distributions that have a high frequency of DIM Factor values around 9, being most of them between 6 and 15, with a non-symmetrical tail skewed to the right, and a small percentage of tenders with dense cargo (DIM factor under 6, average density over 167 kg per CBM) which accounted for around 5 to 15% of the cases. Such results show noticeable scarcity of dense cargo that limits consolidation potential. The fact both RFI datasets belong to an industry where we may expect cargo to be relatively light, may represent a bias in the results. With that being said, we should highlight the fact that being able to have access to DIM Factor values of customers from other industries with heavier average DIM Factors and density within the same o-d lane may help us build a more well-rounded depiction of consolidation opportunities.

When addressing the goodness of fit test regarding gross weight data for the current bids of the Air Rates, we obtained the best goodness of fit according to the Kolmogorov-Smirnov test with a Generalized Gamma Distribution with parameters a=0.38684 b=1906.8 g=100.0. As a result, we obtain a distribution where the higher the gross weight, the lower the likelihood of the frequency of such γ_w value. In other words, likeliness of occurrence keeps an inversely proportional relationship with respect to gross weight of the tender. The most frequent values of γ_w occur for smaller shipments in terms of gross weight.

As a combination of both reflections, and as for our study we assumed independence between variables gross weight and density / DIM Factor of the bid, we can conclude that at least as far as the customer is concerned with respect to the textile and footwear industry most bids will represent relatively low gross weight and relatively lighter cargo (DIM Factor > 6, δ < 167 kg per CBM). This assumption is confirmed by observing in the graphs of the visualization tools for a given lane dense point clouds with relatively small γ_w values, and DIM Factor values above 6, mostly between 6 ans 15. Such combination that describes most tenders does not offer any consolidation opportunity, what diminishes the potential consolidation delta Δ.

Since the probability of occurrence for a given value of gross weight or DIM Factor is directly dictated by the cumulative probability of the distribution, which can be calculated as the area under the lognormal or gamma curves, or through an integral of the values from zero to such value, the higher the gross weight and density we seek, the least probable such combination will be.

Given that increasing the level of consolidation will overall boost portfolio or business profitability as a result of the surplus produced by being able to charge twice a rate to different shippers when combining heavier and lighter cargo in the same load, air freight forwarders should focus on pursuing the most attractive business opportunities, i.e., the biggest and denser shipments possible.

Such ideal consolidation efficiency approach does not take into account the fact that operational limitations such as pallet size, pallet allocation policies, etc. may difficult the feasibility of certain consolidation combinations. But such considerations are beyond the scope of this thesis. Regarding tracing the main conceptual lines of the commercial strategy of an air freight forwarder, the assumptions considered in this thesis and embedded in the analytical model should provide managers with a basic conceptual framework that helps to guide decision-making when addressing tenders.

6.4.- Metrics and forecasting tool potential

The performance metrics created assist in the decision-making process by assessing the level of attractiveness of current bids in relationship with the current business for a given o-d lane. Consolidation Performance Metrics as defined in Table 4.4.1 describe whether and to what extent there is consolidation potential, a potential that can be used as an upper bound for a given bid under consideration. Profitability Performance Metrics allow for assessing the economic performance of the current bid after consolidation with the current business, as well as providing guidance on setting reference values for rates β when addressing the different rounds of the tender process.

The level of aggressiveness in bidding Γ estimates how aggressive a bid is in relationship to the rate paid to the carrier ϱ. The lower the aggressiveness level Γ, the more aggressive the approach is. The higher the potential for consolidation Δ and $C\%(f)$ is, the more aggressive and by extension competitive with respect to the RFQ an air freight forwarder can be, while still being profitable.

Profits per lane per year Π_λ is metric of central relevance since the ultimate goal of the company is to be profitable. It can be used a quantitative assessment of economic performance, whereas Profitability per lane per year $\Pi_\%$ can be used as a qualitative assessment of economic success.

Finally, the three threshold provided regarding the rate β can be used as a framework for assessing the appropriateness of different rate levels during the bidding processes. However, some caution needs to be exercised when using them due to the underlying assumption of perfect consolidation efficiency. Such simplistic assumption may underestimate the rates needed to achieve the desired profitability levels, so that they should be used as approximated lower bounds. The consequences of underestimating the β value will particularly affect those cases where cargo is relatively dense, since there are more chances of damaging consolidation efficiency.

ANALYTICAL MODEL EMBEDDED IN SQL

The SQL query adds value by automating most of the data processing and thus minimizing both the work required and the risk of errors. The same calculations would be hard to automate in Excel since it would require interactions within and between several datasets. It would furthermore be complicated by the fact that the datasets would constantly be changing as contracts expire and new bids come in. Those changes are however easily handled in queries by changing the WHERE clauses. As expected it also turned out to be significantly faster than managing the data in Excel, even though our query has multiple steps and performs several calculations it executes instantaneously on a standard laptop computer.

With more data additional metrics could be created, a part we did not work on due to limited data available was target rates, it would however be quite easy to extend the query to incorporate the changes in profitability if the target rate is adopted. Another potential extension would be to incorporate unused capacity on a lane that has already been bought from the carrier, but not yet filled. Since the cost of that capacity would essentially be a sunk cost it would be similar to consolidation where you use capacity that has already been paid for.

6.5.- Visualization tools potential

The visualization tool enables senior management to get an overview of the business opportunities at a glance. In spreadsheets form it is hard to get an overview since there are several relevant metrics that need to be taken into account at once; such as amount of business, density of cargo, and potential for consolidation. To get an overview on large sets of multidimensional data that is presented in spreadsheet format it is hard under the best of circumstances. To expect professionals to do it in the context of a time-constrained meeting is unrealistic. Our dashboard built in Tableau makes the key information readily available. It also allows for high-level comparisons that are unfeasible to achieve without visual aids, such as comparing the business profile of two customers, or different routes.

On top of giving a visual overview the dashboard also gives access to more detailed information for every bid, such as the rates for the different levels of profitability that we developed, by simply clicking the dots in the scatter plot (Figure 6.5.1). It is also possible to retrieve the underlying data in a table format through Tableau (Figure 6.5.2), which means that users can look into the details without having to use the somewhat intimidating interface of SQL Server.

Figure 6.5.1. Detail display upon bid selection with Dashboard 01

85
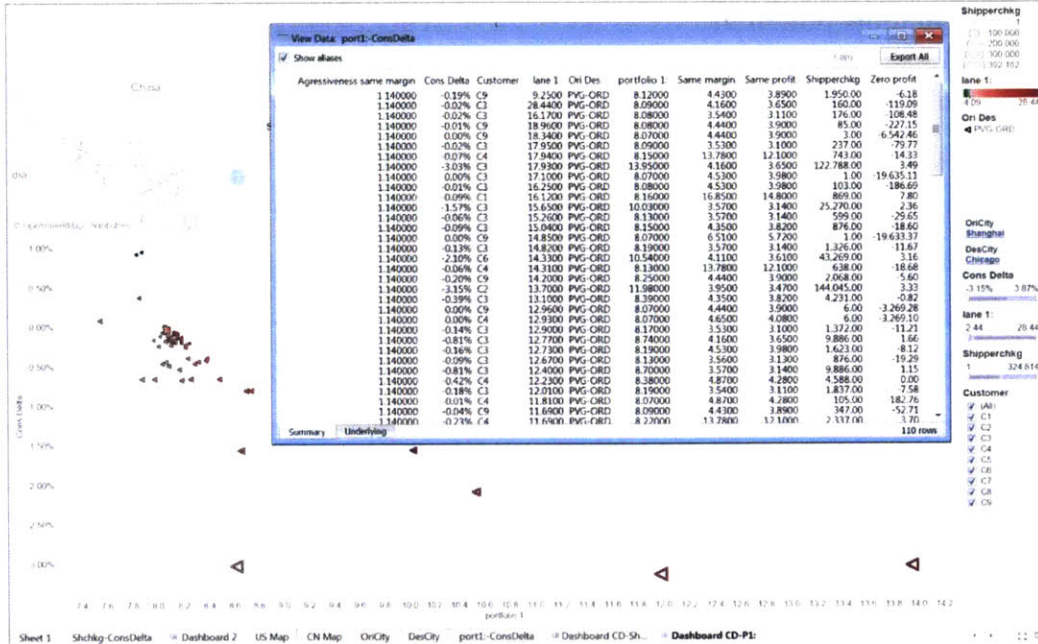
Figure 6.5.2. Display of underlying data table in Tableau

Filters of different types are also easy to use in the dashboard allowing users to focus only on bids with certain characteristics, again without having to change anything in the underlying SQL query. An example is shown in Figure 6.5.3 where the result has been filtered to only show bids for Shanghai to Chicago where the change in consolidation is greater than -1.5%.

Figure 6.5.3. Bids filtered with respect to Consolidation Delta Δ

86

In general we have to be careful when interpreting the pattern of data as shown in our visualization since key parts of it has been generated by us. We can however still comment on some trends that are due to the design of the metrics rather than the specific data that underlies it. One such example is the curves in figure 6.5.4.
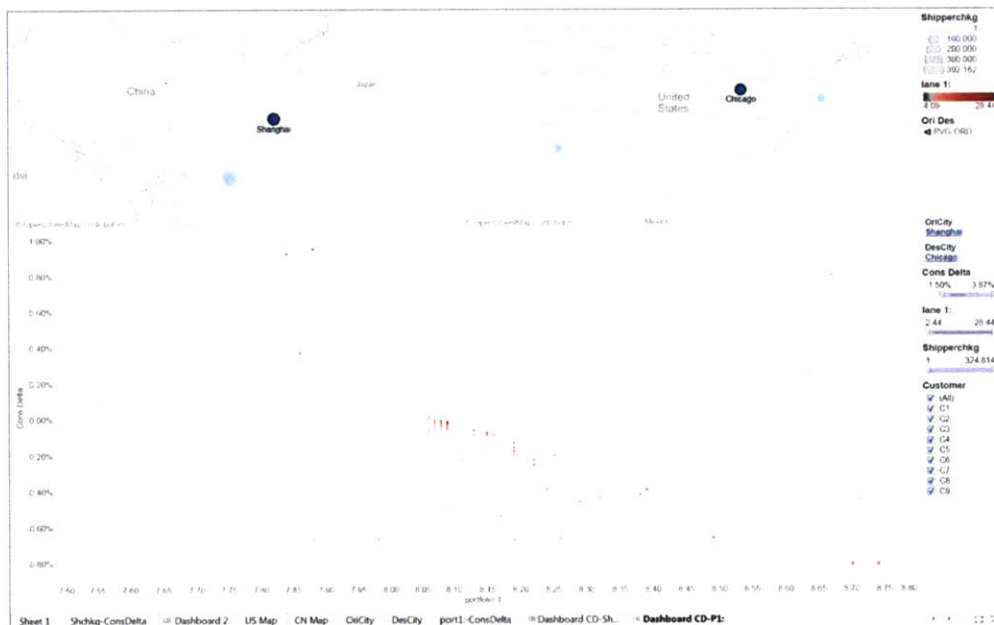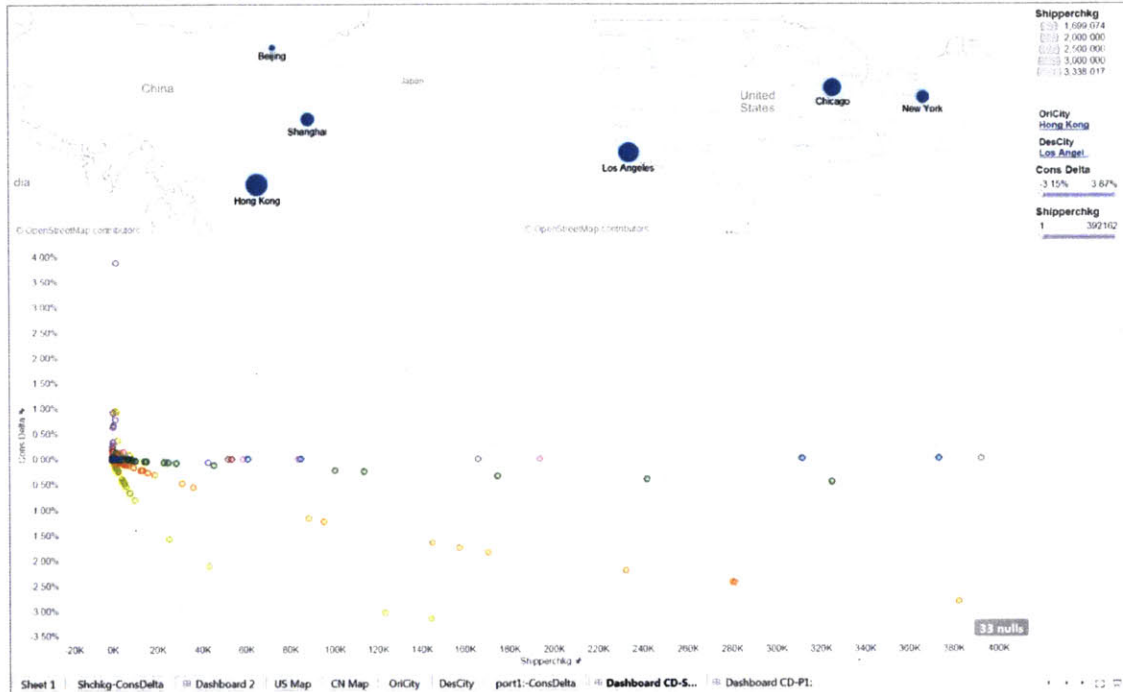

Figure 6.5.4. Consolidation Delta Δ with respect to Shipper Ch kg σ

The bids from each origin destination pair seem to line up on separate curves (Figure 6.5.4). The plot has the change in portfolio consolidation on the y-axis and the size of the individual bid on the x-axis. The reason for the negative curves is that a bigger bid will have a bigger impact on the portfolio. All things being equal for a light bid (DIM Factor > 6) the further right (bigger) it is, the further down it will be, i.e. more negative impact on the consolidation. The reason for the different routes being so clearly separated is that the potential negative change in consolidation is dependent on the original consolidation of the portfolio for that origin destination pair. Furthermore the consolidation cannot be lower than 0, so as the original consolidation level (C%) approaches 0 the potential change in consolidation (Δ) also approaches 0. In other words, if the original consolidation level is really low, it cannot get that much lower.

A somewhat counterintuitive type of bids that show up in the visualization is bids that simultaneously make the portfolio overall denser and decrease the level of consolidation, such as the bid in the lower left in Figure 6.5.5.

Figure 6.5.5. Consolidation Delta with respect to Portfolio DIM Factor (Hong Kong – LA route)

One might think that if the portfolio becomes denser that would improve the consolidation, within our model that is however not true. This is due to a few underlying assumptions. The main one being that our capacity with the carriers perfectly matches our current business and we are able to achieve the theoretically possible consolidation on the operational level. This assumption means that if the portfolio DIM Factor is greater than 6 you can only improve the consolidation if the bid has a DIM Factor that is less than 6. The reason for this is that for such a portfolio the volume capacity is fully used, so every volumetric kg added to portfolio means that more capacity needs to be purchased. If a bid has a DIM Factor greater than 6 it consist of more volumetric kg than gross kg and when combined with light portfolio the amount of new capacity purchased equals the number of volumetric kg, meaning that there is no consolidation of the bid.

However, if the DIM Factor of the bid is lower than the DIM Factor of the portfolio it will still lower the average DIM Factor, even though there is no consolidation. So in theory the consolidation does not improve, that does not necessarily mean that the actual achieved

88

consolidation on the operational level is not improved when DIM Factor decreases. If the constraint regarding achieving the theoretical consolidation potential is relaxed that would mean the actual consolidation would be a function of the theoretical consolidation potential and the consolidation efficiency. It is then entirely possible that a decrease in theoretical consolidation potential is more than offset by an increase in consolidation efficiency so that the outcome is that actual consolidation increases. An interesting hypothesis for future research would be that a decrease in portfolio DIM Factor increases the consolidation efficiency. We cannot say anything regarding the veracity of this hypothesis based on our research. A further implication of relaxing the constraint on consolidation efficiency is that the rates we calculate might be too aggressive for dense cargo and too conservative for certain types of light cargo.

7.- Conclusions

This thesis aimed to establish a framework to build an analytical model for airfreight forwarders to approach the consolidation problem when facing tenders for buying and reselling air cargo space. As a result thereof, we came up with a series of conclusions that synthesize the main findings of our study.

First, throughout the study we wanted to emphasize the relevance of the necessity to create a repeatable, scalable process to gather and store data regarding both past tenders and current bids, especially with respect to some key variables involved in the consolidation processes. By doing so, air freight forwarders will be able to access complete records and keep traceability of data inputs in an effortless way. Such scalable, repeatable process will allow 3PL and especially airfreight forwarders not only to draw an accurate depiction of the consolidation and profitability performance of their current business/portfolio, but also to assess the attractiveness and consolidation potential of current bids under consideration.

Having fragmentary records with incomplete columns with respect to some certain key variables that are explicitly needed to assess consolidation levels and potential meant a significant constraint that limited representativeness of the results of the analysis. However, by performing statistical analysis on those values suffering from lack or scarcity of data inputs via goodness of fit tests we were able to identify significant patterns and populate missing values by running simulation models to create a more robust dataset. Through the process we lost potential internal correlations between certain values, but such counterpart was beyond the scope of this thesis, and did not damage at all the process to create an analytical model and to apply it to the resulting database.

Building a complete, accurate and well-structured database, particularly when having to deal with creating and integrated, holistic database by combining datasets created by different teams can be a challenge for maneuverability of records and data integrity. Being able to integrate the different datasets from different professional teams through a database management system such as SQL Server can help air freight forwarders to acquire a new level of proficiency in data management, by implementing a set of queries to merge pieces of information from different datasets and teams, manipulate data in an efficient way, and create significant metrics to guide decision-making. Air freight forwarder companies will benefit from implementing and refining data management tools that permit to analyze information in a recurrent, automatic and systematic way through calculations

that operate over the incoming RFQ datasets, to synthesize and interpret as quickly as possible the results of the analysis and knowingly participate in the bidding processes.

To do so, we created a series of metrics to both illustrate the current performance and to assess of attractiveness of current bids under consideration regarding consolidation and profitability potential. The performance metrics created assist in the decision-making process by assessing the level of attractiveness of current bids in relationship with the current business for a given o-d lane. On the one hand, Consolidation Performance Metrics as defined in Table 4.4.1 describe whether and to what extent there is consolidation potential, a potential that can be used as an upper bound for a given bid under consideration. On the other hand, Profitability Performance Metrics allow for assessing the economic performance of the current bid after consolidation with the current business, as well as providing guidance on setting reference values for rates β when addressing the different rounds of the tender process.

Increasing consolidation potential permits to boost overall profitability, as a result of the surplus produced by being able to charge twice a rate to different shippers when combining heavier and lighter cargo in the same load. Since dense air cargo (δ > 167 kg per CBM, DIM Factor < 6) is relatively scarce, and frequency of gross weight values is inversely proportional to probability of occurrence, we centered our efforts on creating metrics that enable air freight forwarders to focus on pursuing the most attractive business opportunities, i.e., the biggest and denser shipments possible.

Among other contributions, we focused primarily on the surplus that consolidation permits to achieve by leveraging the increased level of aggressiveness on bidding Γ when addressing the last rounds of the RFQ. To remain profitable not only per se, but also as for the whole current business/portfolio for a given o-d lane, air freight forwarders can use as reference values the Break-even rate for constant profits (β_π), and the Constant profitability per lane rate (β_λ), that define profitability Areas A, B and C. The β_π rate guarantees the same profits in absolute numbers overall per lane. It can be used as a lower bound, threshold or minimum value to determine β; whereas Constant profitability per lane rate (β_λ) guarantees the same level of profitability as the current business. Higher rates, if awarded, would bring not only more profits in absolute terms to the air freight forwarder, but also a level of profitability higher than the current one for a given lane.

The ability to minimize average cost per load of shipments through consolidation will lead to achieve lower values of the Net Achieved Rate for the whole o-d lane or route. This reference can be used to guide attractiveness of current bids when studying incoming RFQs. However, some caution should be beard in mind when taking such approach, since under certain circumstances such as premium services that entail higher nominal costs but also high profit margins the overall cost may increase, boosting the NAR, but the resulting overall profitability might be even higher.

The visualization tool allows senior management to get an overview of the business opportunities at a glance, by providing with fast-processing of information and delivering an illustration of the meaningful metrics we designed in an intuitive and attractive way. In traditional spreadsheets form it is hard to get an overview since there are several relevant metrics that need to be taken into account at once; such as amount of business, density of cargo, and potential for consolidation.

The analytical model developed in this thesis serves to build a methodology to rationalize decision-making criteria when addressing tenders. However, since it primarily focuses on strategic aspects of the commercial policy, there is from our perspective a need to combine such high level approach with operational constraints to build a comprehensive methodology.

Therefore, we suggest that further research in this field should center its efforts on aligning the commercial strategy with operational procedures built upon constraints and limitations that occur at a consolidation execution level. One possible path to be followed would be to combine the analytical model embedded in SQL Server with Lagrangian-relaxation heuristics to solve through software such as CPLEX the optimization of the consolidation problem at an operational level.

Other lines for further research would be to incorporate the effect of taking advantage of discount policies on economies of scale, pallet size, pallet allocation, capacity management, probabilistic analysis on chances of being awarded with a tender with respect to the level of aggressiveness in bidding, big data analysis to create clustering/grouping procedures to increase the understanding of each segment of the market, and increased focus on premium services.

References

Airports Council International - North America (ACI-NA). Air Cargo Guide. Chapter 1:
An historical perspective (2013).

Bowen, John; Leinbach, Thomas. Market Concentration In The Air Freight Forwarding Industry.
Tijdschrift voor Economische en Sociale Geografie – 2004, Vol. 95, No. 2, 2004, p174–188.

Burnson, Patrick. Building air cargo relationships. *Logistics Management*, June 2012.

Burnson, Patrick. Will consolidation change air cargo? *Logistics Management*, March 2013.

CAPA - Centre for Aviation, Boeing, IATA. World Air Cargo Forecast 2014-15 (2014).

Chan, Chi-Kong; Chow, Harry K.H.; So, Sunny K.P.; Chan, Henry C.B. Agent-based flight
planning system for enhancing the competitiveness of the air cargo industry. *Expert Systems
with Applications 39* (2012) 11325–11334.

Donatelli, David J. Evolution of US Air Cargo Productivity. *Massachusetts Institute of Technology*, 2012.

Han, Dong Ling; Tang, Loon Ching; Huang, Huei Chuen. A Markov model for single-leg air cargo
revenue management under a bid-price policy. *European Journal of Operational Research.* 200
(2010) p 800–811.

Harding, William Barkley. Air Transportation Rates. *Harvard Business Review.* Spring 40, Vol. 18 Issue
3, p337. 7p.

Huang, Kuancheng; Chang, Ko-chen. An approximate algorithm for the two-dimensional air cargo
revenue management problem. *Transportation Research Part E* 46 (2010) p426–435.

Huang, Kuancheng; Chi, Wenhou. A Lagrangian relaxation based heuristic for the consolidation
problem of airfreight forwarders. *Transportation Research Part C* 15 (2007) p235–245.

Li, Zichao; Bookbinder, James H.; Elhedhli, Samir. Optimal shipment decisions for an airfreight
forwarder: Formulation and solution methods. *Transportation Research Part C* 21 (2012) 17–30.

Malchow, Matthew B.; Kanafani, Adib. A disaggregate analysis of port selection. *Transportation
Research Part E* 40 (2004) 317–337.

Marketline Industry Profile. Air Freight: Global Industry Guide (2015).

Onghena, Evy; Meersman, Hilde; Van de Voorde, Eddy. A translog cost function of the integrated
air freight business: The case of FedEx and UPS. *Transportation Research Part A* 62 (2014)
p81–97.

Qin, Hu, Zhang, Zizhen; Qi, Zhuxuan; Lim, Andrew. The freight consolidation and containerization
problem. *European Journal of Operational Research 234* (2014) 37–48.

Vinod, Ben; Narayan, Chittur P. On optimising cargo rates to improve the bottom line. *Journal of Revenue & Pricing Management.* Dec2008, Vol. 7 Issue 4, p315-325. 11p.

Wong, Wai Hung; Leung; Lawrence C.; Hui, Yer Van. Airfreight forwarder shipment planning: A mixed 0–1 model and managerial issues in the integration and consolidation of shipments. *European Journal of Operational Research* 193 (2009), p86–97.

Xiao, Baichun; Yang, Wei. A revenue management model for products with two capacity dimensions. *European Journal of Operational Research* 205 (2010) 412–421.

Xiang, Zhu; Ding-you, Lei; Ying-gui, Zhang. Freights loading optimization with balanced and unconcentrated loading constraints. J. Cent. South Univ. (2014) 21: 3386−3395.