Implementing Flight Management Controls for Air Cargo Revenue Management

Eric Jonker
Internship report
Business Mathematics and Informatics

Aard van der Vaart, Maarten Soomer
Vrije Universiteit
Faculteit der Exacte Wetenschappen
Divisie Wiskunde
De Boelelaan 1081a
1081 HV Amsterdam

Joris Wels, Richard Broekman
Polar Air Cargo Headquarters
Europe, Middle-East, Indian subcontinent and Africa
Network, Revenue & Capacity
Triport 1, Room 1148
Evert van de Beekstraat 17
1118 CL, Schiphol Airport

May 2006

POLAR AIR CARGO

Preface

Goals of this Document

The purpose of this document is to report on the results of the investigations I have conducted during my internship at Polar Air Cargo (Polar), Amsterdam, for my Business Mathematics and Informatics study at the Vrije Universiteit (VU) of Amsterdam. The internship is one of the final requirements for this study. The purpose of the internship is, amongst other things, to investigate, analyze, abstract and, preferably, solve a problem that fits within the goals of the master programme.

This document has to serve as a guide to implement revenue management at Polar Air Cargo. It should provide an overview of relevant concepts to business managers as well as explain its practical application in a business-oriented environment to professionals in the air cargo industry, whose primary field of experience is neither mathematics nor statistics.

Problem Definition

The air cargo business is a volatile business. Often, cargo carriers are operating in a highly competitive environment where customer loyalty typically is low and based largely on the pricing of their freighter capacity. Due to the high cost of operating freighter aircraft, the relatively low profit margins and a large volatility in the demand, the operating risk is high and even the smallest increase in revenue generated is likely to contribute fully to the overall operating profit. The revenue management solution proposed means to increase this revenue systematically.

> **Note** This report shows that revenue management is capable of providing a revenue increase of up to € 143,520 for flight PO605/6, per year, reducing total payload carried by around 300 tonnes actual weight and 2,818 m^3. A higher demand will strengthen these results. It is to be expected that all outbound flights from the sales region will show similarly positive results.

The goals of the internship were to implement basic controls for single leg, freighter aircraft flight management. More specifically, the internship efforts were directed at building a model for forecasting demand for cargo only aircraft capacity and giving guidance towards implementing a practical revenue management solution.

This report will show that a systematic, quantitative approach to revenue management at Polar Air Cargo will increase day-to-day operating profit and will provide management with much more accurate information for decision-making.

About the Internship Sponsor

The research was conducted at the Network, Revenue & Capacity department at Polar Air Cargo Europe, Middle East, Indian Subcontinent and Africa (EMEIA) headquarters, Schiphol Airport, Amsterdam.

Polar Air Cargo, Inc. is a global leader in the international air cargo market, specializing in time-definite, airport-to-airport, scheduled freighter service. The company provides a critical link in the international logistics chain by connecting major cargo markets in the Americas, Asia, Europe and the Far East via frequent Boeing 747 freighter services.

Its Network, Revenue & Capacity department provides data quality and data mining services for the financial and operational management, as well as day-to-day flight management and mid term aircraft routing and scheduling for the sales region in cooperation with its counterparts in America and Asia.

A Word of Thanks

I would like to thank my co operators at Polar Air Cargo: Joris Wels, manager of the Network, Revenue & Capacity team, and Richard Broekman, Flight Analyst for Network, Revenue & Capacity, for answering my many questions, their patience and support. Thanks to Aard van der Vaart, Professor of Stochastics with a chair at the VU Department of Mathematics, for the many hours of interesting talks and valuable advice and to Maarten Soomer, PhD for the Optimization of Business Processes research group, for your interest and time. Finally, thanks to my dear uncle Huib Alberts, retired teacher English, for his willingness to spend many hours correcting part of this document.

Executive Overview

Goals of the Investigation

The flight management controls to be implemented are primarily meant to provide sales personnel at Polar with a means to discriminate between requests that should be accepted on a flight or rejected in order to optimize flight revenue. This decision process, called admission control, is the heart of what revenue management is.

Secondly, an accurate forecast of the future demand for air cargo capacity is vital to any revenue management solution, since it supports management in setting a price policy. In addition, it provides valuable management information, since it may be useful in making aircraft routing decisions or directing the sales and marketing efforts.

Finally, the procedures in relation to responsibility for flight performance as well as collection, retrieval and reporting of flight information are to be considered part of an effective revenue management solution. These must eventually be evaluated and changed.

Principles of the Research Design

Central in the research design is a mathematical model of the process in which demand is observed by sales personnel. This process is such that individual space requests arrive one after another with random intermittency, each of which has a certain random weight, volume and revenue. The mathematical model describes when space requests arrive and with what chance its weight, volume and revenue falls within a certain range. The process by which arriving requests are generated is modelled meticulously through statistical analysis.

Knowing with what probability a space request is of certain weight, for example, allows us to generate as many space requests as we require through simulation. Using a computer random number generator, a probability can be drawn, which the model links with a certain weight. An accurate model will produce space requests this way that behave exactly as real space requests. Complete demand sequences are obtained by constructing a timeline of when those space requests arrive. Such a sequence is a single occurrence of the demand that might be observed in practice. A total demand is obtained by summing up the weights, volumes and revenues of the individual requests. By doing this sufficiently often, the behaviour of the demand can be studied. This way, a detailed forecast is produced and a flight performance forecast can be obtained by performing automated admission control. A flight performance forecast is thus only accurate if it was created using the admission control policy that is also used in practice.

The price policy used for admission control is the key to the production of optimal flight performance. The core concept involves valuing the available capacity. The value of the capacity is the revenue that it can be expected to generate. A request should be accepted if and only if the act of accepting it generates more revenue than the value of the capacities it uses. Admission control policies differ mainly in the way the capacity is valued.

The admission control policy in use at Polar is based on what is called the Capacity Access Price. This is essentially a minimum rate per kilogramme a request must generate in order to be considered for admission on the flight. Pak et al (2004, page 8) show that using separate thresholds for a unit of weight and volume, called bid-prices, must be optimal, at least theoretically. They prove this optimality holds if a single pair of bid-prices is used during all of the pre-flight booking period, as long as the bid-prices are set correctly and requests are small in relation to a freighters capacity. The latter condition expresses that it is suboptimal to reject requests just because they do not fit. The policy has the effect of selecting those requests that generate the highest profit per unit of capacity. A policy using a Capacity Access Price has theoretical disadvantages, because it does not fully recognise freighter capacity is two-dimensional and because it cannot be configured flexible enough.

The question remains how these policies perform in practice and how to configure the policies correctly. The latter can be done using simulated demand sequences covering the entire demand spectrum. Each single simulated sequence can be used to compute a single optimal bid-price pair or optimal Capacity Access Price for that sequence, as if it were the future demand observed in hindsight. Since each sequence has equal probability of being observed in real life, the value to be used in practice is taken to be the average of this series of optimal values.

Main Research Results

During the internship, a computer system was created that is capable of the following.

- To estimate parameters for a demand model using historical records of cargo airlifted.
- To simulate an arbitrary number of demand sequences, forecasting total demand.
- To compute an optimal Capacity Access Price for any simulated sequence.
- To compute optimal bid-prices using an algorithm Lenstra et al (1982) have provided.
- To perform automated admission control with both policies, forecasting flight performance.
- To produce single flight performances as if a flight departed with simulated payload aboard.

Two series of such flight performances are produced for 52 PO605/6 flights operated in the past, one using a Capacity Access Price and one using a bid-price pair. This is repeated 100 times. The flight revenues resulting from this are compared with each other in order to establish the relative performance of the policies.

Historical flight performance of PO605/6 was € 156,032 over 66 flights from January 10, 2004, with on an average 103,270 kg payload weight and 567,507 dm^3 payload volume from 30.79 requests per flight. Part shipped requests are contributing to the flight performance of the flight they were first booked for. Therefore, the reported figures might be somewhat different from these reported on the closeout reports of the flights.

On an average per-flight basis, using a static bid-price policy the forecasted performance is € 2,067 higher than historical flight performance and each flight carries 6,105 kg and 55,869 dm^3 less payload. With 52 flights per year, revenue improves by an expected € 107,484 and at the same time, average payload is reduced by 317.5 tonnes and 2,905 m^3, also reducing the cost of airlifting the total payload accepted.

This analysis also shows that a static bid-price policy on average outperforms a policy based on the weekly optimal Capacity Access Price by € 87,672, with 52 departures, under the demand levels being experienced. The policy does this at the cost of carrying an additional 134,368 kg actual weight over 52 flights, but also carries 254,852 dm^3 less volume. Since a static bid-price policy accepts on average only 15.6 requests more over 52 weeks, shipping, handling cost, and administrative overhead are small. The extra revenue accepted is accepted against little more than the absolute lowest differential cost, most notably consisting of fuel and security cost. Since it is reasonable to assume the break-even point with regard to the constant cost has been surpassed, the biggest part of the extra revenue can be expected to contribute directly and fully to operating profit.

For PO605/7, however, the demand levels might be too low during part of the year for the admission control policies implemented to be useful. In my opinion, accepting any cargo on this flight that generates revenue in excess of the cost of carrying it might sometimes prove to be better. It is however well possible a special version of the automated admission control algorithms tailor made for low-demand situations would perform better than using a first-come first-serve policy. This is because the way the admission control policies are configured in the current implementation is such that the revenue requirement is over-estimated if the chance that the flight is unconstrained is high.

For starters, the software system created is capable of supporting in setting an optimal Capacity Access Price for daily use, which is founded on an accurate forecast. This alone is a significant improvement to the current day-to-day flight management practice at Polar Air Cargo already. In addition, an admission control policy using bid-prices instead of a Capacity Access Price will further increase effectiveness of the revenue management solution. The solution also provides a means to establish the relative performance of three different admission control policies, such that the most effective policy can be chosen on a flight-by-flight basis.

Please note the admission control policies use a configuration that is static throughout the week. It is to be expected a daily or more frequently updated bid-price pair or Capacity Access Price will provide additional improvement.

Main Conclusions and Recommendations

The investigation shows a systematic, quantitative approach to revenue management can be expected to make a positive contribution to operating profit during day-to-day operation at Polar Air Cargo. In addition, the demand model constructed delivers more adequate management information than was available without it, regardless of which admission control policy is being used. However, additional improvements must be made in order to forecast demand with improved accuracy. The historical records available were of cargo that has been airlifted, exclusively. Therefore, without correction, the demand total is underestimated. An attempt is made to correct this, but the correction method used is not very sophisticated and time-consuming to apply.

By centralizing admission control and enforcing a bid-price policy aimed solely at improving revenue, around a 6-figure yearly revenue improvement, combined with cost savings due to carrying less cargo, can be expected on a per-flight basis. With several outbound flights from the EMEIA sales region alone, a yearly increase of the net operating profit due to scheduled service operation exceeding € 500,000.00 might be within reach with the limited core schedule in operation at the time of this writing.

The practice of segmenting the capacity in allocations and the free space, each with a sales station being responsible for filling that part of the capacity while meeting the budget is deemed suboptimal, because it prevents enforcing a centralized admission control policy. Sales stations should not be allowed to accept low-revenue cargo aboard simply to meet the station budget. An alternative is to formulate requirements for a sustained performance, instead of a single preset minimum revenue and maximum capacity that is to be targeted on a per-flight basis.

The forecasting process and thus the system producing bid-prices require accurate and timely flight status information. The ideal situation is that admission control can be done on-line using real-time information from across Europe. Developing a system supporting this seamlessly and less labour intensive than today will prove to be invaluable and may reduce overhead cost.

In my opinion, it is a disadvantage that the main flight management information system at Polar is proprietary, a closed system and aimed primarily at administrative and legal information needs. Accurate factual records are the basis of forecasting accuracy. Providing accurate records of small shipments should not be under-estimated, since most requests are small. Especially volume information is not accurate enough. This might partially be because there simply is little awareness for its importance.

It is recommendable to record information about requests that are rejected due to capacity constraints or failed price negotiations and about requests that are never airlifted although they once were booked. Currently, only records are kept of cargo that actually has been airlifted. This additional information will be the single most important improvement to the accuracy of the forecasting model, once it is available. Centralizing the revenue management solution in a single information system, including automated or semi-automated admission control, may greatly support collecting these records. Records of requests that were rejected by admission control due to revenue or capacity constraints can simply be retained, separate from the records of cargo that is billed.

Marketing Polar as the partner of choice for uncommon or large cargo has some adverse effects. The revenue management solution will perform better with smaller requests. Off course, any large cargo offered should be considered for acceptance, but actively attracting it might adversely affect flight performance. Rather, it might be beneficial to encourage customers to offer cargo in smaller units more often. For example, forwarders normally requesting shipment for many small pieces consolidated into a series of pallets could be encouraged to do so for each pallet as soon as it is full.

The automated admission control is straightforward and useful as it is currently implemented, but regardless of which admission control policy is used, improvements can be made by implementing a part-shipment scheme, supporting cargo priorities and off-loading and generally making the solution aware of multiple flights, possibly enabling multi-lane cargo routing. The last improvement might be implemented by asking customers to specify a delivery date when they file a request and subsequently by only booking requests after it is determined on which flight they should move, optimizing performance of multiple flights while respecting the service level agreed upon with the customer. Delivery dates closer in the future might be charged with additional fees, which relates to a possible way of implementing a cargo priority system.

I would like to accentuate that the current method of simulating requests that are rejected or do not move for other reasons is not naturally fit for use with multiple flights, especially if several flights included in the forecast operate under the same flight number. Alternate methods based on newly recorded historical records of lost or rejected cargo are to be implemented first. This will also improve accuracy with only one flight.

Table of Contents

Brief Contents

Contents

1. Introduction

1.1. About Polar Air Cargo

Polar is a global leader in the international air cargo market, specializing in time-definite, airport-to-airport, scheduled freighter service. The company provides a critical link in the international logistics chain by connecting major cargo markets in the Americas, Asia, Europe and the Far East using frequent Boeing 747 freighter services. Polar and its sister company, Atlas Air, Inc. (Atlas), are wholly owned subsidiaries of Atlas Air Worldwide Holdings, Inc. (AAWW), based in Purchase, New York.

Polar is organized in three regional headquarters. Next to the European, Middle-Eastern, Indian subcontinent and African (EMEIA) region headquarters based in Amsterdam, The Netherlands, are the American region headquartered in Long Beach, California and the Asian and Oceanian region headquarters based in Hong-Kong.

Polar EMEIA is further subdivided into regional markets called sales stations, with a number of Polar offices taking responsibility for high volume countries and third party agents called General Sales Agents representing Polar in the other countries.

1.1.1. Time-definite, Airport to Airport, Scheduled Freighter Service

Amsterdam Schiphol Airport functions as the hub or gateway airport for almost the entire EMEIA region. At present, only cargo from the United Kingdom is not commonly transferred through this hub. In America and Asia, several airports have this same role. Cargo is being trucked or interlined[1] from almost any airport to those hubs and then airlifted between those hubs, using long haul Boeing 747 cargo only aircraft – also dubbed freighters.

Almost all flights are scheduled on a weekly basis. Scheduled flights between a fixed origin and destination airport have a common, well-known flight number. Flights with a certain flight number may be operated more than once a week. Nevertheless, any flight can be identified by a flight number and a day of the week. A flight departing earlier or 20 minutes later than scheduled is typically considered a problem.

The schedule of the flights that depart every week is called the core schedule. Flights that are scheduled only once, for example because an aircraft is to be re-positioned due to charter operation or maintenance, are commonly called extra sections.

1.1.2. Standard operating procedures

The sales process at Polar Air Cargo EMEIA is dictated using standard operating procedures dubbed Flight Management and Non Allocated Countries. An overview of those procedures follows.

Customers of Polar are almost exclusively forwarders. They provide Polar with cargo for airlifting on behalf of the actual shipper and pay charges per kilogram.

For each outbound flight, all sales stations receive a budget dubbed allocation they may freely allocate for customers, mentioning a maximum number of pallet positions, volume and weight. The sales budget also details targets for revenue in Euros that must be met on each flight. Countries that are assigned an allocation are called allocated countries. Capacity for cargo originating from other countries may be sold in the unallocated space available, called the free space.

The Revenue and Capacity team is responsible for filling up the free space. In this sense, they are another sales station, responsible for all capacity that is not allocated to any country-bound sales station. They also evaluate bookings from allocated countries that overflow the allocation for that country. The Revenue and Capacity team is mainly responsible for pricing and selling ad hoc shipments as well as overflow shipments from allocated countries.

The Revenue and Capacity team should only accept bookings that generate more revenue per kilogram than the daily determined Capacity Access Price. The Capacity Access Price is set by the Network, Revenue and Capacity department. The policy might be altered to try to stimulate the sale

[1] The cargo is being carried by competitor carriers on a contract called an interline agreement

of cargo that has a much higher or lower density than average, depending on the overall density of the payload already accepted, the amount of available capacity and the time to departure. One goal of the investigation is to provide the Revenue and Capacity team with guidance towards setting the Capacity Access Price or to provide an alternative means by which cargo might be accepted.

The Network, Revenue and Capacity department is also responsible for admitting contracted allotments. Contracted allotments are, however, not very common in the current environment Polar EMEIA is operating in.

The way of operation in the other sales regions are roughly the same, but their practical implementation as well as experience differs at points not detailed here due to the unique challenges of the business environment in each region.

1.1.3. Space requests

Booking requests submitted for cargo capacity– dubbed space requests at Polar – contain information on weight and dimensions of all pieces the shipment consists of. Based on this information, an offer is made and a rate is negotiated per kilogram of cargo to be airlifted. Either party may reject the offer before a booking is made.

If a space request is accepted, all pieces in it become pieces on a single airway bill. Each piece is accurately measured for balancing the aircraft, but its volume is usually copied from customer information. If less than all of the pieces are airlifted on a single flight, the airway bill is said to be part shipped. Only weight and volume of the part of the airway bill flying on a particular leg, the part shipment, is recorded.

The cargo offered for airlifting include amongst others perishables like vegetables, pharmaceuticals, confection, electronics, machinery, live animals in cages or even stables, cars and car parts, aircraft engines, even small aircraft and helicopters and all kinds of goods to be classified as dangerous goods. This diversity has a well visible effect on the diversity of sizes, weight, dimensions and revenue potential of airway bills.

A good part of the cargo consists of standard sized colli. A collo is cargo built on standard pallets, like Euro pallets, to form a single standard sized piece. Several pallets can be built on a unified loading device, which fits the rails inside the aircraft for supported loading of almost any cargo. Forwarders often consolidate many small pieces of cargo originating from many different shippers into manageable chunks. This kind of cargo is called a consolidation or consol.

1.1.4. Chargeable Weight and the Pivot Concept

It should be clear that, in the air cargo industry, weight is of primary importance. The payload weight of an aircraft greatly influences the amount of fuel burnt. In addition, if the cargo inside the aircraft is not positioned well, this increases capacitance. Because fuel is the biggest variable cost involved, cargo to be airlifted is weighted and carefully positioned inside the aircraft. The weight measured is called the *actual weight* and is recorded accurate to the kilogram.

However, densities of different kinds of goods vary greatly and cargo that does not weigh much may occupy a relatively large part of the available space inside the aircraft. To compensate for this in an easy way, it is common in the air cargo industry to charge based on a weight called the *chargeable weight* rather than the actual weight.

For shipments heavier than a threshold weight per m^3, the chargeable weight is equal to the actual weight of the shipment. This threshold weight is called *the pivot weight* and is a density measure. For lighter shipments, the chargeable weight is the volume of the shipment multiplied with the pivot weight.
Put otherwise, chargeable weight is the maximum of the actual weight and the volume in m^3 times the pivot weight. This I call *the pivot rule*.

All aircraft Polar operates are Boeing 747 freighters. The least capable of these aircraft can lift around 100 tonnes of payload and have a cargo deck interior of around 600 m^3. Therefore, Polar uses a pivot of 166 kg/m^3. Polar recognises a higher pivot might have been defendable due to a higher maximum payload the newest aircraft can uplift.

The *Système Internationale* uses kg/m^3 as unit of density, but at Polar, m^3/tonne is regularly used instead to ease computations. This is exactly the same as dm^3/kg. This unit is also used throughout this document. Because a density of 166 kg/m^3 corresponds with 6 dm^3/kg, cargo with a density of 6 is said to be at pivot.

> **Note** The default unit of density used in this document is dm^3/kg rather than the *Système Internationale* default of kg/m^3. If no unit is specified, assume density is in dm^3/kg.

Because cargo is loaded onto unified loading devices of standard size before being loaded aboard the aircraft, the exact dimensions of the pieces determine the number of pieces that fit on the loading device and hence usable space occupied. Also, pieces are often not as high as the cargo deck and may not support stacking on top or below other cargo or at all combining with other cargo on the same unified loading device. This reduces efficient use of the available space. To accommodate for this, sales representatives may negotiate a chargeable weight in excess of the default chargeable weight resulting from the pivot rule.

A shipment of up to 1000 kg is considered a small shipment, for which special rates apply by default.

1.1.5. Revenue, rates and surcharges

Based on the information in a space request, an offer is made and a rate is negotiated per kilogram of cargo to be airlifted. The customer pays this rate multiplied with the chargeable weight in kilogramme. This is most often the only condition negotiated. Due to its importance, it is plainly called the rate.

However, the customer also always pays charges meant to contribute towards covering fuel cost and security cost. Those are called the fuel and security surcharges and are rates computed over the actual weight rather than the chargeable weight. Recently, the fuel surcharge has been increasing steadily, but for day-to-day purposes, fuel as well as security surcharges are constant and apply without exceptions.

The concept of surcharges is also used towards covering various special cost, often dependent on the nature of goods the airway bill consists of, like dangerous goods surcharges. In this document, only fuel and security surcharges are considered special. All other surcharges are included in the rates reported.

> **Note** In this document, the revenue generated by airlifting cargo is the rate multiplied with the chargeable weight plus fuel and security surcharges multiplied with the actual weight.

1.2. Solution Overview

1.2.1. Goals of the Investigation

The goal of the internship was to provide usable advice for implementing revenue management controls for day-to-day flight management at Polar Air Cargo. The following goals have been formulated.

- To construct a forecasting model capable of forecasting demand for air cargo capacity for a flight, detailing the demand development from any moment during the pre-flight booking period until departure of the flight and a suitable quality measure for this estimate.
- To establish a price or admission control policy given the demand development during the remaining part of the pre-flight booking period, the remaining capacity, the time a request is made and the properties of the cargo for which movement is requested.
- To investigate in what way the available capacity must be allocated to sales regions.

Initially, the investigation was concentrating on forecasting for and improving revenue generated by that part of the capacity that is not allocated to any sales region, called the free space. The investigation has however generalized, and aims to optimize total flight revenues. This has been a substantiated decision.

The research focuses on Polar Air Cargo flight PO605. PO605 is connecting Amsterdam Schiphol Airport with Chicago O'Hare International Airport every Saturday and Sunday of every week, almost without cancellations, for at least the last two years. Saturday flight is commonly referred to as PO605/6 and the Sunday flight as PO605/7.

1.2.2. Requirements of the Forecasting Model

Seen from the perspective of a single freighter carrier shipping almost exclusively on behalf of forwarders, such as Polar, the demand for air cargo capacity is observed in large chunks that must move either entirely or not at all. Chunks for which movement is requested, the airway bills, are of highly random size and weight. Combined with the relatively low frequency at which space is requested for a particular flight, this has the effect that the total demand observed for any flight is highly volatile. For example, a single, large airway bill accepted on a flight may have the effect of radically changing our expectations for the take-off total revenue for that flight.

Forecasting future demand of course means forecasting average demand. This average is the demand that is to be expected, but the actual demand development will hardly ever follow our best estimates. A structural difference between estimated average demand and the demand observed is exactly what defines demand volatility. The challenge of constructing an informative forecasting model is to capture this volatility as well. Such a forecasting model will answer the obvious questions with regard to a flights overall performance, but will provide valuable additional insight. This insight might point towards a high chance of experiencing an operating loss, or opportunities for routing additional capacity on a certain lane profitably. The benefit of modelling the volatility thoroughly is that everything in between is covered as well.

Preliminary analysis has shown that trend analysis using time series, regression or some smoothing method are too limited to be of real use. They cannot duly capture the volatility of the demand, because they are only capable of producing point estimates. In addition, many of these methods work by assuming a systematic dependency on the historical demand that is too narrow. Moreover, they often lack flexibility, since they are so trivial they do not support any configuration, which is useful for answering hypothetical questions in interesting "what if" scenarios. Finally yet importantly, these approaches are of little help in a simulation, which will prove later in this document to be a limitation of vital importance.

1.2.3. Constructing a Forecasting Model

The approach followed while constructing a forecasting model is that of in depth mathematical model building. Instead of forecasting total weights, volumes or revenues for a whole flight, an approach is put to use that models the process by which space requests arrive and by which they obtain their properties. The number of space requests arriving, as well as their individual weights, volumes and revenues are modelled one-by-one. These principal properties of any space request arriving, although their value is unknown in advance, are asserted to follow systematic laws.

It is these laws that are expressed in the forecasting model, using methods from the fields of mathematical statistics and probability theory. The model also describes the influence of some of the variables on others. Finally, the trivial semantics are modelled of how space requests arrive and build up over time to form the total demand for air cargo capacity for a single flight in a certain period, such that it becomes possible to describe the demand development essentially exactly as it behaves in real life. A detailed description of the forecasting model is given in the chapter 3.1. The Forecasting Model.

The chapter 2.3. Demand Forecasting describes the principles of mathematical statistics. It shows that the intuitively clear concept of frequency quotients, obtained from historical data, is the basis for describing the hidden, but systematic laws to which a variable like space request weight obeys. A central assumption made is, that the variables are generated from a probability distribution, which captures most if not all immeasurable or unobserved reasons, for example for a space request to have that size it has. It might be surprising to see how well the historical records fit the mathematical theory available when they are presented the right way.

The demand until departure of a particular flight consists of a series of space requests arriving. Such a series of space requests is called an arrival stream. An arrival stream consists of all requests arriving, not just those that are admitted on the flight. The forecasting model can be used to generate an arrival stream through simulation, using a computer random number generator. The chapter 2.3. Demand Forecasting describes how this works. Every single arrival stream that is simulated using an accurate forecasting model is one that might well occur in practice. It simply lists how many and when space requests arrive during a selected period and what their individual weights, volumes and revenues are. Simulation allows us to observe realizations of the future demand repeatedly as to gain insight in its behaviour. This insight will allow us to act well informed.

1.2.4. Forecasting Demand

The model of the demand process, described above, can be used to mimic space requests arriving through simulation. Streams of arriving requests, called arrival streams, can be generated as often as required quickly enough. A forecast is produced this way that not only provides an estimate for the total demand, but is also informative about the full range of demand that may be observed. In addition, insight in the demand development over time is obtained, because the time a request is filed is simulated along with the main request properties – actual weight, volume and revenue.

What is even more, each arrival stream can be subjected to some form of admission control or another, to produce a flight performance forecast in addition to a demand forecast. The difference between total demand and flight performance is after all just the number of requests that are rejected. This procedure turns out to be very useful, since it allows different forms of admission control to be compared using flight performances generated using the same series of simulated arrival streams. A sustained difference in the flight performances resulting from different admission control policies indicates that the one is to be preferred over the other.

The chapter 3.4. Flight Performance presents the solution implemented by means of a step-by-step execution of the procedure required to produce a forecast for a particular flight. It also reports on the effects of the solution on flight performances, in relation to historic flight performances and compares of the effects of the admission control policies implemented.

All statistical analysis is conducted using records of cargo that has been airlifted. Records of rejected, cancelled or lost requests are not available. Hence, the forecasting model would underestimate the total demand if it were used without correcting this. The correction method used in the absence of adequate historical records is based on a first-come first-serve admission control policy, under which requests are never rejected due to revenue. Using this policy, historical flight performance is compared to forecasted performance and extra requests are simulated in order for those to match. This method is explained and defended as a workaround in the chapter 3.3. Constructing Complete Arrival Streams.

1.2.5. Practical Revenue Management

A practical revenue management solution provides sales personnel with a means to discriminate between those requests that should be admitted on the flight and those that should be rejected in order to maximize revenue. The question remains how admission control is best performed. The core concept is that a request should be accepted if and only if the revenue it generates is equal to, or in excess of the reduction of revenue achievable in the future. This reduction of revenue achievable in the future is due to using the capacity required for that request. It is called the opportunity cost. Admission control policies differ mainly in the way the opportunity cost is established.

Strictly enforcing a minimum rate per kilogramme chargeable weight through the daily-determined Capacity Access Price would be an admission control policy, for which the opportunity cost equal the chargeable weight times the Capacity Access Price, plus fuel and security surcharges times the actual weight. This admission control policy has been implemented and is evaluated against simulated arrival streams.

Another admission control policy that has been implemented and evaluated is one called a bid-price policy. The idea of this policy is to define a threshold price for a unit of capacity, such that all those requests generating more revenue fill the flight as completely as possible. We want to select those requests that use the least capacity for the profit they generate. The policy has exactly that effect and therefore it is called greedy. The threshold price is called the bid-price. Because capacity on a freighter aircraft is two dimensional, a bid-price is constructed for each capacity dimension.

The opportunity cost of a request under this policy is the sum of the bid-prices multiplied with the weight and volume it uses. If a single pair of bid-prices is being used throughout the pre-flight booking period, the policy is called a static bid-price policy, since the bid-prices never change. Pak et al (2004, page 8) show that this policy is optimal for the cargo revenue management problem when capacity and demand increase proportionally and the bid-prices are set correctly. The former condition means the policy is approaching mathematical optimality when requests are small in relation to a freighter's capacity. Practically, however, it is better when the bid-price levels are regularly renewed, because requests are not that small. The capacity use development throughout the booking period provides advancing insight that is to be acted upon.

The chapter 2.1. Air Cargo Revenue Management gives an overview of the concepts of and prerequisites for effective use of revenue management and the chapter 2.2. Admission Control gives a more detailed definition of admission control in general and admission control using bid-prices in particular.

1.2.6. Configuring the Admission Control Policies

The last step implementing revenue management is to configure the admission control policy of choice to produce the best flight revenue. If a Capacity Access Price is used, what value has a kilogramme chargeable weight? If a bid-price policy is used, what bid-prices must be set?

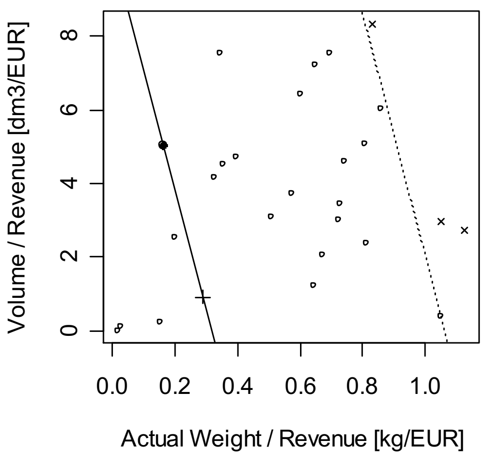
In this situation, simulation is used again to provide the necessary information. If the requests that are to arrive for a certain flight were known, it would have been possible to compute the optimal Capacity Access Price or bid-prices for that demand. The future demand is yet unknown, but recall that a single simulated arrival stream is a reproduction of a possible future demand. A series of simulated arrival streams together cover the full range of possible future demand developments, as long as enough replications are generated. Each such an arrival stream is constructed by using the forecasting model to simulate how many requests will arrive, what the individual actual weights and volumes are and the revenue each request generates.

This series of simulated arrival streams can be used to obtain a series of Capacity Access Prices, or bid-price pairs, that produce optimal flight revenues, one for each arrival stream. Each arrival stream has equal likelihood to be observed. Therefore, configuring the admission control policy is done by simply averaging the values computed to be optimal for the individual simulated streams.

Computing the optimal Capacity Access Price given a single simulated arrival stream is straightforward. Note that the Capacity Access Price is essentially a minimum rate per kilogramme chargeable weight, excluding fuel and security surcharges, that a request must generate to be accepted. For each request in the arrival stream, its chargeable weight is computed from the simulated actual weights and volumes using the pivot rule and the fuel and security surcharges are deducted from the revenue the request generates, at the fuel and security surcharges rate set by company management. The revenue of each request excluding fuel and security surcharges is divided by the chargeable weight. All requests in the arrival stream are then ordered by decreasing value of this rate per kilogramme chargeable weight, such that those requests having the highest rates per kilogramme chargeable weight can be accepted first. Requests are accepted until insufficient capacity remains to accept the next request or until all requests are accepted. The optimal Capacity Access Price is taken to be the rate per kilogramme chargeable weight of the request that was accepted last in order of decreasing rate per kilogramme chargeable weight.

Computing an optimal bid-price is considerably more complex. The idea is to weight each capacity dimension and select those requests that have the highest profit per weighted capacity requirement. The weight ratio of a capacity dimension corresponds with how restricted the capacity dimension is compared to the other. For a single flight, a given set of requests may be graphically represented by points on the flat plane with coordinates corresponding with their weight and volume use per Euro of revenue. Points close to the origin correspond with requests that generate the most revenue per unit of capacity used. Hence, we prefer requests closer to the origin.

Determining bid-prices is done by drawing a line through the flat plane. Requests below the line are accepted, the others rejected. The slope of the line corresponds with how the capacity dimensions are weighted in relation to each other. The height of the line dictates the height of the bid-prices. If the line is further from the origin, the bid-prices will be lower. This slope is selected first. Next, the line is moved upward from the origin, accepting requests in the order they are encountered, until no more requests fit. Refer to Figure 1 for an example. Values from which the bid-prices are computed are then read from the two axes. Since the axes display values in kg/€ and dm^3/€, computing the bid-prices corresponding with a chosen line involves inverting those values, such that they are expressed in €/kg and €/dm^3. This is done repeatedly, each time resulting in a bid-price pair. The optimal bid-price pair is the one resulting in the highest total revenue.



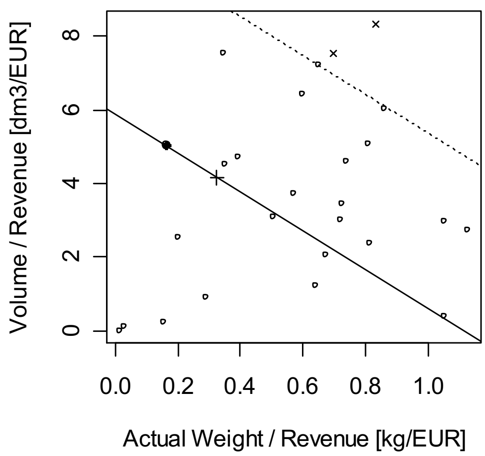


Figure 1 The second and third iterations of the algorithm for obtaining bid-prices. Each dot represents a simulated shipment. The algorithm selectively evaluates request pairs to determine the relative weight of the capacity dimensions (solid), then accepts requests until the flight is full (dotted). Each request pair evaluated produces a single bid-price pair.

Lenstra et al (1982) provide an algorithm implementing this approach efficiently. They do this by skipping evaluation of slopes for the line that have the effect of accepting requests in exactly the same order as a slope already evaluated. Notice that a line through two points corresponding with space requests exactly has a slope for which the requests swap places in the order they are accepted. Making the line less steep has the effect the right hand side request in the pair is accepted first, if it is made steeper, the one on the left is accepted first. In addition, a bid-price pair can only be computed if the line intersects with both the axes. This corresponds with requiring the line to have a negative slope. Therefore, horizontal lines or lines with positive slopes are not evaluated.

The algorithm evaluates requests from left to right, making a pair with each request still further to the right and determining first whether the request further to the right is below the current request. Only then, the slope of the line is computed and the requests are sorted in the order the line encounters them if it moves up from the origin. This is done until the remaining payload capacity is insufficient to accept the next request. The bid-price is then computed from the intersection of the line with the two axes.

This algorithm is described in detail in the chapter 3.2. Determining Bid-Prices. The description given here is implementation ready, providing systematic instructions for each phase of the algorithm up to selecting the optimal bid-price pair from all pairs generated.

1.2.7. Comparing the Admission Control Policies

In this report, results are presented of comparing admission control policies based on a static bid-price and based on the Capacity Access Price, when both are configured for optimum performance. The admission control policies are evaluated against a series of simulated arrival streams. Using the forecasting model to simulate arrival streams based on which optimal bid-prices as well as an optimal Capacity Access Price is computed, the comparison can be made very objectively, because it will be based on exactly the same data.

Theoretically, an admission control policy based on a Capacity Access Price is quite similar to a static bid-price admission control policy. The following similarities hold.

- Both policies can be called greedy, in the sense that they can be enforced by computing opportunity cost based on the capacity use of an arriving request and accepting requests if their direct revenue exceeds this. The former can be called greedy towards chargeable weight, the latter is strictly greedy towards weight as well as volume.
- Under both policies, the opportunity cost increase proportionally with the size of a request. Under a static bid-price policy, the opportunity cost is linear in both weight and volume. Under a policy based on a Capacity Access Price, the opportunity cost is linear in weight as long as only dense cargo or cargo at pivot is considered. It is linear in volume but also in weight, due to the fuel and security surcharges, as long as only volume cargo is considered.
- Both the optimal Capacity Access Price as well as the optimal bid-prices can be computed by accepted requests in order of increasing revenue per unit of capacity. The former is based on revenue per kilogramme chargeable weight add fuel and security surcharges. The latter is based on total revenue per kilogramme as well as total revenue per litre.

In the chapter named 2.2. Admission Control, arguments are provided for an admission control policy based on static bid-prices to be asymptotically optimal when capacity and demand increase proportionally. This means that the policy would be optimal if requests were never rejected because they do not fit. In practice, this will however happen regularly, because requests are big in relation to a freighter's capacity. Due to the impact on total flight performance of accepting a large request, re-establishing the bid-prices regularly during the pre-flight booking period will improve the performance of the policy despite the theory. The knowledge obtained by accepting larger requests has to be acted on.

Both policies under investigation however have their problems. The following arguments are against an admission control policy based on a Capacity Access Price. The policy ignores volume completely if a request is dense. For example, the chargeable weight is equal for requests of equal weight with densities 1 and 5 m^3 per tonne. In addition, fuel and security surcharges are valuable revenue, but only play a part after the request is already accepted. An admission control policy that is aware of fuel and security surcharges will perform more adequately, simply because it is better informed. Finally yet importantly, the pivot ratio dictates how chargeable weight is computed and thus dictates how valuable volume capacity is in relation to weight capacity. Since the pivot ratio is fixed, skewed availability of weight or volume capacity, or high demand for either will not be reflected in the opportunity cost resulting from this policy. The admission control policy based on a Capacity Access Price must be concluded not to generate optimal revenue from the available capacity by design.

The following analysis shows in detail why a policy based on a Capacity Access Price must be suboptimal.

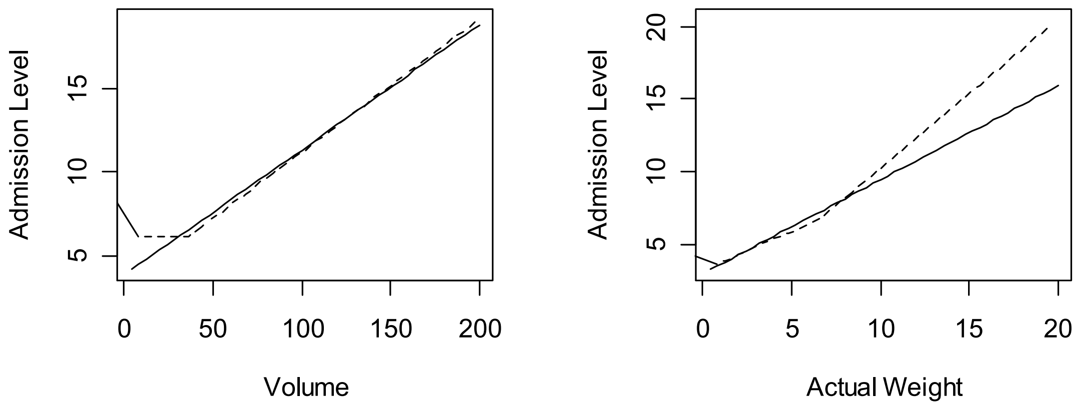


Figure 2 The minimum revenue shipments must generate to be considered for admission. The minimum revenue is displayed for shipments with a fixed weight of 5,870 kg (left) and a fixed volume of 40 m^3 (right) using a static Bid-Price policy (solid) and using a Capacity Access Price (interrupted) for the PO605/6 flight of week 13, 2005.

Let *w* be the weight, *v* be the volume and *r* be the total revenue including fuel and security surcharges of a request. In addition, the following parameters define the admission control policies.

- *pivot* is the pivot ratio in effect
- *surcharges* is the total of fuel and security surcharges per kilogramme actual weight in effect
- *cap* is the Capacity Access Price
- b_v and b_w are the bid-prices for the volume and weight capacity constraints

The revenues required for the request to be accepted under each admission control policy can now be compared against the revenue of a request. Requests are accepted under one of the following conditions, depending on what policy is being used and on the density of the request.

- $r \geq w \times (surcharges + cap)$, using a Capacity Access Price for dense cargo.
- $r \geq w \times surcharges + (v \div pivot) \times cap$, using a Capacity Access Price for volume cargo.
- $r \geq w \times b_w + v \times b_v$, using a pair of static bid-prices for any cargo.

All three requirements are linear in both weight and volume and hence, they form planes in three-dimensional space. If the revenues required are the same for any request, the former two planes will intersect with the last one. They will each do so exactly once whenever they are not parallel.

Theoretically, it is possible the intersection takes place for negative values of weight or volume or for values that are too large to be of practical importance. In practice, they usually will intersect and the lines along which these planes intersect follow from taking the opportunity cost under different policies to be the same. Because the lines define a relationship between weight and volume, their slopes correspond with densities.

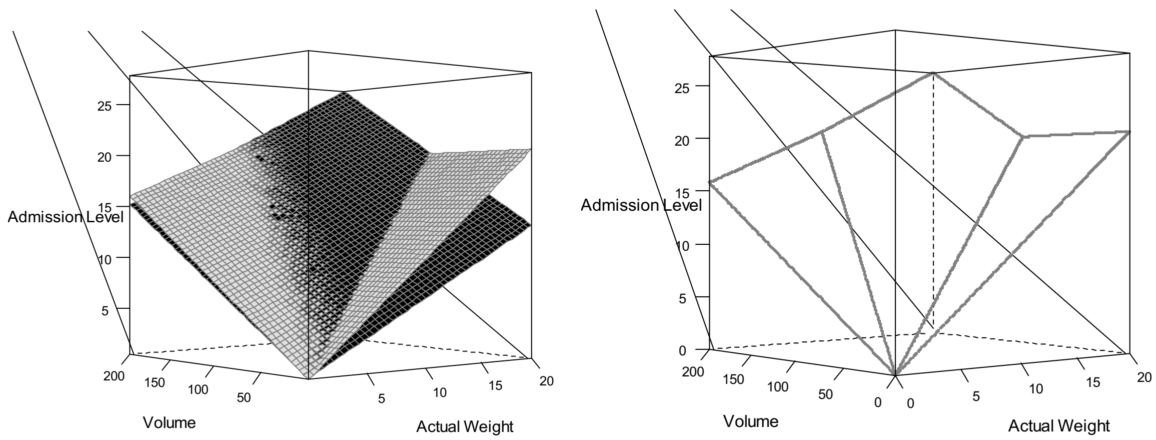


Figure 3 Minimum revenue levels for requests to be considered for admission (left) and the areas where a policy requires the highest revenue (right), using a static bid-price admission control policy (dark grey, centre diamond) and using a Capacity Access Price (light grey, triangles).

Taking the required revenues under both admission control policies to be the same yields

- $w \times (surcharges + cap) = w \times b_w + v \times b_v$ for dense cargo and thus
  $v = w \times (surcharges + cap - b_w) \div b_v$
- $w \times surcharges + (v \div pivot) \times cap = w \times b_w + v \times b_v$ for volume cargo and thus
  $v = w \times (b_w - surcharges) \div (cap \div pivot - b_v)$

These relationships can be used to compute those densities for which the two admission control policies require the same revenue when the optimal Capacity Access Price and bid-prices are computed for a certain flight.

> **Example** For the PO605/6 flight of week 13, 2005, the optimal bid-prices for weight and volume computed are 0.645 and 0.075, respectively. The surcharges are 0.55 and pivot is 1:6. The optimal Capacity Access Price computed is 0.474. Hence, the planes intersect at densities 5.09 and 21.1.

Most notably, the revenue requirement of an admission control policy using Capacity Access Prices is very high for dense cargo. The denser a request is, the more disproportionate the revenue requirement is under this policy. In contrast, it is somewhat low for large volume cargo requests with moderate densities. Note, however, that for these requests the absolute revenue requirement deficit definitely will be considerably large, because the requests are known to be large.

1.2.8. Restrictions of the research

The research is focused on only one flight number, constituting of two departures a week. Those departures are both during the weekends and are investigated separately from each other, although based on the same model. It is shown that the performances of those flights in a certain week are related to each other, at least to a certain extent. Most of the reporting done in this document is exclusively for the Saturday flight, which is the flight typically experiencing the highest demand, to reduce verbosity. The appendices however contain figures and tables concerning both flights.

The admission control is centralized to cover all requests regardless of originating sales station. Originally, the goals of the investigation referred to the free space specifically and exclusively, as they were formulated. However, optimizing flight revenue generated exclusively from the free space is technically hardly possible to start with, amongst others because the demand for free space capacity will show even larger volatility than the total demand. In my opinion, meaning to optimize revenue generated from just the free space capacity will require a forecasting model to condition on whether individual sales stations will overflow their assigned allotments. Hence, the total demand will have to be forecasted anyway. Since segmenting the capacity in allotments and optimizing for an allotment individually can be proven to be less optimal than to optimize revenue from all of the available capacity, the latter has later been set as the goal of the investigation.

Strategic or political targets of a specific sales station or customer, those related to building presence in a region for example, are ignored. Balancing requirements of this kind with optimizing flight revenue is today, and should remain, a management decision. However, in my opinion, such goals should not be formulated using a strict per-flight budget. The forecasting model built may aid in analyzing consequences of such a decision.

The revenue management solution proposed means to optimizing total flight revenue. No special attention has been paid to the cost and profit structure. Hence, the assumption is made that optimizing total flight revenue will also optimize operating profit. Most notably, trucking and interlining cost and administrative cost, those related to obtaining permits for example, are completely ignored. Much work has been done with regard to trucking cost already, but the information becoming available regarding this could not be used with the revenue management solution yet.

The research is using historical records of cargo that has been airlifted, exclusively. Records of requests that have been rejected for any reason and records of requests that are lost because of other reasons were unavailable. Due to this, the model underestimates total demand, if used without correction. A correction method is incorporated into the model, but this method can only be seen as a workaround most suitable in the absence of required historical records. A more precise and appropriate method is suggested for day-to-day use. This must be based on additional records yet to be collected. An alternative to collecting full records of requests rejected requiring less effort is discussed briefly in the chapter 2.4. Counting Space Request Arrivals.

The demand model constructed is quite generic. It ignores for example customer, region and commodity specific differences in the behaviour of space request properties. Inclusion of some a-priori knowledge of this kind might improve forecasting accuracy. How this can be done is hinted towards in the chapter 3.1. The Forecasting Model and 2.4. Counting Space Request Arrivals. The current implementation of the model is capable of estimating for multiple models and simulating multiple arrival streams that are later merged to this end, although this aspect of the implementation has not been fully evaluated or tested yet.

An alternative to the model presented might have been obtained by joining the arrival streams of the PO605 flights, possibly combined with other flights, and diverting booking requests onto one of the flights only after they arrive. This can be done using simple rules. The customer's preference must be accounted for with these rules. For example, it is well possible to record whether the customer is indifferent with regard to the exact flight their cargo is moving on and if not, on which flight the cargo has to be moved. In the former case, the customer may only be interested in getting the cargo at its destination before a certain date, which provides Polar with the opportunity to, for example, select the least constrained flight departing in time. This approach may be incorporated into the model laid out in this document relatively easy and would have been very attractive if, for example, a third PO605 would have been operated on, say, day 4.

This preceding model change would enable flights to be included into the model quickly after their inclusion into the flight schedule, with presumably acceptable forecasting quality. It may also help forecasting for extra sections. Please note, however, that in the current model setup, new flights may also be incorporated with ease by adding arrival streams for the new flight and simulating these in addition to the arrival streams for PO605/6 and PO605/7.

A prerequisite for incorporating other flights either way is that the space request properties do not deviate much between the flights already incorporated in the model and the flight to be added. Although research in this direction points towards this to be a save assumption, additional research should provide a definitive answer. Another PO605 can at least be safely added this way.

2. Discussion of Relevant Literature

2.1. Air Cargo Revenue Management

Revenue management is the practice of selecting those customers that generate the maximum revenue from a fixed and perishable capacity. For Polar, this means accepting those requests for space usage on a flight that will maximize the total revenue at take-off, subject to the weight and volume constraints of the aircraft. A trade off has to be made between accepting a request with certain revenue and waiting for a more profitable request that may or may not come.

There is a lot of literature about revenue management. However, not so much of it discusses the special case of its application to cargo transportation. It is to be noted that a general definition is not agreed upon, partially because of the widespread applicability of the principles involved. Please refer to Pak et al (2004) for a comparison of revenue management for passengers and cargo and an in-depth discussion of bid-price controls for air cargo revenue management. Refer to Hendricks et al for an outline on implementing revenue management throughout the enterprise from a business perspective.

For revenue management to be applicable, the following criteria must be met.

- Different people must be willing to pay a different price for a unit of capacity
- Capacity must be fixed and perishable
- Variable cost should be relatively low

In addition, demand should be random. In an air cargo environment, demand is highly random. Not in the last place because of the diversity in shipment sizes that are commonly airlifted. For the cargo only aircraft carrier Polar is, this is partially because freighters support such large shipment sizes. In addition, the special needs of the European cargo market are considered to reinforce this diversity.

Finally, Hendricks et al outlines the key criteria of having systems in place to provide centralized control of capacity and to record historical and real time information against which the revenue management system can operate. It should be no question that an adequate information technology infrastructure is paramount, especially when the sales apparatus is geographically diverted.

2.1.1. Willingness to Pay

If different people would not have been willing to pay a different price for a unit of capacity, it would not matter much to discriminate between requests, hence revenue management would be less useful. However, Hendricks et al mention aiming at reducing cost or the importance of individual clients might make revenue management functional.

Different people must be willing to pay a different price for a unit of capacity. Although this might be true, the willingness to pay is primarily dependent on the properties of the cargo to be transported and only in the second place on the customer requesting the capacity. This adds a theoretical argument for the diversity in willingness to pay.

It is important to note that a unit of capacity cannot be given a single number to indicate its size, because freighter capacity is two-dimensional. Both weight and volume are of importance. If measured with enough precision, every single shipment is different from any other. For this reason, weight and volume of a space request are taken to be continuous and random, as opposed to countable and fixed as seats on a passenger plane are.

Although a series of requests might generate the same revenue per kilogramme of cargo, their density is poised to be different. Hence, every single request also generates a different amount of revenue per unit of capacity.

2.1.2. Fixed and Perishable Capacity

If the capacity were not fixed, but could be changed inexpensively and quickly, the problem shifts from selecting customers to efficiently changing capacity. If the capacity were not perishable, the problem shifts to balancing turnover and load factor. The latter would require service level to be not an issue.

Capacity is fixed as long as a single departure is concerned. When the flight departs, the remaining unsold capacity cannot generate any more revenue and is considered perished. The capacity available depends on the following factors.

- Capacity may be in use for cargo not originating from the current take-off airport region. This might be normal sales traversing multiple hubs. In addition, capacity may be reserved for company materials or cargo booked on a previous scheduled departure is overflowed to the next flight, due to overbooking for example. These capacities might or might not generate an amount of revenue. It is an advantage if the model to be constructed can cope with these situations.
- The aircraft positioned to fly the service. The aircraft make and model are of primary influence to the typical maximum payload weight and unambiguously determines the maximum payload volume. Also, safety regulations and permits affecting maximum allowable payload are tied to the tail number of the aircraft positioned to flying the service the next departure.
- The flight distance to the next airport. Although almost all Polar flights are long haul, actual flight distances are quite different between different flights. In addition, individual departures may occasionally be routed differently although they fly under a well-known flight number. The amount of fuel required and the maximum fuel load determines the maximum payload weight for the aircraft.
- Forecasted weather conditions. The weather influences the amount of fuel burnt and may require the aircraft to fly an additional distance to avoid bad weather. This might reduce the maximum payload weight.

Although available capacity is thus uncertain up to a short time from departure, it is assumed to be known to be able to focus on the booking control policy.

If less capacity is sold than was available, it would have been beneficial to sell the remaining capacity at any price in excess of the low variable cost, preferring any revenue over nothing.

Another risk is that if low-revenue cargo is accepted early, higher-revenue requests might have to be turned down due to capacity constraints.
The ability to overflow cargo to the next flight exists, but is often undesirable from a customer service perspective and may still mean higher-revenue cargo is being turned down on either flight. Therefore, this ability is ignored in this paper. It is, however, possible to use common overbooking techniques in combination with a booking control policy as discussed here.

2.1.3. Low Variable Cost

As stated above, if less capacity is sold than was available, it would have been beneficial to sell the remaining capacity at any price in excess of the low variable cost.

For this reason, Koole (2004) and Venrooy (2002) state that for an effective application of revenue management, variable cost should be low, to ensure that selling a single unit of capacity generates as much profit as possible once the break even point has been met with regard to the constant cost.

The fact that constant cost are high means the risk is high, but low variable cost increase the potential for relatively large gains in net operating profit once the break-even point has been surpassed.

Due to the high cost of operating aircraft, relatively low profit margins and a large volatility in demand, the smallest increase in profitability is likely to contribute fully to the overall operating profit. It is for this reason that the aircraft industry has been at the cradle of revenue management.

2.2. Admission Control

After formulating the goal of selecting those space requests that maximize the expected revenue at takeoff and having forecasted demand for air cargo capacity, controls for discriminating between arriving space requests need to be established.

Space requests are evaluated against a single policy when they arrive and either admitted on the flight or rejected. This is called admission control. The discussion of admission control is focused on admission control using bid-prices. The concepts however generalize well to other forms of admission control. In this chapter, admission control using bid-prices is explained and an implementation suitable for practical use is detailed.

2.2.1. Bid-Price Admission Control Policies

The idea of bid-price admission control policies is to determine a value for which a unit of capacity can be sold at a certain point in time. This value is the bid-price.

The revenue a customer has offered to pay for using a part of the remaining capacity is called the direct revenue. It is denoted by r in this chapter. This is called direct revenue because if the request is accepted on the flight, the revenue realized on the flight immediately increases with that amount.

If an amount of weight and volume is still available on a flight, the best revenue still achievable with that remaining capacity is called the optimal expected revenue. This revenue is still to be realized.

- At any point in time, this value is equal to or larger than zero.
- Moreover, while time passes, this value will decrease and reach zero when the flight departs. This expresses the fact the capacity is perishable.
- Finally, this value decreases whenever either the remaining weight or the remaining volume decreases due to a shipment having been accepted on the flight. If either the maximum payload weight or the maximum volume for the flight has been reached, the optimal expected revenue will also become zero.

A general bid-price acceptance policy works by comparing the direct revenue of a request with weight m and volume v with the reduction in optimal expected revenue due to the remaining capacities being reduced by m and v. This reduction in optimal expected revenue is called the opportunity cost.

The baseline of admission control is that a request should only be accepted if and only if the direct revenue r earned with accepting the request equals or exceeds its opportunity cost. If the optimal expected revenue before accepting a shipment exceeds the optimal expected revenue after accepting a shipment plus the direct revenue of that shipment, the shipment should be rejected.

> **Note** Requests should only be accepted if the direct revenue earned with accepting the request equals or exceeds its opportunity cost.

Because capacity on a freighter aircraft is two dimensional, a bid-price is constructed for each capacity dimension. The opportunity cost of a booking request can be approximated by the sum of the bid-prices of the capacities it uses (Pak et al, 2004, page 7).

> **Note** The opportunity cost of a booking request can be approximated by the sum of the bid-prices of the capacities it uses in each capacity dimension.

Bid prices are studied extensively for passenger revenue management by Talluri and van Ryzin (1998). They show that a static bid-price policy is asymptotically optimal when the capacities and the demand increase proportionally and the bid-prices are set correctly. This means the following.

- If the bid-prices are correct, the direct revenue of a shipment can be evaluated against opportunity cost computed using bid-prices. These do not change over time. They are static.
- A static bid-price policy will have the effect of accepting all requests that arrive during the booking period generating revenue per unit of capacity in excess of a fixed threshold. For this reason, a static bid-price policy is called the greedy solution.
- If all space requests where the same size, the greedy solution would definitely be optimal. The fact that space request sizes differ causes a part of the capacity to remain unused due to no requests arriving anymore that fit in that last, hopefully small, part of the aircraft. Smaller

requests that would fit may have been turned down earlier because they did not generate direct revenue in excess of the bid-price threshold.

- In a perfect world, space requests are small relative to a freighters capacity. If this were the case for all requests, high-revenue large requests would not be turned down just because they do not fit. It is evidently suboptimal to have to turn down requests due to the plane being too small. Hence, the mentioned optimality condition holds.

Pak et all (2004) argue why the asymptotic result mentioned holds for the 2-dimensional problem as well as for the single dimensional problem as demonstrated by Kan et al (1993). That is, also for the cargo revenue management problem, Pak et al (2004, page 8) show a static bid-price policy is asymptotically optimal as demand and capacity increase proportionally and the bid-prices are chosen correctly.

2.2.2. Bid-Prices for Cargo

Let r, m and v be the revenue, weight and volume of a space request. We want to select those space requests that do not use a lot of capacity for the profit they generate. That is, we want to select those space requests for which m/r and v/r are small.

The idea is to weight each capacity dimension and select those requests that have the highest profit per weighted capacity requirement. Each capacity dimension is given a weight ratio, whose relative value corresponds with how restricted the capacity dimension is compared to another.

Let a_m and a_v be the weight ratio for the weight and volume dimensions of capacity. Space requests to be accepted are the ones that have the highest ratio

$$d = r / (a_m m + a_v v)$$

For a given set of space requests, one can simply select the requests that have the highest ratio d. If the future booking requests are not known, a threshold value z has to be specified in advance. Then, when a request arrives, it is accepted if its ratio d exceeds the threshold z. That is, if

$$r / (a_m m + a_v v) >= z$$

Provided all numbers in the above inequality are non-negative, this is equivalent to

$$r >= z\ (a_m m + a_v v) = za_m m + za_v v$$

This means the above can be seen as a bid-price policy, with bid-prices $b_m = za_m$ and $b_v = za_v$ for the weight and volume capacity dimension, respectively.

The algorithm discussed in the next section provides a way to obtain the optimal bid-prices for a given set of booking requests. In practice, however, the bid-prices have to be set before the demand is known. Bid-prices can then be obtained by simulating a series of arrival streams, computing the optimal bid-prices for each of these using the algorithm below and then setting the bid-prices to the average value over all sequences.

2.2.3. Constructing Bid-Prices

We want to select those space requests that do not use a lot of capacity for the profit they generate. That is, we want to select those space requests for which m/r and v/r are small. Because of this, the algorithm is called greedy.

For a single flight, a given set of requests may be graphically represented by points with coordinates $(x=m/r;\ y=v/r)$ on the flat plane. The optimal bid-prices for weight and volume are placed on the horizontal and vertical axis at za_m and za_v, respectively, and connected with each other. The bid-price acceptance policy now dictates that requests whose points lie at or below this line are to be accepted and requests whose points lie above this line are to be rejected.

Determining bid-prices is done by repeatedly choosing a slope for the line and moving it upward from the origin, accepting requests in the order they are encountered, until no more requests fit. The optimal bid-price is the one resulting in the highest revenue when using the bid-price under evaluation.

In order to do this effectively, it is important to note that two slopes are practically identical, if they have the effect of accepting requests in the same order. Therefore, only slopes that actually change the ordering of requests are to be considered.

Lenstra et al (1982) provide an algorithm to determine all possible orderings. The algorithm makes use of the fact that the line that passes through two requests provides exactly that slope for which those requests swap places in the ordering. In order to obtain all orderings, we only have to evaluate those lines that connect two booking requests. In addition, only negative slopes are to be considered. Finally, when iterating all requests, we only have to consider pairs of requests once. This can be achieved by ignoring request pairs of which the second one is above and to the left of the first one.

For each ordering, the solution of the greedy algorithm can be computed by accepting the requests until the capacity is full and recording the slope and height of the line at that point. If the slope is β and the intercept of the line with the y-axis is α, the bid-price for the weight dimension is $-\beta/\alpha$ and the bid-price for the volume dimension is $1/\alpha$.

The chapter 3.2. Determining Bid-Prices includes graphics to visualize this method and describes the algorithm in more detail and implementation-ready.

2.2.4. Multiple Flights

An interesting aspect of the algorithm laid out above is that it is equally well suited for multiple flights. The changes it requires to accommodate for multiple flights are explained here.

All variables defined may be considered scalar values for the algorithm to be used for a single flight. In the case of n flights, the revenue of each request is still the scalar value r, but its weight m and volume v become vectors with n coordinates. Each coordinate now specifies the capacity usage requirement on one of the n flights. If a certain flight is not used, the corresponding coordinate in m as well as v must be zero.

Also, a_m and a_v both become vectors of n relative weights for the weight and volume capacity dimensions of the n flights, respectively. b_m and b_v then both become vectors of n bid-prices. The threshold value z remains a scalar value.

Finally, the products $a_m m$ and $a_v v$ are meant to be inproducts and produce a scalar value. The opportunity cost is computed for all flights the request uses and are then combined by summing them up. Opportunity cost of unused flights will be zero. The total revenue earned must again at least match the total opportunity cost.

> **Example** For flight i, the ith coordinate of a_v is multiplied with the ith coordinate of v. Those individual products are computed for all n flights and summed up to produce the in-product of a_v with v.

Although a graphical representation becomes difficult if $n > 1$, the algorithm detailed by Lenstra et al (1982) can easily be extended to handle multiple flights. Note that when the threshold value z is increased, the bid prices for all flights are increased with it. The relative weights of the capacity dimensions are meant to express differences, not only in bid-prices for different capacity dimensions, but for different flights, too.

2.3. Demand Forecasting

Statistics is the art of modelling situations, often referred to as experiments, in which chance plays a role and of drawing conclusions from realizations of such experiments. (Oosterhof and Van der Vaart, 2003, page 2)

Two forms of statistics should be distinguished. Descriptive statistics concerns collecting and summarizing of data in a transparent manner. Mathematical statistics concerns methods to analyse data, based on probability models. Mathematical statistics therefore is strongly connected to and makes extensive use of concepts from probability theory. Together with the non-deterministic decision theory, these three fields are called stochastics (*stochastiek)*.

In probability theory and mathematical statistics, experiments play a central role. Probability theory defines so called random variables (*stochastische variabelen*), which are connected to the possible outcomes an experiment may produce. In statistics, outcomes of an experiment are called realizations of the random variable. Of primary importance is that those random variables have a probability distribution (*kansverdeling*). This distribution defines with what probability the random variable will produce outcomes within a specific range of all possible outcomes.

> **Example** Suppose the experiment is to roll a fair dice once. The random variable X we are interested in is the number of eyes turning up. The probability distribution of X is dependent on the experiment as well as the result of interest that it measures.

Outcomes of an experiment are called realizations of a random variable. Suppose, however, the probability distribution of this variable is unknown. With mathematical statistics, the real distribution of the random variable is to be determined. Then we can forecast new realizations and explain historical outcomes.

2.3.1. Frequency quotients

The nature of experiments in the context of stochastics is such that outcomes are produced with a great deal of randomness. It might be surprising that such a process is still systematically driven. This can be concluded when an experiment is repeated under constant conditions and its successive outcomes are used to produce a series of frequency quotients. This frequency quotient is the fraction of experiments producing a certain outcome divided by the number of experiments completed up to that point. (Harn and Holewijn, 1997, page 7)

Successive frequency quotients will after just a few experiments show large fluctuations, but when the number of experiments grows, the frequency quotient will converge towards a constant number and settle. This is a repeatable result for any experiment and is called *the empirical law of large numbers*.

The set of possible outcomes of an experiment is called the state space (*uitkomstenruimte*) or support. Sometimes, not just a specific event is of interest, but rather a subset of the set of possible events, like in the example below. Such a subset is called an eventuality (*eventualiteit*) in Dutch. It would be intuitive to define the probability of an eventuality to be the value of the frequency quotient after repeating the experiment infinitely often, allowing it to fully settle. This value of the frequency quotient is called the *limiting probability*.

> **Example** Throwing a fair dice will produce an even number (2, 4 or 6) on average every second roll. The more rolls we perform, the more clear this will be visible. Because we know the probability distribution, we know the limiting probability the quotient converges to is 1/2.

Probabilities are often written shortened using the measure P. Such a measure assigns probabilities to eventualities. If Ω is the state space of a random variable X and A is a subset of Ω, $P(X \in A)$ is read as 'the probability of X in A' and $P(A)$ is read as 'the probability of A' (*de kans op A*).

The following three axioms are the basis for the probability theory[2].

- The probability that the outcome of the experiment falls within its state space is 1.

---

[2] In short, $P(\Omega) = 1$, for any $A \in \Omega$, $P(A) \geq 0$ holds and for disjoint eventualities $A_1 \ldots A_n$ the relationship $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$ holds.

- Given any eventuality, the probability that the outcome of the experiment falls within the eventuality is zero or a positive real number.
- The probability of the union (*vereniging*) of a set of disjoint eventualities (having no values in common) is equal to the sum of the probabilities of those disjoint eventualities individually.

They are called axioms because they are left without formal proof. They are truisms or clichés.

> **Example** The state space Ω of throwing a fair dice is 1…6. The frequency quotient of this eventuality is fixed to be 1 because 100% of all experiments will produce outcomes in this range. If we take a subset of 1…6, the frequency quotient is zero for the empty set, or positive for any other subset of 1…6. The frequency quotient of the eventuality of throwing an even number is the sum of the frequency quotients of throwing 2, 4 and 6.

Eventually, those axioms will lead to quite powerful theorems, amongst which both the weak and strong laws of large numbers and the *Central Limit Theorem*, which support the choice of modelling of probabilities based on frequency quotients. Those theorems have tremendous influence in today's application of stochastics.

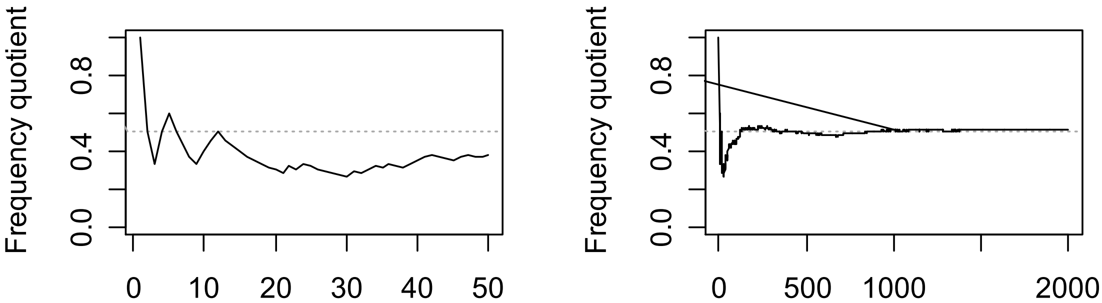


Figure 4 Subsequent frequency quotients of a series of throws with a dice, observing the fraction of throws producing an even number, with left the first 50, right the first 2000 quotients.

2.3.2. The probability distribution

As said, in mathematical statistics, it is of primary importance that random variables have a probability distribution. In probability theory, the probability distribution is defined as a function. The distribution function details with what probability an outcome of an experiment will be equal to or smaller than a fixed value.

> **Definition** If the function F is the distribution function of the random variable X, it equals the probability that X is smaller than or equal to its argument x. In short, $F(x) = P(X \leq x)$.

Every distribution function starts at 0 and ends at a value of 1. In between, it never decreases. Put differently, while x gets bigger, $F(x)$ never produces smaller probabilities.

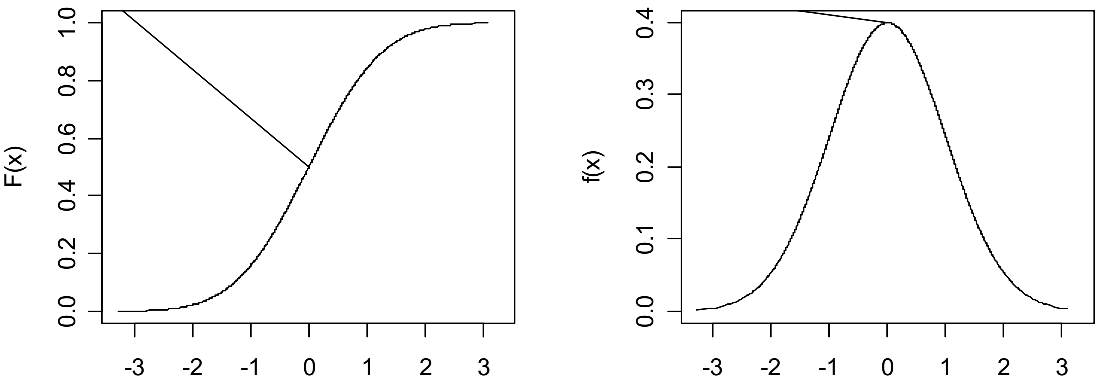


Figure 5 The standard normal distribution, with its distribution function (left) and its density function (right).

The distribution function is closely related to the density function (*kansdichtheid*). It is defined to be the derivative (*afgeleide*) of the distribution function. Therefore, if the density function is known, the distribution function can be obtained by determining the surface below the density function.

> **Definition** If the function f is the density of the random variable X, then X is said to be distributed according to f, in short $X \sim f$. Moreover, the distribution F can be obtained by taking $F(x)$ to be the surface below f left of and up to x.

Density functions nowhere return negative values; they are either zero or positive for any x. They can, however, produce values infinitely much larger than 1. Nevertheless, their surface is required to be 1 on their support (*domein*) which is required to be the state space Ω.

An important distinction is to be made between random variables that only take on a limited, countable number of possible outcomes and ones that take on an unlimited number of possible outcomes. The former are said to have a discrete distribution, the latter are said to have a continuous distribution.

> **Example** If X is the number of eyes turning up when throwing a fair dice, X has a discrete distribution, because it can only take on 6 different values.
> If X is the result of drawing a random number between zero and one, it has a continuous distribution, because there is an unlimited number of distinct values it may take on.

> **Example** In this paper, weight, volume and revenue of a single space request are random variables assumed to have continuous distributions, whereas the number of requests arriving during a booking period has a discrete distribution.

For discrete distributions, determining the surface below the density function corresponds to summing up probabilities. In fact, the distribution function is commonly called the cumulative distribution function. For continuous distributions, determining the surface requires integrating the density function[3].

> **Note** Although this might be intuitive, it is usually an error to interpret the values of a continuous density function as real probabilities. Probabilities never exceed 1 to start with.

The density function uniquely defines the distribution of the random variable. It may be used to determine all kinds of interesting information, like its expectation, standard deviation (*spreiding*), safety margins (*betrouwbaarheidsinterval*), conditional probabilities and more.

> **Example** If X is the number of eyes turning up when throwing a fair dice once, then its density is 1/6 at 1, 2, 3, 4, 5 and 6, it is zero elsewhere. Hence its distribution $F(x) = x/6$ for $x \in \{0 \ldots 6\}$. Below 0 it is 0 and above 6 it is 1. For floating point z, $F(z)$ equals $F(x)$ for the largest $x \in \{0 \ldots 6\}$ smaller than z. Recall that $F(x) = P(X \leq x)$, so for example $F(0.5) = F(0)$.

> **Example** If X is the result of drawing a random number between 0 and 1, its density $f(x)$ is constant for x between 0 and 1, because each number between 0 and 1 is as likely as any other to be observed. Therefore, X is said to have a uniform distribution. The value of the constant follows from the requirement that the surface below the density function has to be 1. Hence $f(x) = 1$ for x between 0 and 1 and its distribution $F(x) = x$ for x between 0 and 1.

2.3.3. The empirical distribution

Data that has been collected by repeatedly measuring the outcome of an experiment is called empirical data. This data may be used to construct the empirical distribution function. The empirical distribution function is based on frequency quotients.

> **Definition** If $X_1 \ldots X_n$ is a series of n realizations of a random variable X and F_n is its empirical distribution function, F_n equals the fraction of observations smaller than or equal to its argument x. In short, $F_n(x) = 1/n \times \#(X_1 \ldots X_n \leq x)$.

The similarity between the probability distribution function and the empirical distribution function lies within the fact that both measure $P(X \leq x)$. Consistent with single frequency quotients converging towards the so-called limiting probabilities, the empirical distribution will eventually look like the real distribution of X. The objective of mathematical statistics is to determine the real distribution of a random variable connected to an experiment.

> **Example** The empirical distribution function $F(x)$ of the weight of a series of airway bills recorded in a database details the fraction of airway bills in the database having weight x or weighting less.

---

[3] The existence of continuous distributions that are not absolute continuous is ignored for clarity.

The empirical distribution is a discrete distribution, because it is always based on a limited number of observations. This is visible because the empirical distribution function is a stair function. Its value jumps straight up a distance of 1/*n* at values that are observed in the empirical data and is horizontal at all other values in the state space that were not observed.

2.3.4. Quantiles and simulation

Simulation is the process of generating realizations without actually executing an experiment. It has become a very useful technique in many ways and in many situations thanks to the broad availability of fast computers.

Simulation is a suitable tool when models are very complex, making it very hard to give a closed form of the desired distribution. A complex experiment may be decomposable into smaller parts. These smaller parts, being experiments, can then be simulated and the resulting realizations put together using simple rules to give a single realization of the random variable of interest.

> **Example** Simulating a single flights payload weight comprises simulating a series of space requests arriving and summing their respective weights after they are admitted on the flight.

Simulation is done by generating probabilities between 0 and 1 using a computers random number generator and then computing what realization would correspond to this probability according to the distribution function.

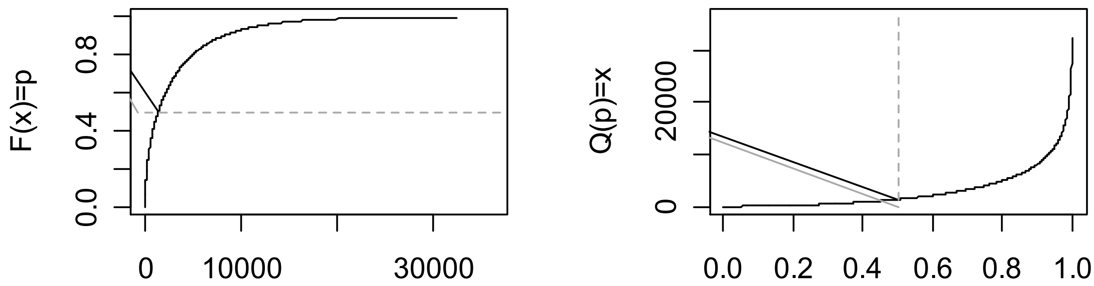


Figure 6 The gamma (½, 6000) distribution. The weight of a space request has a similar distribution. Its distribution function is shown left and its quantile function right. A dotted line is drawn at the median.

The distribution function specifies what probability corresponds to a certain value the random variable may take on, so its inverse details what value corresponds to a certain probability. This value is now called a quantile and the inverse of the distribution function is called the quantile function.

> **Note** The quantile function *Q(p)* is the inverse of the distribution function *F(x)*. If *F(x) = p* then *Q(p) = x*. Upper quantiles are used in this document for parts of *F* that are horizontal. The upper *p*-quantile of *F* is defined to be the largest value of *F* for which *F(x) = p*.

The graphic of the distribution function of *X* may be used to simplify this. Computing a probability corresponds to selecting a value on the horizontal axis and reading the probability on the vertical axis. Computing a quantile is done by selecting a probability on the vertical axis and reading the corresponding realization from the horizontal axis. If the distribution function is vertical at this probability, the biggest value is used. This *p*-quantile is therefore also called the upper *p*-quantile.

> **Example** The median of a sample is in essence the realization in the sample that has as much bigger realizations above it as smaller ones below it. Hence, the probability that a realization is smaller than the median is 50%, and its value may be read from the empirical distribution function graphic at distance 1/2 from the horizontal axis. In fact, it is often defined to be the 50% quantile. However, if the distribution is vertical here, the true sample median is not the upper *p*-quantile, but the average between the upper and lower *p*-quantiles.

2.3.5. Statistical models

Up to a certain extent, the direction of statistics is exactly opposite of that of the probability theory. Instead of knowing the probability distribution and doing computations with it, some results of the experiment are known, but the underlying probability model is completely or partially unknown and has to be derived from the observations. A statistical model is therefore defined to be a collection of all distributions that are considered possible for the realizations observed[4]; the statistical model is an expression of our knowledge of the experiment that led to recording the observations.

> **Example** Measurement errors in physics are often modelled to have a normal distribution with a mean of zero and a standard deviation of some $\sigma > 0$. The statistical model describing measurements errors in physics is hence the collection of all normal distributions with zero means and positive standard deviations.

Suppose the distribution of a random variable has been estimated. This estimate is never exact and hence estimation errors are introduced. These are equal to the difference between the observed outcome and the value that should have been the outcome based on the distribution function estimated. This means they are seen as measurement errors. Those errors, called residuals, are often required to be asymptotically distributed with the well-known normal distribution because of the central limit theorem referred to earlier.

If this turns out to be really the case, this is a strong indication that there is little reason to doubt the correctness of the assumptions made in the statistical model of the original experiment.

[4] The concept of probability measures is ignored for clarity.

2.4. Counting Space Request Arrivals

While forecasting demand for freighter capacity, it is necessary to forecast the number of space requests that will be offered by customers. When a customer submits a space request, its weight, volume and revenue are of course recorded. However, the number of space requests offered during the pre-flight booking period will ultimately dictate overall flight performance in concert with the properties of individual space requests.

The number of space requests offered during the pre-flight booking period is regarded to be of instrumental impact on flight performance. Whereas the behaviour of the properties of individual space requests is poised to change only slowly over more extended periods, this number might even change overnight under a host of influences of which seasonality is only one. Therefore, extra attention is paid to modelling of the number of space requests offered.

This chapter discusses some theory about counting processes, which deal with counting events called arrivals, and the assumptions commonly made when modelling how arrivals occur. The discussion of the theory is largely based on Harn and Holewijn (1997) and Ross (1972). Since we are interested in the number of space requests that will be offered, we say we count space request arrivals.

2.4.1. Counting processes

Counting space request arrivals is a counting process. The state of a counting process is recorded by a stochastic variable, say N, at each point in time. This is written as N_t. The value of N_t denotes the number of arrivals recorded from the start up to a certain time t. This count of the number of arrivals adheres to the following rules.

- It starts at zero.
- It is never decremented.
- When incremented, it is incremented with a positive, whole number, usually 1.

The time between two arrivals is called the inter-arrival time (*tussenaankomsttijd*), or waiting time. Rather than this being regarded a fixed amount of time, the inter-arrival time is in most useful models subject to uncertainty. Therefore, it may be measured repeatedly in an experiment and modelled to have a probability distribution.

The average number of arrivals during a single unit of time is called the arrival rate (*aankomst intensiteit*).

The following two important properties counting processes might have can now be formulated.

- A counting process has *independent arrivals* if the number of arrivals during non-overlapping intervals is independent. This means observing a number of arrivals during a certain hour provides no information as to when or how often arrivals occur during the next hour.
- A counting process has *stationary increments* if the arrival rate is independent of time. This means there is no trend with regard to the speed at which arrivals occur.

This is for the following reasons.

- If the number of space requests arriving in the morning is independent of the number later on the day (or take any other two distinct parts of the day for that matter), the process has independent arrivals.
- If the inter-arrival time distribution remains the same during the day, the process has stationary increments regardless of whether space requests arrive evenly distributed throughout the day or not. Individual inter-arrival times vary due to chance rather than due to differences in the way they are generated.

Several arguments support modelling the arrival of space requests to have the above properties.

- Customers are indifferent about what time of day they file a request.
- Customers will not coordinate with each other the time they file a request.
- Customers will not coordinate with each other in what order they file a request.

2.4.2. The Poisson process

The only counting process having both independent arrivals and stationary increments is the Poisson process. This can be proven with the fact that, for its arrivals to be truly independent, its inter-arrival time distribution must be memoryless (*geheugen vrij*). The amount of time elapsed since the last arrival must provide no information about the amount of time until the next arrival.

> **Example** Cheap ballpoints are known to contain just enough ink for only a certain amount of writing. When counting the number of ballpoints worn off, knowing that the one currently in use has already served for several weeks provides information that it certainly will run out much faster than a new one.

There is only one distribution being memoryless. This is the exponential distribution. Its default interpretation is the lifetime of a part experiencing practically no wear off, such as a bulb. The amount of time the part has been in service is of no practical influence on its remaining lifetime.

A counting process is a Poisson process with rate L, with $L > 0$, if its inter-arrival time has the exponential distribution with rate L and arrivals occur one at a time. If the latter is not true and arrivals are allowed to arrive in batches, the process is called a compound (*samengesteld*) Poisson process.

> **Note** Space requests are assumed to arrive one at a time.

This has the following equivalent implications.

- Its inter-arrival time distribution is memoryless, and hence the counting process has independent arrivals.
- The number of arrivals during any interval of length s is Poisson distributed with mean sL, regardless of when that interval starts, so it also has stationary increments.
- On average, L arrivals occur during a single time unit. Put differently, arrivals are on average $1/L$ time units apart.
- The risk of a request arriving in the next short time interval is constant.

The failure rate function of a distribution is defined as the instantaneous rate of failure at time t during the next instant of time. For the inter-arrival time distribution, it returns the risk of a request arriving in the next second or so. The failure rate function of a distribution also defines it. The only one that is constant belongs to the exponential distribution.

> **Example** Aircraft engines are known to be highly reliable. As long as one is serviced regularly, it operates as if it were new; due to their high reliability, a breakdown may be attributed to bad luck rather than being increasingly expected. Hence, the probability that it breaks down in the near future is equally big just after a service than after extended periods of operation.

The Poisson process and the exponential distribution take a central place in reliability and renewal theory and are widespread in their use in modelling for example births in a city, the number of animals in a certain area, radiation in physics and modelling of other highly random events.

2.4.3. Practical consequences

The number of arrivals during any interval of length s is Poisson distributed with mean sL. Hence, the average number of arrivals observed during a time interval, divided by the length of the time interval may be used to define the counting process.

Suppose it takes time to respond to a space request. The average number of requests being processed equals the expected response time multiplied with the arrival rate. This is the formula of Little and is often referred to with PASTA (Poisson Arrivals See Time Averages). As this abbreviation suggests, the formula of Little is valid for systems having inter-arrival times that are exponentially distributed (Poisson arrivals). Service times may be exponential, but can be of a more general form, e.g. constant.

The sum of two Poisson distributed random variables is again Poisson distributed; merging two Poisson processes with rates L_1 and L_2 produces a Poisson process with rate $L = L_1 + L_2$. This is a result of the arrivals of each process separately, but also joined, being truly independent.

> **Note** Please note that the arrival rates of the two Poisson processes have to be aligned. If one is based on average arrivals per day and another on average arrivals per week, the arrival rate of the former must be multiplied by seven.

Arrivals of space requests from a particular customer may be separated from the main arrival stream. It might be useful to change the expected flight performance based on whether a particular customer has offered any space requests yet or not. Often, however, customers that book once very often might book not at all or multiple times occasionally. If this is the case, it is well possible the customers booking behaviour is better described with a Poisson process having a low rate, instead of conditioning on the customer having offered a space request.

Note Although arrivals of individual customers will often turn out to be Poisson distributed, they do not have to be. On the other hand, if they are not, this does not directly mean the overall arrival stream is incorrectly modelled to be Poisson. The effects of the arrivals of a particular customer not being Poisson might be masqueraded.

Suppose it turns out a certain fraction p of all space requests are rejected. Because we do not know the properties of the requests before they have arrived, the probability that this happens might be considered fixed during a whole day, especially when the flight is nowhere near full yet. If the main arrival stream has rate L, requests will be rejected with rate pL and accepted with rate $(1 - p)L$. This provides a good starting point for recording data on rejections, especially if requests rejected due to insufficient capacity available are left out of this model. This approach is also useful for modelling cancellations and no-shows.

For a Poisson process with rate L, the expected number of arrivals during a time interval with length s is sL, but its variance is sL, too. Hence, its standard deviation is the square root of sL.

Note To get an indication of whether arrivals occur according to a Poisson process quickly, compare mean and variance of a series of arrivals during for example a series of days. If they are nowhere near one another, the assumption of Poisson arrivals might be wrong.

Requests in a fixed interval of certain length, for example a day, are spread out the same as a random drawing from the uniform distribution. This has the following results.

- Arrivals tend to group together.
- Personnel handling arrivals will notice strong variations in workload compared to average workload during a day.
- There is one or more inter-arrival time that is likely to be (much) bigger than the average inter-arrival time.
- When starting at a random time, waiting for the next arrival, one is more likely to start waiting during a longer inter-arrival time. Hence, by recording the time elapsed until the next arrival this way, the arrival speed is under estimated for it seems the inter-arrival time is longer than expected. This is called the inspection paradox.

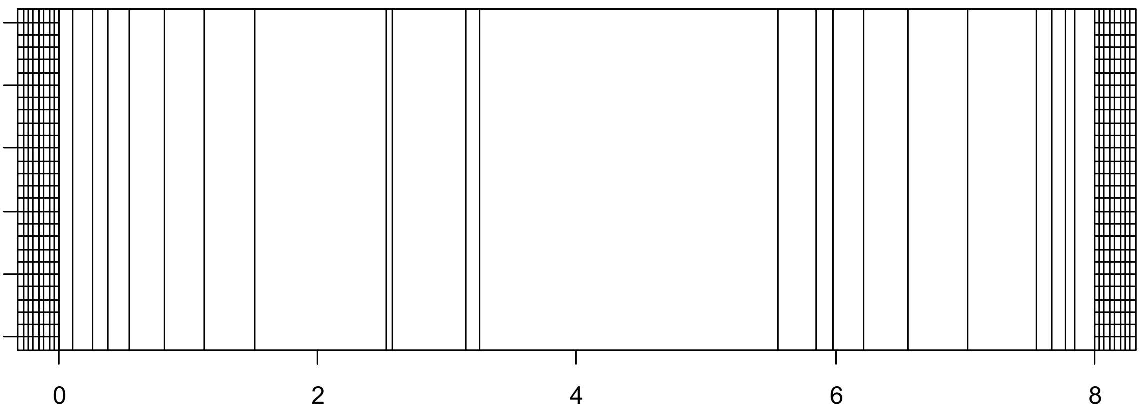


Figure 7 Arrivals simulated with rate 16 per working day plotted on a timeline.
There are 22 arrivals recorded.

2.4.4. The Binomial Process

Although the homogenous Poisson process is the only counting process with independent arrivals and stationary increments, and the exponential distribution is the only memoryless distribution, a process exists which defines those properties in a discrete environment. This process is the binomial process. A simple experiment of throwing coins may be constructed that follows its rules.

Instead of measuring time continuously as we are used to, it may also be measured in terms of turns as with a board game. Every turn, the game time is advanced with one turn. How long a turn exactly takes is defined to be a fixed amount of time, like an hour. This way, time has become discrete and countable.

Suppose the same unfair coin is thrown once every turn. An unfair coin produces heads with a probability different from 1/2. If it produces heads, a score is recorded. Otherwise, the turn passes. A counting process is used to keep track of the score.

- The counting process has independent arrivals when we measure time in terms of turns, because the scores awarded in a number of consecutive turns (which is a discrete time interval) is only dependant on the number of turns. It is not dependent on whether scores were recorded during other turns or on how much turns ago it was that a score was recorded for the last time.
- The counting process has stationary increments, because the same coin is thrown every turn exactly once. So the probability to score during any turn remains the same, regardless of how long the game is in progress.

The number of turns required to score has a geometric distribution, because the experiment matches the default interpretation of this distribution exactly. It is interpreted as the number of independent Bernoulli trials, each with success chance p, which is required to reach the first success.

> **Example** Suppose the coin used produces heads when thrown with probability p and tails with probability $q = 1 - p$. Let X be the number of turns required to throw heads.
>
> The probability that we need more than k throws to throw heads for the first time equals the probability that we throw tails the first k throws. Thus $P(X > k) = q^k$; since throwing tails has probability q, throwing it twice in a row has probability $q \times q = q^2$ and throwing tails k times in a row happens with probability q^k.
>
> Now, suppose we have been throwing tails k times in a row from the start.
> The probability to throw tails at least the j throws following those k, too, equals
> $P(X > k + j \mid X > k) = q^{k+j} / q^k = q^j$
> which equals the probability to throw tails in the first j throws[5].
>
> The probability of throwing tails in the next j throws is equally big regardless of the value of k, being the number of throws already observed to produce tails. It is independent of past performance, and hence is called memoryless or memory free (in a discrete sense).

Recall that turns are defined to take a fixed amount of time. The Poisson process may be obtained from the above experiment, if the number of turns is not too small and p is not too large, by performing the following steps.

- Let the number of turns grow very large.
- Keep the total game time (measured as real time) constant.
- Reduce the chance to score in one turn proportional with the reduction in turn length.

Put otherwise, if n is the number of turns, s is the turn length and p is the probability to score in one turn, $n \times p$ is the average score during each such a game and $n \times s$ is the total game time.

Let n go to infinity while $n \times p$ and $n \times s$ are kept the same. Both p and s will become very small. The value of $n \times p$ is related to the arrival rate of the resulting Poisson process, since it indicates the scoring frequency per unit of real time. The approximation is better starting with small values of p and large values of n.

---

[5] Conditional probabilities: when throwing a dice, throwing 6 has probability 1/6, but if we were told we threw an even number, it would be 1/3, because $P(X = 6 \mid X \text{ even}) = P(X = 6) / P(X \text{ even}) = 1/6 / 1/2 = 1/3$. This formula, when generalized, holds only for independent pairs of eventualities.

3. Implementing the Revenue Management Solution

3.1. The Forecasting Model

3.1.1. Counting Space Requests Arriving during the Booking Period

The pre-flight booking period is divided into eight labelled intervals in accordance with the day-to-day practice at Polar. Those are the seven days of the week, scheduled flight date being T-0 and the day after the previous flight, six days before scheduled flight date, being T-6. The period before this day is the pre-booking period and space requests arriving in this period are called pre-bookings. It is sometimes useful to indicate this period as being yet another day, labelled T-7, although this virtual day extends as long as necessary into history. Pre-bookings are those bookings that occur a week or more before scheduled flight date and for which the customer specifically instructs Polar to lift the cargo on another than the next flight.

Space requests arriving during the seven pre-flight days are modelled to do so one-by-one, at a constant rate throughout that day and independent of another. The arrival rate is the expected number of arrivals during the whole day. For a single flight, arrival rates are different for each day. In addition, the number of pre-bookings is modelled similarly, with a mean matching the expected number of pre-bookings.

The modelling of how space requests arrive, the theory underlying this modelling and the practical consequences of the choices made are explained and defended in the chapter 2.4. Counting Space Request Arrivals. The chapter 1. Space Request Arrival Rates discusses the results of the statistical analysis of the arrivals.

3.1.2. Modelling Space Requests

Space requests are modelled with the following properties.

- The density classification indicates the cargo is below, exactly at or above pivot density.
- The actual weight is the weight of the cargo as it is measured on a scale.
- The density allows the volume of the request to be computed, using the actual weight.
- The revenue modelled is the total revenue excluding fuel and security surcharges.

Each space request has a density classification independent of any other property of the request. The density classification is explicitly modelled, because the weight, density, volume and revenue of a request behave quite differently depending on its density classification.

The modelling of the density of a space request is especially different for each density classification. If a shipment is classified as cargo at pivot, its density is 6, since Polar uses a pivot ratio of 1 tonne per 6 m^3. If it is classified as dense cargo, its density has values in a range from 0 to 6. The density of volume cargo is larger than 6. The modelling of densities is quite different for dense and volume cargo, apart from the range in which the density lies.

The revenue modelled is revenue excluding fuel and security surcharges, because these surcharges are strictly proportional with the actual weight of the shipment and can thus be computed. The reason why space request volume is not explicitly modelled is that it can be computed when the density is known and density may be modelled more accurately than is possible with total request volume.

The results of the statistical analysis, as well as the exact modelling of the space request properties, are discussed in section 4. Statistical Analysis, in the chapters 4.2. Density Classification, 4.3. Actual Weight, 4.4. Density and Volume and 4.5. Revenue. In addition, the influence of time on these properties is made visible in the chapter 4.6. Trend and Seasonality.

3.1.3. Fine-Tuning the Model

The above properties of space requests are the only ones that are modelled. For this reason, the forecasting model is quite generic. It ignores, for example, possible differences between space requests originating from different regions, offered by different customers or containing different commodities. The model can however be extended by incorporating such knowledge into it, without requiring to incorporate additional variables for all space requests. In this section, space requests from a particular customer[6] are separately modelled to demonstrate the opportunity of improving the model and doing so highly selectively.

Suppose requests from a particular customer are all for movement of aircraft engines which have similar weights and volumes. In addition, those shipments are all accepted at a fixed rate per tonne chargeable weight. Finally, requests from this customer arrive at an almost constant rate throughout the year. Knowing well what this customer wants to have airlifted provides an opportunity to fine-tune the model.

> **Note** Any definite a-priori knowledge of a certain customer's behaviour or of properties of the space request the customer offers is candidate for incorporating separately into the model, depending on whether doing so improves accuracy, or well-informedness.

For this customer, space requests are verified to arrive according to a Poisson process. Therefore, the rates at which requests originating from this customer arrive on each day are deducted from the arrival rates estimated for booking periods T-0 to T-7. Because the arrival rates are quite low, they are combined into a single, weekly arrival rate.

> **Note** If a particular customer is known to offer space requests at or around a fixed time, this a-priori knowledge should be modelled, rather than using Poisson arrivals. However, be sure not to neglect the fact that although there can be expectations as to when requests arrive, this does not at all mean the time of arrival is suitably modelled when it is regarded fixed.

All space requests from this customer weigh from 5500 kg up to 7000 kg. They are known to take up around 38 m^3 on the aircraft. Their revenue is a fixed rate per tonne. Those requests originating from this customer having the properties mentioned are excluded from the data used to estimate parameters for distributions of other space request.

When simulating arrival streams to obtain bid-prices, space requests from this customer will be simulated separately and then merged with the list of space requests obtained using generic means. Note that the order of arrival of space requests is of importance, especially near scheduled flight date. For this reason, arrivals are simulated using the inter-arrival time distribution rather than simulating only the total number of weekly arrivals using the Poisson distribution.

3.1.4. Data Collection and Usage

The bookings accepted on a flight are consolidated into build lists, which are used by ground operations to load the aircraft. At this point, actual warehouse deliveries are verified and the cargo is weighed and loaded. A detailed manifest is the result. This information then flows back to the data quality department as to verify this information against the booking information. A procedure dubbed the verification and closeout is executed to account for every kilogramme flown and eliminate any discrepancies with respect to the manifest. This may result in changes to the amounts invoiced as well as to actual or chargeable weights of cargo that is lifting off. This two-stage procedure delivers the information that is used in the internal reporting and is the input to the demand-forecasting model.

Although the weights and invoiced amounts are meticulously checked, any related information is passed through verification virtually untouched and is practically left alone from the moment of initial entry. This includes the volume of the airway bill wholly or partially shipped. In the reporting stage, some errors can still be detected and fixed, but this does not very often include numeric data and almost never volume. The volume of shipments is hence often not accurate.

For volume cargo, a difference between actual weight and chargeable weight typically provides a means to compute the volume of a request. However, will this be useful then the pivot rule must have been applied exactly as it is defined. This is often not the case, especially for shipments below 1000 kg. The volume recorded for those small shipments is often the least accurate. This and the fact that the volume is often not specified or fixed to be one leads to relatively large differences between actual and recorded volume. This is the primary reason why a significant fraction of airway bills is reported to be exactly at pivot as a reasonable default.

---

[6] The customer is Geo, shipping Aircraft Engines from the United Kingdom.

Moreover, where the volume computed based on chargeable weight is different from the volume recorded in the database and the latter is subject to suspicion, e.g. because it is very large, the most modest volume of the two is assumed to be correct and used in the analysis.

Around 7% of all volume cargo has a volume corresponding with a chargeable weight of 100 kg, i.e. its density is 600 divided by its actual weight. This is because a chargeable weight of 100 is often used for billing if the request is smaller. It would have been better if the true weight and volume were recorded. Similar series are at 500 kg and 1000 kg chargeable weight, respectively 1.2% and 1.4%, to be exact.

Because verification and closeout is focused solely on cargo that has actually been flown, there is no good indication of the number or quantity of cargo requests that have been turned down by sales personnel, due to capacity constraints or failed price negotiations. In addition, there is no good information on weight or revenue missed due to cancellations or cargo not showing up before the departure cut-off. The risk of the latter would be mitigated by a suitable overbooking policy. The chapter 3. Constructing Complete Arrival Streams provides a means to correct the arrival streams with regard to this missing information, but the solution proposed herein can only be seen as a workaround. Ideally, the missing information is recorded and included in the estimation procedures.

3.2. Determining Bid-Prices

The section 2.2.3. Constructing Bid-Prices in the chapter 2.2. Admission Control provides overview of the algorithm Lenstra et al (1982) have proposed for finding optimal bid-prices. This algorithm is explained in more detail in this chapter. The algorithm needs the actual weights, volumes and total revenues, including fuel and security surcharges, of requests arriving for a single flight. This means these numbers must be known in advance. Therefore, simulation is used to generate series of space request arrival streams for each of which an optimal bid-price is determined by executing the algorithm provided here. The bid-price to be used in practice is the average of this series of bid-prices.

In the chapter referred to, pointers are also given towards extending the algorithm to compute bid-prices for multiple flights. This extension is however ignored here.

3.2.1. Graphical Representation

The algorithm takes the following as input. Except for the first two, these values are vectors containing one value for each shipment on the flight.

- Maximum weight (w_{max}). The maximum payload weight in kg.
- Maximum volume (v_{max}). The maximum payload volume in dm^3.
- Revenues (r). The revenues of the shipments, including fuel and security surcharges.
- Weights (w). The actual weights of the shipments in kg.
- Volumes (v). The volumes of the shipments in dm^3.

Using this, the algorithm starts by computing the capacity use per Euro of revenue.

- Weight use in kg per Euro ($x=w/r$)
- Volume use in dm^3 per Euro ($y=v/r$)

These values are plotted on the x-axis and y-axis, respectively, as shown in Figure 8. The vectors r, w, v, x and y are ordered by increasing value of x, to be able to iterate them in that order quickly. Each space request in the arrival stream may now be identified by its position in these vectors, the leftmost in Figure 8 being request 1 and the rightmost being request n.

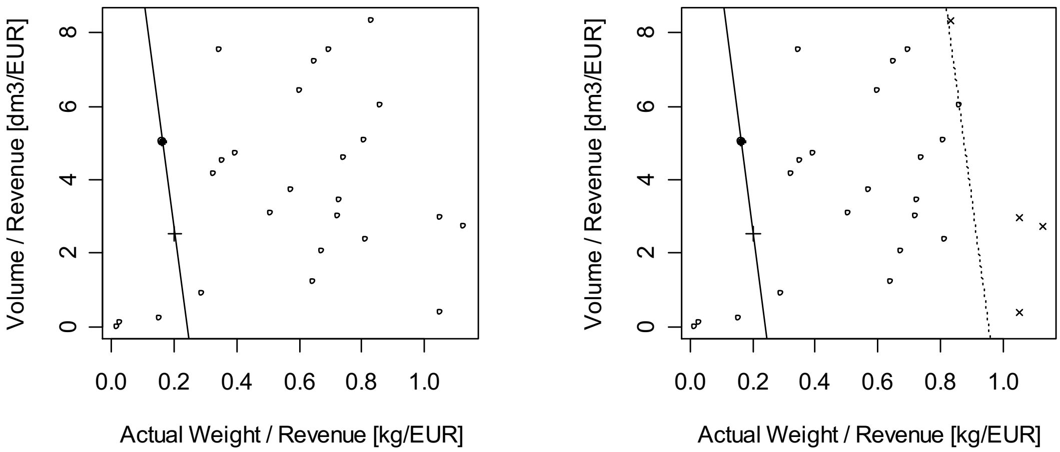


Figure 8 The start of the first step of the algorithm for obtaining bid prices. The plots show actual weight used (horizontal) and volume used (vertical) per Euro of revenue. Each dot represents a single shipment. These shipments where simulated to arrive during week 13, 2005, for movement on PO605/7.

3.2.2. Eliminating Request Pairs

Recall from the description of the algorithm in the chapter 2.2. Admission Control, that the line through two space requests provides exactly that ratio between the bid-price for weight and the bid-price for volume, for which the two requests swap places in the order in which they are accepted. Therefore, all requests are to be paired with each other to determine these ratios. If there are 30 requests, which is a moderate number, 900 pairs are evaluated when using brute force. This would take a long time.

An obvious measure is to remove all pairs of requests for which the reverse combination of requests is already evaluated. In addition, only those lines with negative slopes will ultimately produce valid bid-prices. This is because the line through a pair of requests will only intersect with both the x-axis and y-axis if the slope of this line is negative. See Figure 8 for an example. The algorithm does this as follows.

- Iterate all requests to become the first – or left hand side – of the request pair. Do this from left to right, and call the current request i.
- Skip the first i requests and iterate the remaining to become the second – or right hand side – of the request pair. Do this from left to right, and call the current request j.
- The horizontal and vertical distances between the two requests can now be computed, the horizontal distance being $\Delta x = x_j - x_i$ and the vertical distance being $\Delta y = y_j - y_i$. A positive value for Δx indicates request j is to the right of request i. A positive value for Δy indicates request j is above request i.
- The slope of the line through request i and j is $\beta = \Delta y / \Delta x$.
- If β is not negative, skip this right hand side of the request pair and go to the next j.

At this point, the leftmost three points in Figure 42 are skipped due to the rectangular region to their right and below containing no other points. With these three space requests, no pair can be formed that produces a line with a negative slope.

> **Note** Δx is definitely positive due to the requests being iterated from left to right and the condition $j > i$ is always true, but Δy can have any value and if Δy is positive, so will β be.

- The line through request i and j is now known. The equation of the line is $y = \alpha + \beta \times x$. It intersects with the y-axis at an y-value of $\alpha_j = y_j - \beta \times x_j$. This formula can be generalized to $\alpha = y - \beta \times x$, using the vectors x and y and the scalar β, the slope of the line. This produces a vector of intercepts, one for each space request.

At this point, the graphic on the left in Figure 8 can be constructed completely.

3.2.3. Determining Which Requests are Accepted

The values in the vector α are small for space requests close to the origin and are large for space requests far from the origin. Requests are accepted in the order in which this line encounters them if it moves up from the origin. Hence, if the intercepts in α are sorted, the request being accepted first is at position 1 and the request being accepted last is at position n. Requests are accepted in order of increasing intercept, until there are no more requests or the flight is full.

> **Note** It is more useful to not sort the coordinates of α, but rather order them by creating a vector with n coordinates, the first one being the index into α that contains the smallest value and the last one being the index into α that contains the largest value. This will allow the weights, volumes and revenues of the corresponding requests to be found at the same index into w, v and r, respectively.

- Order the α-vector into a vector called o. If $o_1 = 2$, the second request from the left has the lowest intercept. For the requests in Figure 8 $o_1 = 1$. An alternative notation for o_i used here will be $o[i]$.
- Iterate from 1 to n, call this number k.
- Using the indexes from o, create sums of all weights $w_{o[1]} \ldots w_{o[k]}$ and volumes $v_{o[1]} \ldots v_{o[k]}$, as long as these sums do not exceed the maximum payload weight w_{max} and maximum volume v_{max}, respectively.

> **Note** If all requests can be accepted, the flight is unconstrained, but the relative weights of the bid-prices for weight and volume are still informative.

- The value of k for which the maximum payload weight and volume are not yet exceeded is the number of requests accepted using the relative bid-price weights corresponding with the line through requests i and j.

3.2.4. Computing Flight Statistics and a Bid-Price

The last step in the algorithm computes the bid-prices for weight and volume using the slope of the line through request *i* and *j*.

- There were *k* requests accepted in order of increasing intercept.
- Let the corresponding intercept value be denoted by $\alpha = \alpha_{o[k]}$. This intercept is in dm³/€.
- The corresponding bid-price in €/dm³ for the volume capacity is $1/\alpha$.

At this point, the graphic on the right in Figure 8 can be constructed completely. The interrupted line shows the maximum weight and volume use per Euro for which requests are accepted. The requests above the line are rejected and marked with a cross. The requests below the line are accepted.

- Using $\alpha + \beta \times x = 0$, the intercept with the *x*-axis is at $x = -1 \times \alpha/\beta$. This intercept is in kg/€.
- The corresponding bid-price in €/kg for the weight capacity is $-1 \times \beta/\alpha$.
- The algorithm has completed evaluating the request pair *i* and *j*. Continue with the next *j*, or if request *j* was the rightmost request, i.e. $j = n$, continue with the next *i*.

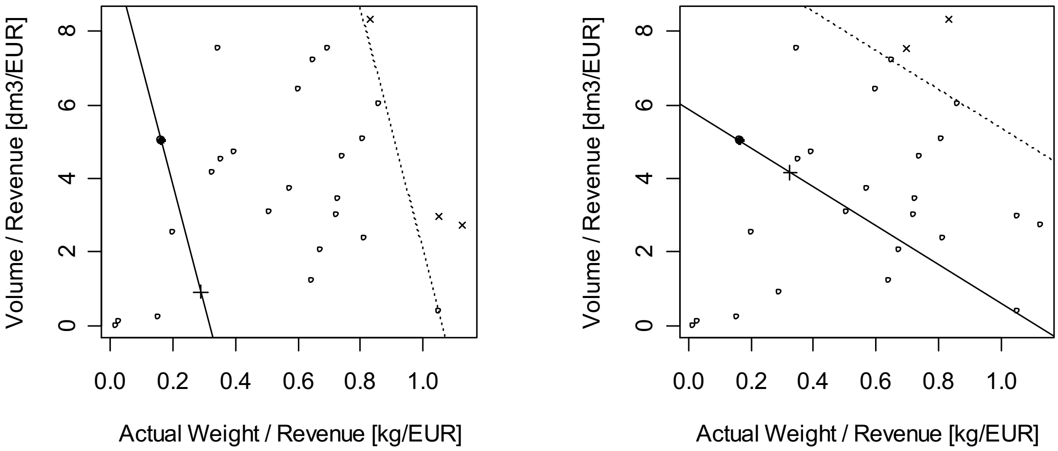


Figure 9 The second and third step of the algorithm for obtaining bid prices. The plots show actual weight used (horizontal) and volume used (vertical) per Euro of revenue. Each dot represents a single shipment. These shipments where simulated to arrive during week 13, 2005, for movement on PO605/7.

3.2.5. Selecting the Optimal Bid-Price

Knowing how many and exactly which requests are accepted makes it possible to produce a flight summary. The total revenue is the revenue that would have been generated if the arrival stream would have been life and the bid-prices computed were in effect.

In addition, a lower bound for the expected optimal revenue can be given by computing revenue as if all requests that were accepted generated revenue exactly at the level of the bid-prices. That is, their weights and volumes are multiplied with the bid-prices for the capacity constraints.

- Record the total revenue of the requests accepted and the lower bound for the expected optimal revenue corresponding with the current bid-prices.
- After all pairs of requests have been evaluated, find the bid-prices producing the highest lower bound for the expected optimal revenue. These bid-prices are the optimal bid-prices for the current arrival stream.

This lower bound is useful in determining the optimal bid-price ratio if there are multiple pairs of bid-prices that have the effect of accepting exactly the same requests, leading to the same total revenue. This lower bound is especially useful for analyzing arrival streams that completely fit on the flight. In this case, the sum of the revenues of the accepted requests is always the same. In this situation, the bid-prices determined using any request pair are most probably over-estimated.

3.3. Constructing Complete Arrival Streams

Using the forecasting model, arrival streams are simulated. Each simulated arrival stream is to contain all space requests arriving during the pre-flight booking periods of a single flight. This includes the following.

- Requests to be accepted that will move normally.
- Requests to be accepted that will not move, for example due to not making it to the freighter aircraft before the departure cut-off.
- Requests rejected due to volume or weight.
- Requests rejected due to revenue. This means the request would not pass admission control.

However, the space requests from which the model is constructed were all accepted and did move, at least as a part-shipment, on the flight they were accepted for. Note that all data is collected from closeout reports.

> **Definition** An arrival stream containing all space requests offered by customers throughout all of the pre-flight booking periods is called complete. It is complete because requests in all of the aforementioned categories are present. Therefore, it is a true arrival stream. Any arrival stream from which space requests are missing for any reason is called a partial arrival stream.

> **Definition** Arrival streams simulated using model parameters that are unchanged after estimating them are called uncorrected arrival streams. If some form of correction has been used to produce the arrival stream, the arrival stream is called a corrected arrival stream.

Partial arrival streams are simulated based on records of space requests that were actually moved, only. These uncorrected arrival streams are hence not complete arrival streams. Nevertheless, they will on average be at par with closeout performance as long as they are not subjected to admission control. Figure 10 shows on a per-flight basis they do.

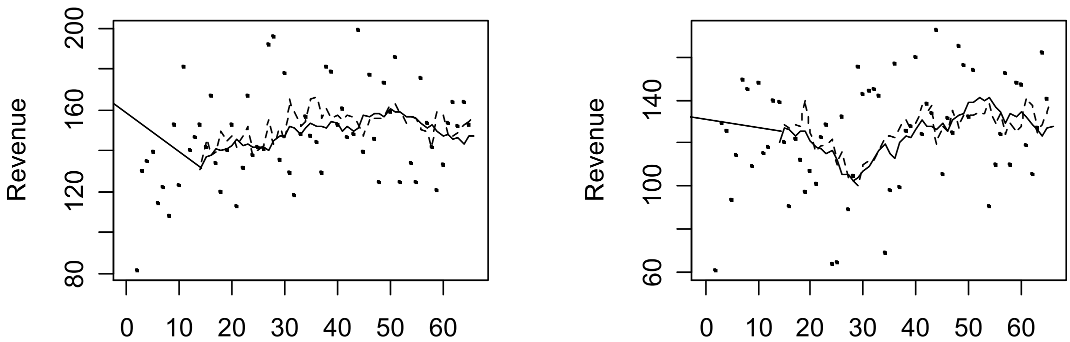


Figure 10 Simulated partial arrival streams (solid) and realized flight performance (interrupted) for PO605/6 (left) and PO605/7 (right), based on 12 weeks of collected data. Revenue is revenue including surcharges, divided by 1000. The value at the horizontal axis is the number of weeks from January 1, 2004.

The figure shows average realized flight performance, alongside average total revenue within the uncorrected arrival streams of 100 simulated flights. For each week, this number of flights was simulated based on a single model, for which parameters were estimated using recorded bookings from the preceding 12 weeks. The average realized flight performances are averages of the 12 flights during the same weeks. The dots represent individual realized flight performances. That is, they represent totals of the revenues gathered from the closeout reports.

> **Note** Although total revenues of uncorrected arrival streams are plotted alongside averages computed over multiple weeks, the revenues of the uncorrected streams are simulated for a single week, only. It is not an objective for them to match exactly the values of the moving averages displayed, although the observations used to compute model parameters are those from the same weeks as were used for computing the moving averages.

3.3.1. Comparing Uncorrected Streams and Realized Flight Performance

Figure 11 shows mean and a 90% confidence range for the total revenue generated by uncorrected arrival streams. The dots represent realized flight performance of individual flights. Note that the realized flight performance is consistently higher than the lower bound of the confidence range. The confidence range is wider than that of realized flight revenue, because of the following reasons.

- The arrival rates used during simulation are under-estimated. The Revenue and Capacity team was able to signal flights poised to be low-revenue and had the choice to allow lower-revenue cargo on board to improve total flight revenue, whereas those extra requests are not present in the simulated arrival streams. Hence, the lower bound of the confidence range is very low.
- The higher bound of the confidence range is higher primarily because capacity constraints are ignored. Very big requests, perhaps even a few of them, within some streams and streams generating many space requests, not all of which will fit, are affecting the high bound of the confidence range. In practice, part of these shipments will have to be part shipped, chartered, offloaded to the next flight, or rejected.

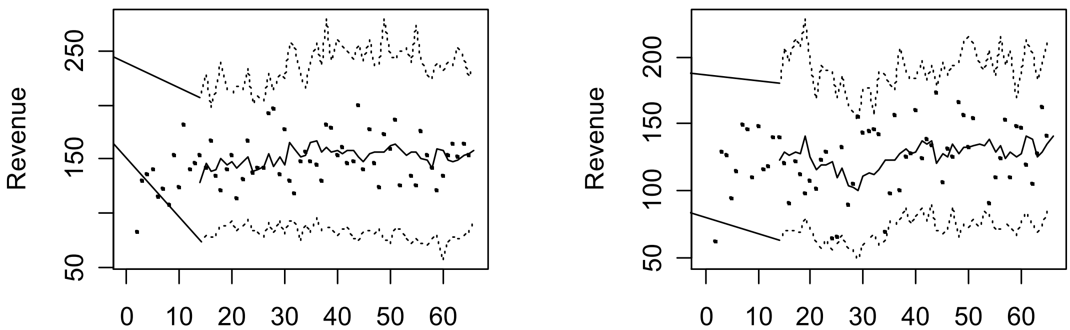


Figure 11 Mean (solid) and a 90% confidence range (dotted) of the total revenue of streams simulated based on 12 weeks of collected data for PO605/6 (left) and PO605/7 (right). Revenue is revenue including surcharges, divided by 1000. The value at the horizontal axis is the number of weeks from January 1, 2004.

Figure 12 shows the chance of observing a total payload weight, volume or revenue. This is done for both real and simulated flights in the same graphic. The simulated total revenues exceeding realized flight performance would turn out to be unachievable, because no admission control has been performed yet, capacity constraints are ignored and these total revenues will correspond for a big part with arrival streams whose total weight and volume will never fit on a single freighter aircraft.

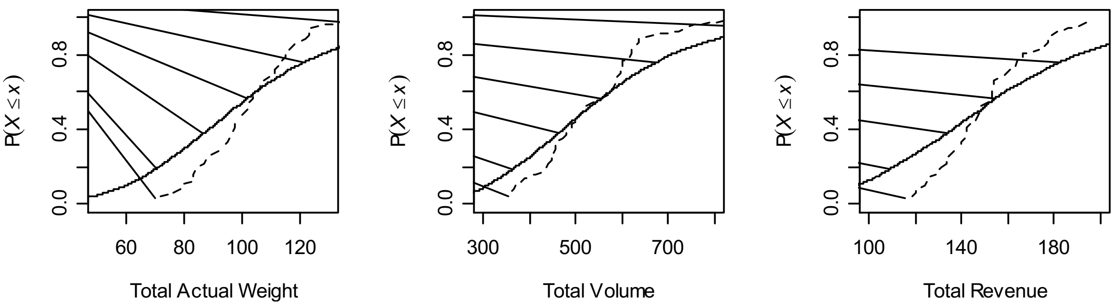


Figure 12 Total weight (left), total volume (middle) and total revenue (right) of the 64 PO605/6 flights for which data was collected (interrupted) and of 1000 uncorrected arrival streams simulated using all observations (solid). The vertical lines correspond with maximum payload capacity.

3.3.2. Accounting for Rejected Space Requests

When performing admission control using uncorrected arrival streams, requests will be rejected, due to both capacity as well as revenue, whenever a flight is constrained. This will happen regularly. Some simulated streams will contain many small requests, while others will contain one or a few very large requests, for example. Whenever the total weights or volumes of requests in a single arrival stream exceed the configured maximum payload weight or volume, the simulated flight is constrained and requests will have to be rejected.

Because on an average volume, weight and revenue of all requests in a series of uncorrected streams will match realized flight performance, rejecting requests for any reason will have the consequence that the performance will drop below the historical level. The forecasted flight performances will be too low. Therefore, a solution is required to change the simulations in such a way that they will produce complete arrival streams with acceptable accuracy, given that data of requests that did not move is unavailable.

In order to correct the simulated partial streams, rejected requests have to be accounted for in some way or another. Those requests that were once accepted, but that do not move for some reason are ignored here, because they are expected to form a small fraction of the total cargo and because they are to be accounted for using a suitable over-booking solution rather than admission control alone. The following arguments can be used in order to find a solution to simulating true, complete arrival streams.

- Foremost, the accuracy of the simulated streams is acceptable, apart from them being partial. The reason why the streams simulated are only partial streams is well understood.
- Simulation using the model generates space requests with weights, volumes and revenues that behave exactly as they should for accepted requests. Weights, volumes and revenues of arbitrary simulated requests are observed as frequent as they should for accepted requests.
- Because a considerable fraction of all flights for which data was collected were unconstrained, low-revenue requests are available to the extend they were accepted in practice and therefore will be simulated, although probably with smaller chance and slightly higher revenue, on an average, than is to be expected during day-to-day operation.

The above means that, assuming requests are rejected independent of their weight, volume and revenue, simply simulating arrival streams using an increased arrival rate will produce a complete arrival stream. Since revenues of individual requests are simulated as if they are already accepted and requests are never rejected due to capacity for flights that are unconstrained, this assumption holds for unconstrained flights.

3.3.3. Arrival Stream Correction using First-Come First-Serve

The partial arrival streams simulated produce space requests whose individual weights, volumes and revenues behave exactly as they should for a partial arrival stream containing only accepted requests. Therefore, performing admission control using a first-come first-serve policy, in which requests are accepted in the order they arrive and requests are not rejected as long as they fit, will have the following effects.

- Volumes and weights of individual requests will be distributed correctly, as long as a simulated flight is unconstrained. In this situation, no single request is rejected under a first-come first-serve policy.
- Revenues of individual accepted requests will be distributed correctly regardless of how constrained the simulated flight is, because requests are never rejected due to revenue.
- When a simulated flight is constrained, smaller requests will still be accepted, but larger requests will be turned down due to a lack of capacity. This creates a deviation between the weights and volumes of requests in uncorrected arrival streams and accepted requests under first-come first-serve, because extreme weights and volumes will be filtered out by this policy, whenever a flight is constrained.
- Since revenues of individual requests are distributed correctly, total average flight revenue is at par with realized flight performance, as long as no requests are rejected. Whenever requests are rejected, average simulated flight revenue will remain roughly at par with historical performance, save for differences in revenue between large and small requests.

The general idea of applying a first-come first-serve policy is that the additional requests simulated due to the higher arrival rate replace those requests that are rejected due to capacity constraints. The more requests available in the arrival stream, the bigger the chance a request is available that is small enough to still fit on the constrained flight. The objective is to find an increased arrival rate at which average revenue under a first-come first-serve policy matches average realized flight performance.

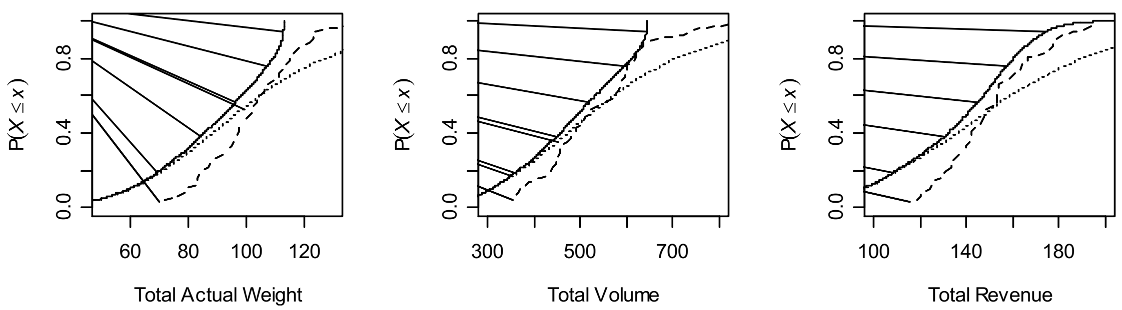


Figure 13 Total weight (left), total volume (middle) and total revenue (right) of the 64 PO605/6 flights for which data was collected (interrupted), of 100 uncorrected arrival streams (dotted) and accepted from these using a first-come first-serve policy (solid).

Figure 13 shows the performances of PO605/6 achieved using a first-come first-serve policy with uncorrected arrival streams. Achievable totals using this policy are strictly lower than the realized flight performances, whereas they should match for unconstrained flights. This indicates PO605/6 flights were hardly ever unconstrained. This is very different with PO605/7.

3.3.4. Correcting the Simulated Arrival Streams

The method described is used to correct the arrival rates estimated from the historical records by simulating a series of arrival streams with arrival rates that are increased by various percentages. Each of these arrival streams is subjected to admission control using the first-come first-serve policy. Afterwards, the percentage is selected that produces flight performances under first-come first-serve that matches historical flight performances closest.

For the PO605 flights operated during the year starting Q2 2004, arrival streams are simulated with arrival rates as they are estimated from the closeout reports as well as with arrival rates that are increased by 10%, 15%, …, 50% and 60%. Each of those is subjected to admission control using first-come first-serve. The flight performances thus obtained are compared to the flight performances computed with a 12-week moving average. The percentage corresponding to the smallest difference between those is selected.

Those percentages are reported on for PO605/6 as well as PO605/7 in the appendix D. Additional Graphics and Tables, under Arrival Stream Correction Using First-Come First-Serve, on page 104. The forecasting can be made more accurate when arrival streams are simulated with arrival rates increased by additional percentages and by incorporating a variance measure into the comparison of performance under first-come first-serve and historical flight performance. The former is a very time consuming process, however, at least when it is done for 106 departures at once.

The numerical data presented in the appendix shows PO605/6 is considerably more constrained than PO605/7. The percentages by which the arrival rates are increased to match historical performance are much higher for PO605/6 than for PO605/7. Whereas additional simulations using percentages in the range 50-70% may be expected to improve forecasting accuracy for PO605/6, percentages in the range 0-15% would increase accuracy of the forecast for PO605/7. The percentages matching historical flight performances closest show large differences depending on the time of year. In addition, the figures are reason to state that low overall demand has its effects on PO605/7 first before affecting PO605/6.

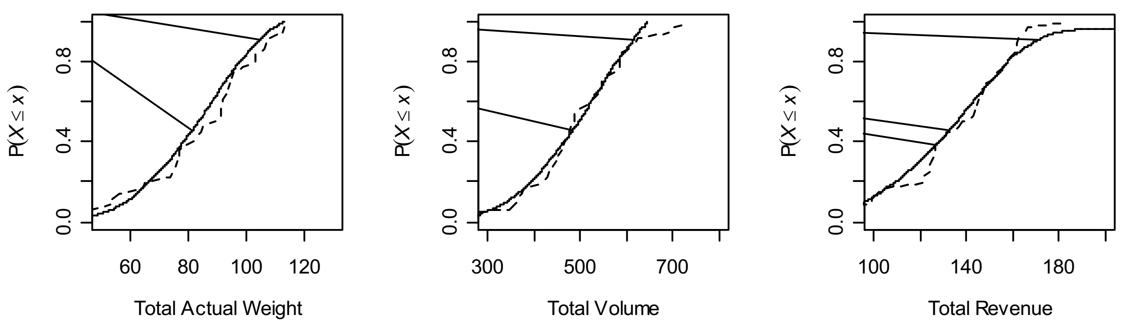


Figure 14 Total weight (left), total volume (middle) and total revenue (right) of the 64 PO605/7 flights for which data was collected (interrupted) and of the unconstrained arrival streams out of 100 simulated with 20% increased arrival rates, accepted using a first-come first-serve policy (solid).

Figure 14 shows the performances of unconstrained flights under first-come first-serve using arrival streams with 20% increased arrival rates. A 20% increased arrival rate means Polar on average accepts about 5 of every 6 requests available to the company.

Figure 15 shows the effects on average total flight revenue of increasing the arrival rates by two different percentages for PO605/6. The lower percentage of the two is more applicable about the first 6 months of the year, whereas the higher percentage fits better throughout the second half of the year.

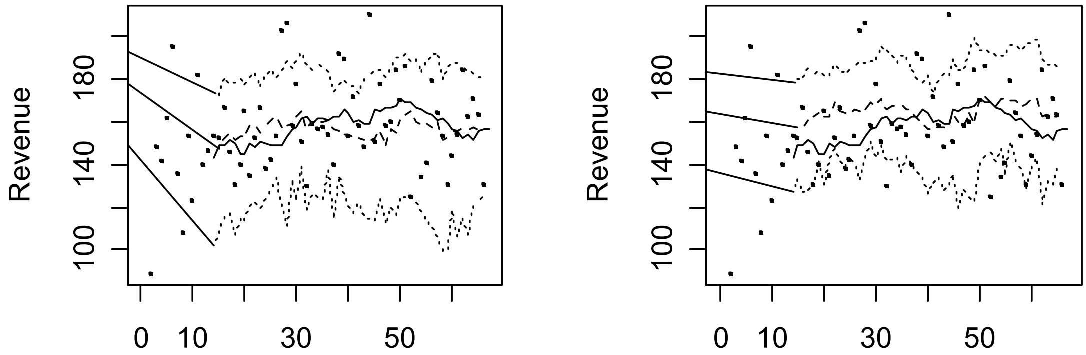


Figure 15 Mean (interrupted) and a 90% confidence range (dotted) of the performances under first-come first-serve of corrected arrival streams with 25% (left) and 50% (right) increased arrival rates against average realized flight performances (solid) of PO605/6. The value at the horizontal axis is the number of weeks from January 1, 2004.

The average arrival rate increase computed for PO605/6 is 32.4%, meaning 75.6% of all requests offered to its sales personnel and agents throughout Europe are accepted. For PO605/7, the average increase is 6.89%, meaning 93.6% of all requests are accepted. These values are averages throughout the whole data collection period.

3.4. Flight Performance

3.4.1. Configuring the Forecasting Model

In order to start forecasting flight revenue for a particular week or a part of a week, the following information has to be collected.

- Total payload capacity of the aircraft flying the service
- Total payload capacity already sold and the revenue already achieved
- The booking periods remaining to be forecasted for
- The length of the period to use data from for estimating model parameters
- A choice for a correction method for simulating complete arrival streams
- The current rate of the fuel and security surcharges in effect
- The number of arrival streams to simulate

When forecasting for the week to come, for example, pre-bookings will already be known. In this situation, booking period T-7 is to be excluded. The total weight, volume and revenue of the requests already accepted can be specified to produce a forecast with a complete closeout report. Alternatively, total payload capacity can be reduced by the amount of weight and volume already in use. This will produce a forecast of only those weights, volumes and revenues still to be expected.

If a single working day is not to be simulated, for example due to it being a national holiday, the arrival rate of this day can be set to zero as to make sure no arrivals are recorded during this day. Whatever effects this has in practice on the arrival rates of the surrounding days is ignored by default, but these can be incorporated by changing the arrival rates after parameters are estimated. Parameters are always estimated using all data selected using the method described below.

The length of the period to use data from for parameter estimation can be specified as a number of weeks. The weeks used will be those weeks preceding a certain date, for which observations are present. Any weeks for which no observations are present will be skipped and not counted towards selecting the requested number of weeks.

The correction method used for simulating complete arrival streams may be any one or a combination of the following. These are listed in order of increasing quality.

- A correction factor for the arrival rates estimated from the observations when arrival rates are corrected using first-come first-serve
- Estimates of the arrival rates of requests that are expected to be rejected or cancelled
- Recorded weights, volumes and revenues of requests rejected or cancelled may be included as individual requests into the dataset used for estimating model parameters

It is important the observations are all forming complete weekly closeout reports or complete arrival streams, depending on the correction method used for simulating complete arrival streams. Any observations missing will influence the estimated arrival rates as well as the simulated weights, volumes and revenues. The first two correction methods are applicable under the assertion the observations available are from closeout reports only. The last method is useful under the assertion the observations form complete arrival streams.

The number of weeks to use is important for several reasons. As a rule of thumb, selecting more observations for parameter estimation will increase the quality of the estimates. However, this also means the observations are collected over a longer period and hence trend and seasonality play a more significant role. Any trends or seasonal influences, if present within the period selected for parameter estimation, are ignored, which means their effects are flattened out. This can be mitigated, at least with regard to arrival rates and actual weight, by assigning weights to observations based on from which week they are and computing weighted averages, or by producing single-week estimates first and extrapolating the behaviour over time of this series of estimates.

For density, volume and revenue, the former approach is however not suitable, because the estimates are not based on averages. The latter may become too complex and slow for practical use, because a series of maximum likelihood computations will be necessary. In addition, because estimation of parameters for dependent properties is similar to segmenting the data often, any of the bins created by segmenting the data will contain only a few samples. Hence, individual single-week estimates may not be of sufficient quality.

> **Note** A single estimation of quality for density and revenue requires a relatively large number of observations to be input to the estimation algorithm. Estimates regarding actual weight, arrival rates and density classification are less sensitive.

The number of weeks of data used determines the number of observations available. Obtaining maximum likelihood estimates based on one or a few weeks will be possible without errors as long as there are sufficiently many space requests recorded during those weeks. The quality of the estimate will however decline with the decreasing number of observations. If the observations form complete arrival streams, less weeks are required compared to when these are from closeout reports exclusively. A reasonable absolute minimum is 3 weeks of data if the observations are from closeout reports only and 2 weeks if they form complete arrival streams. This means estimates are based on somewhere between 50 and 100 observations. PO605/6 will require fewer observations than PO605/7 for estimates of similar quality, simply because the arrival rates are higher for the former flight.

> **Note** The algorithm used for computing maximum likelihood estimates sometimes violates the bounds set for the parameters it estimates, causing errors. Usually, however, a next execution of the procedure will complete without errors. The chance the algorithm fails, for example due to the shape parameter of the Gamma distribution being set to a non-positive number, seems to increase when the number of observations it has as input is smaller.

The choice for a certain number of weeks will have to be established based on an assessment of the quality measures presented throughout the results section, like graphics showing the goodness of fit.

- Simulated average flight performance will match realized historic performance better if data from fewer weeks is used, because trends and seasonal influences then play a smaller role.
- Lower and upper bounds of the confidence range of simulated flight performance may be less extreme if more data is used. A confidence range that is much larger than reported in this document is hinting towards individual distribution estimates not being of adequate quality. If enough data is used, the confidence ranges will stabilize at a range corresponding with the variance and standard deviation of the underlying probability distribution.
- The quality of the individual distribution estimates for booking properties, like the revenue of a request, will match the empirical distribution of the observations better when there are more observations. This is both due to the empirical distribution being closer to the underlying 'real' distribution, as well as due to the estimates based on this data being of higher quality.

The number of flights simulated to produce a single forecast will also influence forecasting quality. The more arrival streams are simulated based on which bid-prices are computed as well as admission control is performed on, the better the simulated performance will match with the parameter estimates and the more stable forecasted flight performance will be.

> **Note** Although during this investigation often only 100 flights are simulated to produce forecasts, it is recommendable to simulate many more during day-to-day forecasting. Producing say 1000 simulated flights instead of 100 will only take a minute more if a single forecast is to be produced.

Recall simulation is done because the compound experiment of a single flights departure cannot be replicated by theoretical means with adequate precision. Each simulated arrival stream produces a single realization of this experiment. Simulating more realizations will improve the quality of the empirical distributions of the measures of interest, like the average optimal bid-price and forecasted total flight revenue.

> **Note** Quality of estimates can be regarded proportional with the square root of the number of observations used. For example, simulating four times more flights will produce total flight revenues whose empirical distribution is on average half the distance away from the underlying 'real' distribution being simulated by simulating the compound experiment in small steps. This is a result of the *Central Limit Theorem*. If subsequent simulations with exactly the same parameters and configuration produce very different results, this is an indication too few realizations have been generated.

Simulating a series of arrival streams can be done in an amount of time that is increasing almost linear with the total number to simulate, although the significant memory use of the current implementation may disturb this relationship, especially on less capable computers. Quality will increase linear with the square root of the total number to simulate. A simulation sequence 10 times larger is expected to produce distribution estimates with approximately 3.2 times smaller residual standard error. This is the spread of the differences between the estimated distribution and the underlying 'real' distribution.

> **Note** The current estimation and simulation algorithms, which are implemented in script, can be replaced with a compiled version. This will dramatically increase speed and may reduce memory requirements considerably. Third party software components are widely available to aid with this. The R Project for Statistical Computing for example can be used from almost any programming language on Windows operating systems to do the difficult work.

3.4.2. The Forecasting Process

In its current implementation, the forecasting process works by executing the following steps.

- Filter the observations based on flight number and day of week
- Select observations of the configured number of preceding weeks
- Estimate model parameters using the selected observations
- Increase arrival rate estimates, using arrival rate correction with first-come first-serve

  Note The percentage by which the arrival rates are increased must be set in advance or be determined by comparing historical and forecasted performance under first-come first-serve.
- Reset any arrival rates to zero of booking periods not to be simulated
- Simulate the number of arrival streams configured, each one belonging to a single simulated flight identifiable by a sequence number
- Determine optimal bid-prices for each of the simulated flights
- Compute a single pair of bid-prices for weight and volume to be used in practice
- Simulate another series of arrival streams, the verification series
- Performing admission control on the verification series, using the single pair of bid-prices computed
- Report on the optimal bid-price pair selected, total flight revenues, total payload weights and volumes and quantities of accepted requests as well as requests rejected, due to either capacity or revenue

The fraction by which the arrival rates are increased, or marked up, are manually determined in advance and taken to be the fraction by which the arrival rates of the preceding week were increased.

> **Note** The forecasting process can be configured to perform admission control using a Capacity Access Price and using a first-come first-serve policy in addition to performing admission control using a pair of bid-prices easily and simultaneously.

3.4.3. Determining the Optimal Bid-Prices

After the configured number of arrival streams is simulated, all of the simulated requests are input to the program computing the bid-prices. Table 1 shows part of the output generated by this program.

```
Computing Bid-Prices for PO605/6 of 2005WK13

Capacity   113,000 kg,   645,000 dm³
Achieved         0 kg,         0 dm³, EUR       0
Available  113,000 kg,   645,000 dm³

Sequence,Iterations, [Total Offered  ] [Total Accepted ] [Bid-Prices ] Flight Status
Flight 1,  467 iter, [52,125, 799,219] [41, 97, 637,187] [1.018,0.053] Full
Flight 2,  155 iter, [35, 45, 276, 83] [35, 45, 276, 83] [0.897,0.023] Unconstrained
Flight 3,  659 iter, [61,124, 794,227] [55,110, 602,198] [0.091,0.164] Full
Flight 4,  299 iter, [45,158, 873,229] [40,113, 611,182] [0.578,0.097] Full
Flight 5,  319 iter, [39, 85, 478,146] [39, 85, 478,146] [0.784,0.003] Unconstrained
Flight 6,  149 iter, [34,113, 570,172] [33,104, 505,162] [0.307,0.129] Heavy
Flight 7,  198 iter, [36, 74, 461,124] [36, 74, 461,124] [0.526,0.060] Unconstrained
Flight 8,  259 iter, [41,125, 661,191] [36,112, 573,177] [0.544,0.109] Full
Flight 9,  179 iter, [41,112, 646,146] [37,110, 632,144] [0.383,0.114] Volumed Out
```

Table 1 A series of reports, one for each simulated flight with a unique sequence number, amongst others stating the optimal bid-price computed for that simulated PO605/6 flight.

The program reports the number of iterations required to find the optimal bid-prices, the total quantities of the requests in the arrival stream and the total of those that were accepted with the optimal bid-price pair. The bid-prices reported are the optimal bid-prices for the weight and volume capacity constraints, respectively. The four numbers in the other two number groups are the number of requests, total weight, volume and revenue, in that order.

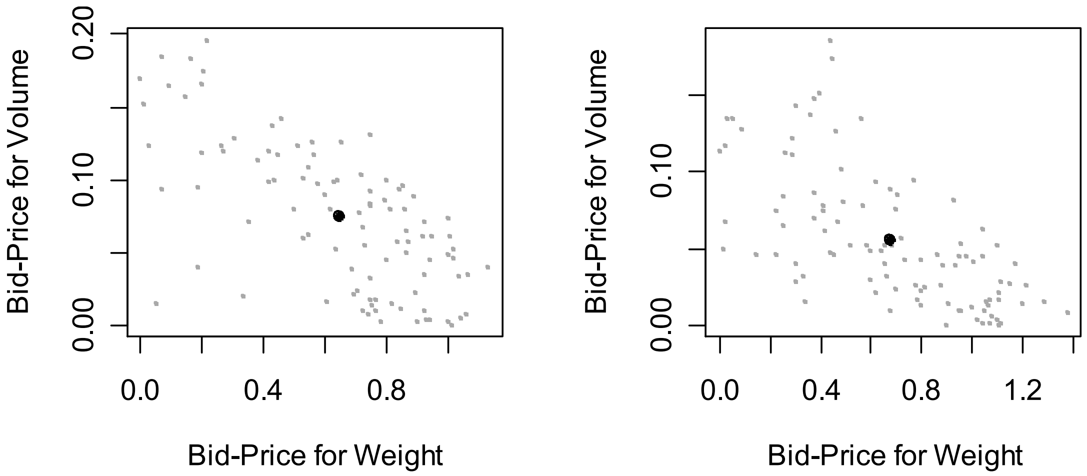


Figure 16 Optimal bid-prices computed for 100 corrected arrival streams for the PO605/6 (left) and PO605/7 (right) of week 13, 2005. The black dot is the average optimal bid-price.

Figure 16 shows the 100 bid-prices computed for the PO605/6 and PO605/7 of week 13, 2005, to give an idea graphically of the locations of the optimal bid-prices computed for a series of arrival streams. The bid-prices computed to perform admission control with are standing out in this graphic. These are obtained by computing the mean of all of the bid-prices that were optimal for individual simulated arrival streams.

3.4.4. Determining the Optimal Capacity Access Price

The optimal capacity access price for a single arrival stream is computed using the following rules.

- The pivot ratio is fixed to be 6.
- Chargeable weight is computed for each request using the pivot rule. That is, chargeable weight is taken to be the maximum of actual weight and volume divided by 6.
- The rate of a request is its revenue excluding surcharges divided by its chargeable weight.
- Requests are accepted in order of decreasing rate.
- The optimal capacity access price for a single stream is the rate of the last-accepted request.

The capacity access price used for performing admission control is the average of the capacity access prices that were optimal for each of the individual arrival streams simulated.

3.4.5. Performing Admission Control

Table 2 shows a closeout report as it is generated by the program performing admission control. The program can perform admission control with a static bid-price policy as well as a policy based on a Capacity Access Price. The first-come first-serve policy is equal to a static bid-price policy with bid-prices of zero for both the weight and volume capacity constraint.

```
Performing Admission Control for PO605/6 of 2005WK13
(...)
Flight  94 [ n,Act, Vol,Eur] Heavy (37 requests)
Capacity   [   113, 645    ] maximum payload.
Achieved   [     0,   0,  0]
Available  [   113, 645,121] revenue at Bid-Price levels.
Offered    [37,136, 611,192]
Shipment   [14,  4,  55,  6] ( 5,629- 6,879=-1,251>=0) rejected due to revenue.
Shipment   [20, 16,  82, 13] (12,509-16,315=-3,806>=0) rejected due to revenue.
Shipment   [35,  4,  18,  5] ( 5,127- 3,948= 1,179>=0) rejected due to weight.
Rejected   [ 2, 20, 136, 18] due to revenue.
Rejected   [ 1,  4,  18,  5] due to capacity.
Accepted   [34,111, 456,168] at Bid-Price levels [kg,dm³]=[0.644817, 0.074504].
Perished   [     2, 189, 15] revenue at Bid-Price levels.
Closeout   [34,111, 456,168]
```

Table 2 A closeout report generated by the admission control program for a simulated PO605/6 flight.

Admission control is performed on a single arrival stream that is constructed by simulating arrivals during all of the booking periods defined or a subset of these. The flight number reported is the sequence number identifying one such arrival stream. The total weight of the 37 requests in this example exceeds the configured maximum payload weight less weight already accepted. Hence, the flight is reported to be heavy. In addition, the closeout report reports on the following.

- **Capacity**. The configured maximum payload weight and volume.
- **Achieved**. Total payload weight, volume and revenue of requests already accepted during previous booking periods when not all booking periods have been simulated.
- **Available**. Maximum payload weight and volume less weight and volume already accepted. The reported revenue is the threshold revenue of the policy in effect of a request that exactly fits this remaining capacity.
- **Offered**. The number of requests simulated to be arriving during the booking periods that have been simulated and their total weight, volume and revenue.
- **Shipment**. A summary of a request is displayed for each request that is rejected. The position of the request rejected within the arrival stream is reported first, next to its weight, volume and revenue. Moreover, its total revenue is compared against the minimum required revenue according to the admission control policy in effect. Negative values indicate the request does not meet this revenue level. If a request does not meet the required revenue level and does not fit, it is reported to have been rejected due to capacity.
- **Rejected due to revenue**. The total number of requests rejected due to their individual revenues failing to meet the revenue requirement of the policy in effect and their total weight, volume and revenue. A request is never reported to be rejected due to revenue if it does not fit.
- **Rejected due to capacity**. The total number of requests rejected due capacity and their total weight, volume and revenue.
- **Accepted**. The total number of requests simulated to be arriving that are accepted during the booking periods that have been simulated and their total weight, volume and revenue. A summary of the policy in effect is also reported here. For a static bid-price policy, the bid-prices for the weight and volume constraints are reported. For a policy based on a Capacity Access Price, the Capacity Access Price being enforced is reported along with the rate of the fuel and security surcharges as well as the pivot ratio.
- **Perished**. The payload capacity that is lost due to the flight departing without its maximum payload capacity having been sold completely. This is the maximum payload weight and volume less totals of weight and volume accepted after admission control has been performed.

The reported revenue is the threshold revenue of the policy in effect of a request that exactly fits this remaining capacity.

- **Closeout**. The total number of requests simulated to be arriving that are accepted during the booking periods that have been simulated and the total weight, volume and revenue of these requests plus the total payload weight, volume and revenue of requests already accepted during previous booking periods when not all booking periods have been simulated.

3.4.6. Comparing Flight Performance under different Admission Control Policies

Although a policy based on a Capacity Access Prices must be suboptimal because it is not a strictly greedy policy, the question is whether this theoretical results holds in practice. Table 3 shows it will.

The table shows total revenue, weight and volume as well as the number of requests within the arrival streams on which admission control is performed. For each week from week 14, 2004 to week 13, 2005, 100 arrival streams are simulated, based on data collected using the preceding 12 weeks. The averages shown are thus averages over 5300 simulated flights.

In hindsight, 28.1% of the arrival streams could not completely fill the maximum payload capacity available, whereas 71.9% of the arrival streams did. For 10.1% of the simulated arrival streams, the total weight of the requests exceeded maximum payload weight. For 4.4% of them, total volume exceeded maximum payload volume.

Performance Measure	**All Arrivals**	**Accepted**			
All Simulated Flights	100%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	211,356	157,140	156,413	158,099	1,686
Payload Weight	138,982	102,061	94,581	97,165	2,584
Payload Volume	758,051	550,506	516,539	511,638	-4,901
Requests	40.22	34.11	33.48	33.78	0.30
Unconstrained Flights	28.1%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	139,584	139,623	129,655	131,515	1,861
Payload Weight	88,738	88,751	78,906	80,791	1,886
Payload Volume	474,443	474,247	427,201	427,789	588
Requests	34.33	34.32	32.16	32.59	0.44
Constrained Flights	71.9%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	239,503	163,985	166,907	168,524	1,617
Payload Weight	158,686	107,261	100,728	103,586	2,858
Payload Volume	869,275	580,301	551,575	544,521	-7,054
Requests	42.53	34.02	34.00	34.24	0.24

Table 3 Summary of the effects of different admission control policies on the flight performances of 53 simulated PO605/6 flights, based on arrival streams simulated based on estimates over 12 weeks of collected data.

The column CAP shows the totals accepted under a policy based on a Capacity Access Price. The column BP shows the totals accepted under a static bid-price policy. The last column shows the difference between the last two.

Historical flight performance of PO605/6 was € 156,032 over 66 flights from January 10, 2004, with on average 103,270 kg payload weight and 567,507 dm^3 payload volume from 30.79 requests. On an average per-flight basis, the forecasted performance using a static bid-price is € 2,067 higher than historical flight performance and each flight carries 6,105 kg and 55,869 dm^3 less payload. With 52 flights per year, the policy improves revenue by € 107,484 and at the same time reduces payload airlifted by 317,5 tonnes and 2.905 m^3.

Throughout the year starting Q2 2004, a static bid-price policy on average outperforms a policy based on the weekly optimal Capacity Access Price by € 87,672, over 52 departures. The policy does this at the cost of carrying an additional 134,368 kg actual weight over 52 flights, but also carries 254,852 dm^3 less volume.

Similarly, of on average 14.6 flights departing during this year for which no requests are rejected due to capacity, a static bid-price policy accepts on average € 27,193 more revenue at the cost of carrying 27.6 tonnes more weight and 8.6 m^3 less volume.

The remaining 37.4 flights that are heavy, volumed out, or both, generate € 60,456 more revenue using a static bid-price policy and carry 106.8 tonnes more weight and 263.7 m^3 less volume.

Since a static bid-price policy accepts on average only 15.6 requests more over 52 weeks, shipping, handling, security cost and administrative overhead are very small. The extra revenue accepted is accepted against little more than the absolute lowest differential cost, most notably consisting of fuel and security cost. Since it is reasonable to assume the break-even point with regard to the constant cost has been surpassed, the biggest part of the extra revenue can be expected to contribute directly and fully to operating profit.

3.4.7. A Shortcut to Forecasting Flight Performance

In this document, a single flights departure is seen as a probabilistic experiment. As such, measures like take-off payload weight and volume as well as total flight revenue, amongst others, have been defined to be stochastic variables with probability distributions and expectations.

Because the probability distributions of these stochastic variables are unknown and cannot be easily determined, a single flights departure is broken down into space requests arriving and being accepted or rejected. These events have been modelled with probability distributions and are subsequently simulated, ultimately to produce empirical probability distributions of the measures that are interesting at the level of a single flights departure.

Whenever a forecast is produced, the results will still incorporate the uncertainty inherent to not knowing future events. The total flight revenue that can be reported at the end of the week is just one realization of a probabilistic experiment. In other words, it is a quantile of the empirical distribution produced by simulating a series of flights using the forecasting model, evaluated at a more or less random number between 0 and 1.

The efforts of the sales force, amongst others, must be seen as if it has the effect of increasing the expected arrival rates and revenue potential of individual requests. Without it, the expected flight revenue will rapidly decline towards zero. A successful carrier without a sales force will soon enough simply cease to be a carrier. With it, the uncontrollable factors that will affect flight performance can be and are suitably captured in the forecasting model. Major efforts out of the ordinary that are conducted by the sales force, for example promotional activities, must be incorporated as accurately as possible into the forecasting model in order to avoid surprises.

Another reason why single flight departures have been broken down was to be able to provide flight management controls. Most notably, this means an effective admission control policy, but also a flexible, configurable forecasting model. However, when only a single statistic is of interest, like total flight revenue, a useful shortcut exists for forecasting it quickly. It is based on the *Central Limit Theorem* and normal distribution theory.

It can be proven very easily using this theorem that total payload weight, total payload volume and total flight revenue must have a distribution matching the normal distribution sufficiently close if the weekly arrival rates are at normal levels, since these are the sum of a number of identically distributed independent stochastic variables. An estimate can be produced in any of several ways.

- By conditioning on the number of requests accepted, these performance measures can be seen as the sum of a series of stochastic variables that are distributed with a normal distribution. Such a sum has a normal distribution itself, its mean is the sum of means and its standard deviation is the square root of the sum of variances.
- The performance measure is a Poisson distributed number of requests multiplied with a stochastic variable whose mean and variance can be easily obtained. The mean of the product is the product of the means and the variance of the product is the product of variances.

The above methods are not very accurate in all situations, although they are useful for quickly producing a flight performance forecast. They perform better generally for PO605/7 than for

PO605/6. This may be attributed to this method being not sophisticated enough to respect payload capacity constraints.

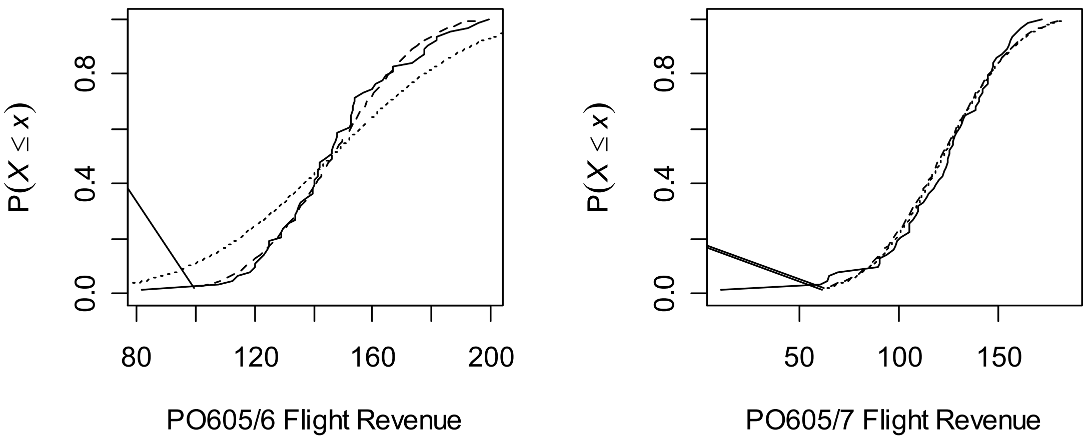


Figure 17 Estimates of total flight revenue at flight departure for PO605/6 (left) and PO605/7 (right), with empirical distributions (solid) and normal distribution fits, based on historic flight performances (interrupted) and based on the estimated weekly arrival rate distribution multiplied with the revenue per request (dotted).

Figure 17 shows realized total flight revenues of all flights for which data was collected, along with several normal distribution fits. One of those was solely based on 63 historic flight performances. This normal distribution has a mean and standard deviation matching those of the 63 realized flight performances. The second normal distribution fit was created using the latter method described above.

> **Example** Forecasting total flight revenue for the PO605/6 flight of April 2, 2005, data is used from flight closeout reports starting week 1, 2005 and ending week 12, 2005. The estimated arrival distribution is a Poisson distribution with arrival rate 32.8. Hence, its mean and variance both are 32.8. The 394 requests that moved during these weeks have average revenue € 4,494 with a standard deviation of € 6,259.
>
> A rough estimate of total flight revenue is, that it has a normal distribution with mean 32.8 × € 4,494 = € 147,561 and standard deviation the square root of variance 32.8 × $6{,}259^2$, which is € 35,863. This distribution suggests total flight revenue is between € 113,000 and € 182,000 twice as often as it is not.
>
> Simulation suggests expected flight performance is € 147,621. Only nine out of 100 simulated flight performances are below € 113,000 and only two of those are higher than € 182,000.

3.4.8. Flight Performance Development and the Booking Curve

The booking curve provides insight in how flight performance develops from the start of the first booking period until flight departure. It is a graphical tool for monitoring flight performance.

Figure 18 shows booking curves of weight, volume as well as revenue for the PO605/6 of April 9, 2005. These booking curves are based on 250 flights, all simulated with model parameters estimated not taking into account realized flight performance during week 14, 2005. The realized flight performance turns out to be on the low side during the start of the week, for a large part due to the absence of pre-bookings. Revenue is even outside the 90% confidence range displayed at the start of Tuesday April 5, which is booking period T-4. However, during the Friday before scheduled flight date, the flight revenue is improved in excess even of the expected flight revenue forecasted at the start of the week.

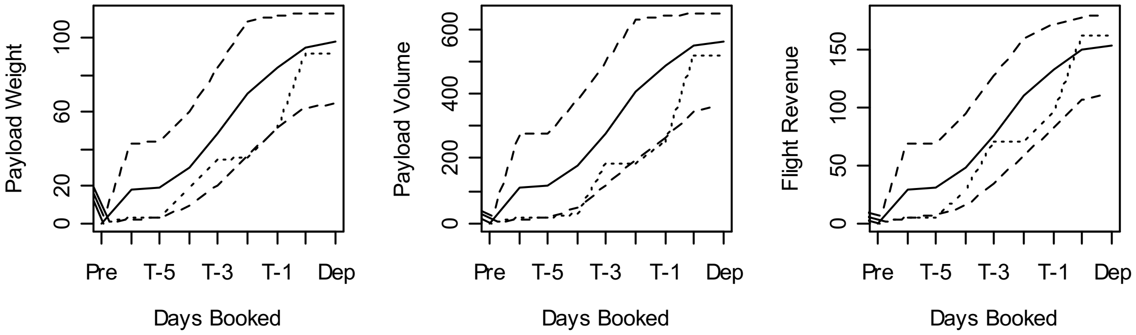


Figure 18 Booking curves of payload weight (left), volume (middle) and flight revenue (right) for the PO605/6 of April 9, 2005, with expected quantities at the start of the booking period (solid) and a 90% confidence range (interrupted), as well as realized quantities at the start of each booking period (dotted) up to flight departure ('Dep').

Since arrival rates during subsequent booking periods and properties of requests arriving after each other are unrelated and request sizes and revenues have large variance, the booking curve is of little use towards rating the realized flight performance. In addition, a booking curve simulated even before the start of a week will be different from booking curves simulated for a remaining part of the week, even if the same admission control policy is used with similar minimum revenues, since the remaining capacity will be known instead of it being randomly simulated.

Booking curves simulated later will be based on simulated arrival streams of which more requests can be accepted when flight performance has been low and of which much less requests can be accepted when flight performance has been excellent. This will affect both the booking curve itself, being constructed from expected flight performances at the end of each booking period, and an arbitrary confidence range, as the one shown in Figure 18 for example. With less payload capacity available for sale, flight performance at departure will show a much smaller spread.

Figure 19 gives another illustration why the booking curve is of limited use, again for the PO605/6 flight of April 9, 2005. Each estimate of flight performance at departure that is displayed in this figure is built acknowledging realized performance up to the end of the preceding booking period. The available payload capacity at the start of each booking period is taken to be the remaining capacity at the end of the preceding booking period and achieved flight revenue is accordingly added to the performance simulated during the remainder of the week.

The figure shows simulated end-of-week flight performance is affected the first booking periods much more moderately than the booking curve suggests. Only when scheduled flight date approaches, a weak flight performance development proves to have significant impact. Figure 19 illustrates this by showing a declining average expected flight performance during booking periods T-3 and T-2, where flight performance increases slower than average. The figure shows beautifully the effects of a single days performance to turn out better or worse than would be expected beforehand. For the booking curve to be informative, it is required to produce a huge number of them in order to compare a single flights performance with it. Otherwise, it is only capable of producing point estimates without quality measure, which will turn out to be of questionable accuracy.

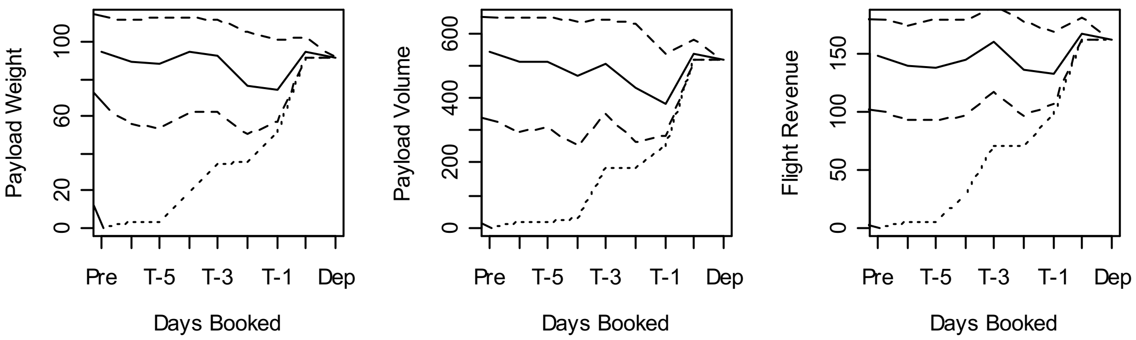


Figure 19 Total payload weight (left), volume (middle) and flight revenue (right) for the PO605/6 of April 9, 2005, with the development of the expected quantities at flight departure (solid) and a 90% confidence range (interrupted), as well as realized quantities at the start of each booking period (dotted) up to flight departure ('Dep').

Note the upper bound of the 90% confidence range for weight and volume displayed in the figure reflect the fact more than 5% of all simulated flights are constrained with regard to weight or volume. The upper bound of the confidence range only differs from maximum payload capacity because not enough very small requests were available within the simulated arrival streams that fit the small amount of remaining capacity.

3.4.9. Admission Control under Definite Overcapacity Conditions

When admission control is performed on a live arrival stream, shipments will be rejected before it is known whether the flight will be constrained. An objective of admission control is to make a near on-line decision, since the customer has to be informed quickly of the decision made. The accept/reject decision is made aiming to maximize expected total revenue at take-off. However, the demand development after rejecting a shipment may turn out much worse than expected, even such that the payload capacity remaining cannot be completely sold. This means revenue is sometimes lost that, in hindsight, proves to have better been accepted.

If a shipment that fits is rejected by admission control, this is because its opportunity cost exceeds the direct revenue earned by accepting the shipment. At the time the request was rejected, this decision was admissible towards maximizing total flight revenue. Whenever the forecasted demand development is accurate, the level at which the bid-prices or the Capacity Access Price is set reflects the amount of demand forecasted. Lower forecasted demand will produce lower levels, hence lower opportunity cost. This will result in more shipments being accepted.

> **Note** The revenue lost by rejecting requests on flights that turn out to be unconstrained is generally offset by the improved revenue of flights that are constrained at take-off.

There is always a positive chance that the next shipment arriving is so large that it almost completely fills the available capacity. Recall that a requests weight, for example, is not bound to a maximum. The probability that even the next single shipment arriving is large enough to fill the flight decreases well below theoretical levels as the flight fills up. This means that whether the flight will turn out to be constrained may remain a practical possibility until very close to departure, even if the demand during the first part of a week has been disappointing.

Figure 19 illustrates this very well. The upper bound of the confidence range for payload volume at take-off, shown in this figure, indicates the probability that the flight will be volumed-out is still over 5% up to T-1. Note the 90% confidence range shown is a two-sided confidence range.

The asymptotic optimality of the static bid-price policy holds, at least as long as the flight is constrained. Practically, however, the level of the bid-prices or the Capacity Access Price computed to optimize total flight revenue may be such that the opportunity cost of some or all shipments drops below the cost of carrying that cargo. In this situation, the opportunity cost of a shipment must be taken to be the maximum of the opportunity cost under the admission control policy in effect and the cost of carrying the shipment.

> **Note** The opportunity cost of a shipment must be taken to be the maximum of the opportunity cost under the admission control policy in effect and the cost of carrying the shipment.

Regularly renewing the computed optimal bid-prices or Capacity Access Price will ensure enforcing the admission policy strictly will indeed optimize total flight revenue at take-off. When the chance that the flight will turn out to be unconstrained is very large due to low demand, however, using a first-come first-serve policy might become attractive. This is due to the following reasons.

- The bid-prices or Capacity Access Price computed for an arrival stream that does not fill the available capacity generally is the highest one for which the request in the arrival stream is accepted that generates the least revenue per unit of capacity. The bid-price or Capacity Access Price computed might in hindsight therefore prove to be too high.
- The bid-prices or Capacity Access Prices computed for a series of simulated arrival streams are averaged to produce the bid-price pair or Capacity Access Price to be used in practice. If most arrival streams simulated produce values that are even slightly too high, the average computed, and subsequently used for admission control, will be too high.

Together, this will consequently have the effect requests are rejected that actually should have been accepted. To mitigate this effect, alternative methods of computing the optimal bid-price pair or Capacity Access Price from a series of values computed for single arrival streams may be investigated. Regardless, the performance of different admission control policies is to be compared to each other on a flight-by-flight basis in order to establish the practical consequence of their application on the performance of the flight being forecasted for.

> **Note** In very low-demand situations, it might become attractive to accept any cargo that generates more revenue than the cost of carrying it. This will have to be determined on a flight-by-flight basis.

The fact that a bid-price policy performs better than a policy based on a Capacity Access Price is for the larger part due to the demand being high enough and because it in this situation leads to accepting the most revenue on flights that eventually turn out not to be constrained, as Table 3 shows. Either admission control policy will nevertheless perform better when the demand is higher. The difference between the policies will also be more outstanding. A policy using bid-prices just performs better than one based on a Capacity Access Price. The performance under first-come first-serve will in contrast be bounded at a suboptimal level under higher demand.

4. Statistical Analysis

4.1. Space Request Arrival Rates

The first step in simulating a single flight departure is to determine the number of space requests arriving. The chapter 2.4. Counting Space Request Arrivals explains that space requests arriving do so according to the rules of a counting process and arguments why this well might be the Poisson process.

If arrivals occur according to a Poisson process, this means that the total number of arrivals during the booking period is distributed according to a Poisson distribution. It also means that the time between two subsequently arriving requests is distributed according to an exponential distribution with a parameter corresponding with the mean of the same Poisson distribution.

Instead of straightforwardly simulating a total count from a suitable Poisson distribution, the number of space requests arriving is determined by simulating from the inter-arrival time distribution. The number of space requests arriving during a day is that number for which the sum of the inter-arrival times fits within that day.

Simulating inter-arrival times provides the best flexibility towards expanding the model, because several arrival streams can be combined this way. It has to be known in what order the arrivals occur in a combined arrival stream, or admission control will not be possible. The chapter named 3.1. The Forecasting Model gives an example of the usefulness of this approach in the section named 3.1.3. Fine-Tuning the Model.

Whether the observed arrivals are suitably modelled with a Poisson process is analyzed in the last subsections of this chapter.

The arrival rates, or speeds, are measured in arrivals per time unit of length one, regardless of whether they are measured per working day of eight hours or per week. In this chapter, the measurement unit used will be explicitly mentioned.

4.1.1. Modelling the Pre-Flight Booking Period

The pre-flight booking period is divided into eight labelled intervals in accordance with the day-to-day practice at Polar. Those are the seven days of the week, scheduled flight date being T-0 and the day after the previous flight, six days before scheduled flight date, being T-6. The period before this day is the pre-booking period and space requests arriving in this period are called pre-bookings. It is sometimes useful to indicate this period as being yet another day, labelled T-7, although this virtual day extends as long as necessary into history. Pre-bookings are those bookings that occur a week or more before scheduled flight date and for which the customer specifically instructs Polar to lift the cargo on another than the first flight departing.

Space requests arriving during the seven pre-flight days are modelled to do so according to Poisson processes having arrival rates that are constant throughout that day. The arrival rate is the expected number of arrivals during the whole day. For a single flight, arrival rates are different for each day. In addition, the number of pre-bookings is modelled to have a Poisson distribution with a mean matching the expected number of pre-bookings.

4.1.2. Graphical Overview

For the sake of providing overview, Figure 20 shows weekly arrival rates for flight PO605/6, departing on Saturdays, and PO605/7, departing on Sundays. The graphics in the figure also show the theoretical Poisson distribution function fitted to the observations.

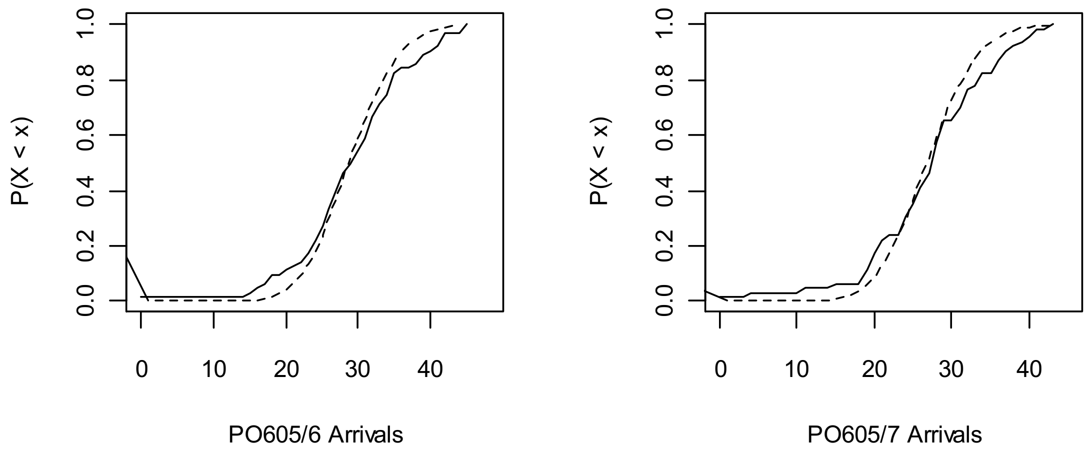


Figure 20 Weekly arrival rates (solid) for PO605/6 (left) and PO605/7 (right), along with a fitted Poisson distribution (interrupted).

The figure below shows the expected arrival rates per day for PO605/6 and PO605/7. The expected number of pre-bookings is plotted to the left the arrival rates of the seven days of the week.

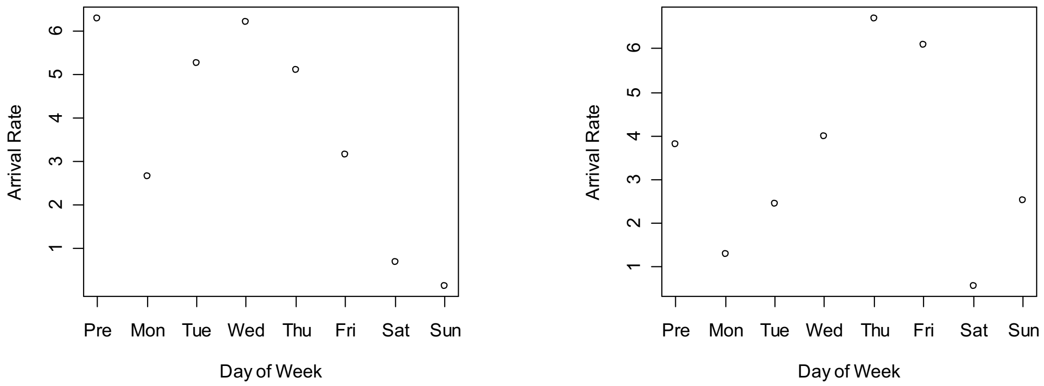


Figure 21 The expected daily arrival rates for each day of the week, as well as for the pre-bookings (Pre). Left arrival rates for PO605/6, right arrival rates for PO605/7.

4.1.3. Numerical Overview

Day of Week	Pre	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
Booking Period	**T-7**	**T-5**	**T-4**	**T-3**	**T-2**	**T-1**	**T-0**	**T-6**	**(all)**
Sample Size	63	63	63	63	63	63	63	63	504
Mean	6,30	2,65	5,25	6,21	5,10	3,16	0,67	0,13	29,47
Variance	8,99	4,84	7,39	9,84	1,86	8,81	3,13	0,24	11.90
Standard Deviation	3,00	2,20	2,72	3,14	1,36	2,97	1,77	0,49	3.45

Table 4 Numerical summaries of the arrival rates for PO605/6. The column Pre concerns pre-bookings.

Day of Week	Pre	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
Booking Period	**T-7**	**T-6**	**T-5**	**T-4**	**T-3**	**T-2**	**T-1**	**T-0**	**(all)**
Sample Size	63	63	63	63	63	63	63	63	504
Mean	3,81	1,29	2,44	4,00	6,70	6,10	0,57	2,54	27,45
Variance	5,61	2,05	4,12	7,48	4,57	3,06	0,86	6,09	10.74
Standard Deviation	2,37	1,43	2,03	2,74	2,14	1,75	0,93	2,47	3.28

Table 5 Numerical summaries of the arrival rates for PO605/7. The column Pre concerns pre-bookings.

4.1.4. Independence of the Arrivals

If the space request really do arrive according to a Poisson process, arrivals during any pair of time intervals should be independent. If arrivals in different intervals really are independent, plotting the arrivals of a pair of booking periods should show points that are randomly scattered. In addition, the correlations of such pairs will be close to zero. Because the graphical method will produce many plots, the latter method is used.

Correlation coefficients are used to compare two series of numbers. They are very useful in detecting linear regression. Correlation coefficients close to one indicate relatively high values in either series are matched with high values in the other. Negative values, close to minus one, indicate high values in either are matched with small values in the other. Values close to zero indicate such relationship is not present. Because any relationship is to be eliminated, only the absolute value is of interest here.

The following table shows in what range the absolute values of the correlation coefficients lie. The fact that half of all correlation coefficients are between -0.1 and 0.1 and almost all correlation coefficients are within -0.3 to 0.3 is reason enough to not doubt the assumption of independent arrivals. There is no linear relationship of significance.

Absolute Correlation	< 0.10	< 0.20	< 0.30	≥ 0.30	Total
PO605/6	14	6	6	2	8 × 7 ÷ 2 = 28
PO605/7	14	8	5	1	8 × 7 ÷ 2 = 28

Table 6 Absolute correlation coefficients indicating whether the arrival rates during different booking periods are related. There are eight differently labelled booking periods. Each unique pair is evaluated.

4.2. Density Classification

The density classification of a space request indicates whether the space request is for dense cargo, volume cargo or cargo reported to be at pivot. In other words, it indicates whether the density of the cargo is respectively smaller, larger or equal to 6 dm^3 per kg. Being a classification, it does not specify the exact density value, but only in what range the exact value is.

The table below lists each density classification with its meaning and the probability that a space request has that classification. This probability is taken to be simply the fraction of airway bills flown having that classification.

Refer to the chapter 4. Density and Volume, below, for information on data quality with regard to the density classification.

Classification	Meaning	Probability
Dense	The density of the space request is less than 6 dm^3 per kg.	36.6%
Volume	The density of the space request is more than 6 dm^3 per kg.	48.2%
Pivot	The density of the space request is 6 dm^3 per kg or the space request is small and the density of the request is unknown or doubtfully correct – e.g. zero or exactly 1 m^3.	15.2%

Table 7 Density classifications with their meaning and the probability of a space request having that classification.

> **Note** To account for round-off errors introduced be the fact that both weight in kg and volume in dm^3 are stored as integers, be sure to allow the density to deviate a small amount from 6. For example, take cargo at pivot to be those space requests having densities that are between 5.99 and 6.01.

The probability distribution of the density classification is an alternative distribution customized to allow for more than one alternative. It can be constructed by labelling the density classifications with a number, for example 1…3, and taking the function to have a value according to the probability that the density classification label is smaller than or equal to the value on the horizontal *x*-axis.

The density classification can be simulated by drawing from the standard uniform distribution on (0, 1) and choosing dense when the number does not exceed 0.366, choosing pivot if the number drawn exceeds 1 – 0.152 = 0.848 and choosing volume otherwise.

4.3. Actual Weight

The actual weight of a shipment is the primary indicator of its size. The actual weight is the real weight of the shipment measured on a balance.

A different distribution is provided for each density classification, because the weight distribution of space requests is different depending on the density classification.

4.3.1. General form of the Actual Weight distributions

All three actual weight distributions are Gamma distributions with different parameters. Refer to the appendix A. Important Probability Distributions for more information on the Gamma distribution and on how to do computations with it like simulating from it.

Although the weight distribution of space requests depends on the density classification, this dependence is of the most straightforward form. Statistical modelling is performed by partitioning the observations based on the density classification and then estimating parameters for each partition as if there was no dependency. When the density classification is known, simply selecting the correct distribution and drawing from it suffices for simulating weights of a series of space requests.

For the Gamma distribution, analytic arguments for estimating parameters have proven to work the best. Simply computing the scale to be mean over variance, the shape to be mean over rate and the rate to be 1 over scale outperforms more difficult techniques consistently with a margin.

The figure below shows the distribution functions of actual weight for each density classification. Space requests constituting of volume cargo are typically the smallest and space requests for dense cargo are typically the largest. Hence, the distribution corresponding with volume cargo requests is on top in the graphic and the distribution corresponding to dense cargo requests is the one assigning the lowest probabilities to requests with matching sizes.

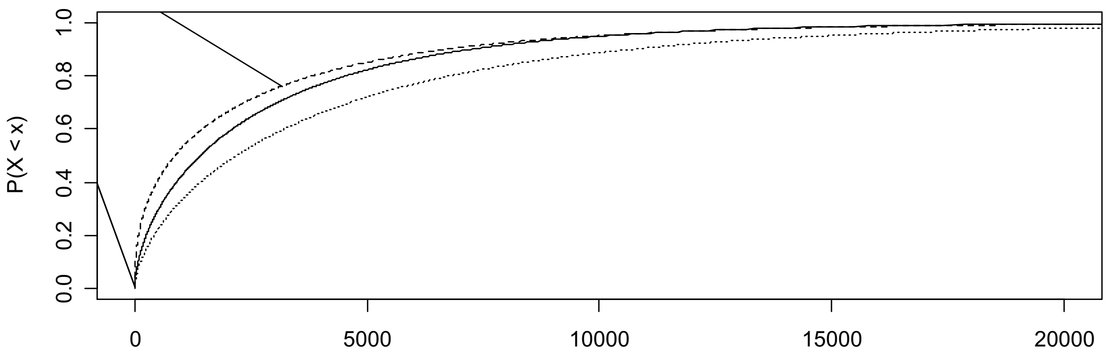


Figure 22 Distribution functions of the Actual Weight of Dense Cargo (dotted), Volume Cargo (interrupted) and Cargo at Pivot (solid). The graphic has been cut-off at 20 tonnes.

4.3.2. Numerical Overview

A numerical overview of the available data on actual weight is provided in the table below. The parameters of the Gamma distribution estimated for each density classification partition are listed first. The sample size is the number of observations included. The data includes all shipments for the first PO605 flight of 2004, with a scheduled flight date of January 4, up to the PO605 flight of March 27, 2005. Mean and variance are the sample mean and variance. The standard deviation is the square root of the variance. The remaining values are quantiles of the empirical distribution, the minimum being the 0% Quantile, the median being the 50% Quantile and the maximum being the 100% Quantile.

Property	Dense Cargo	Volume Cargo	Cargo at Pivot
Gamma Distribution – Shape	0.61	0.40	0.58
Gamma Distribution – Rate	1.5E-4	1.7E-4	2.1E-4
Gamma Distribution – Scale	6,761	5,877	4,752
Sample Size	1,332	1,753	552
Mean	4,115	2,369	2,738
Variance	27.8M	13.9M	13.0M
Standard Deviation	5,222	3,746	3,607
Minimum – Zero	5	1	1
25% Quantile	659	280	282
Median	2,035	905	1,140
75% Quantile	5,467	2,950	3,793
Maximum – Positive Infinity	37,938	45,746	18,535

Table 8 A numerical summary of the actual weight for each density classification.

4.3.3. Extremely Large Shipments

Please note that the exact value in the table above of the maximum actual weights observed is hardly informative, although those observations illustrate the shape of the distribution function very well. The largest maximum observed is found in the partition with the most samples. This is only proof of the fact that with seeing more observations, seeing more extreme ones becomes more likely.

It is best practice to remove extreme points if their influence is so large that their inclusion leads to misrepresenting the sample through the estimated distribution. The question remains whether there are such extreme values and if so, how much of them should be disregarded. To this end, graphical methods are invaluable to start with. Next, verifying estimations of the data both with and without the extreme points may be necessary.

The graphics below give reason to pay special attention to the largest space requests for volume cargo. The extreme points displayed for volume cargo are far more off than the extremes amongst the other density classifications and they are occurring close to each other. Especially the shipments of above 30 tonnes might be incidental observations, whose exclusion might provide better forecasting accuracy.

The largest shipment for example, which is 45,746 kg, has the effect of increasing the mean over 12 weeks of the weight of volume cargo with 140 kg or 6%, and the standard deviation with 714kg or 20%. This effect largely disappears 12 weeks later, when this specific sample does not contribute anymore to the moving average.

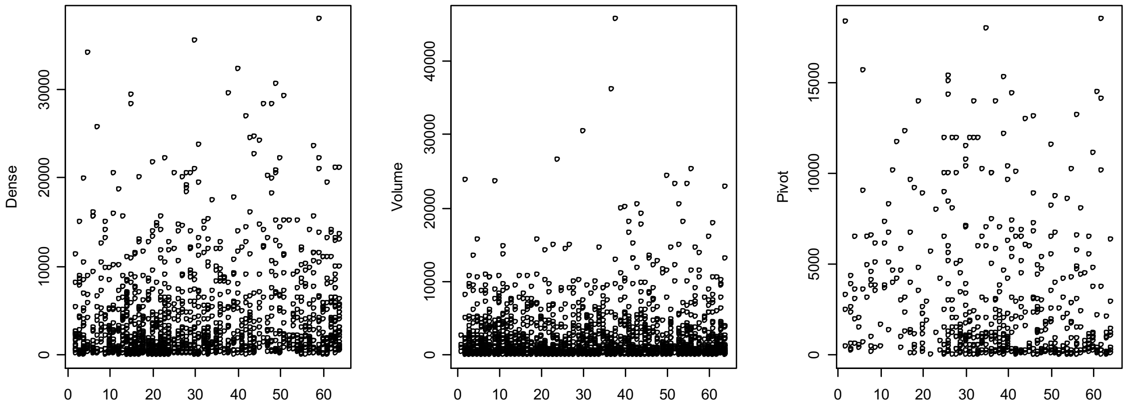


Figure 23 All Actual Weight samples for each density categorization on a timeline.
The value at the horizontal axis is the number of weeks from January 1, 2004.

Next to just removing them, another way to deal with the extreme values is using robust statistics for mean and variance, instead of well known and easy to compute mathematically spoken optimal variants. A robust statistic is a statistic that is relatively insensitive to appearance of extreme values in the data.

For estimation of the parameters of the actual weight distribution with analytic means, the mean can be replaced with an *α-trimmed average* and the variance can be replaced with the square of a variant of the *median absolute deviation* corrected for use with the Gamma distribution. The former converges to the mean without change, if the number of samples increases. The latter converges to the standard deviation after it is multiplied with a constant specific to the Gamma distribution, if the number of samples increases.

4.4. Density and Volume

The density of a space request specifies how much space it takes per unit of weight. The volume in m^3 of a space request can be obtained by multiplying its density with its weight in tonnes. Because of this relationship, only density is incorporated into the model as a property subject to statistical analysis. The volume of the space request can be computed when the density is known.

4.4.1. Graphical Overview

The graphics below give an overview of the densities of all space requests available. The graphic on the left is a box-and-whisker plot. The box encloses 50% of all samples. The thick line in the centre of the box is drawn at the median. The lines extending from the box are called the whiskers. The lower horizontal line is drawn at the minimum density observed. Values that extend beyond the whiskers are called outliers and are all individually represented with a dot.

The graphic on the right is the empirical distribution function of the samples. Note that this graphic is cut-off at a value of thirty. Values of over 30 are not frequent, but no exception either. The graphic shows very clear the reason why a density classification is introduced. The function can be separated into three distinct parts. The linear part on the left shows the densities of dense cargo, the vertical part does the same for cargo at pivot and the curved part on the right for volume cargo.

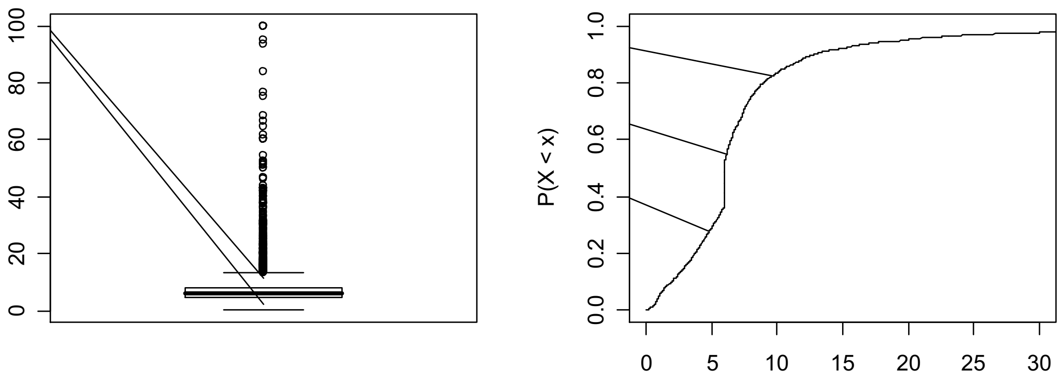


Figure 24 A graphical overview of the density of space requests.
Left a boxplot of all densities, right the empirical distribution.

Plots of weight and density of dense and volume cargo are displayed below to give an idea of the data. In addition, those plots illustrate how a dependency between two random variables can be recognised.

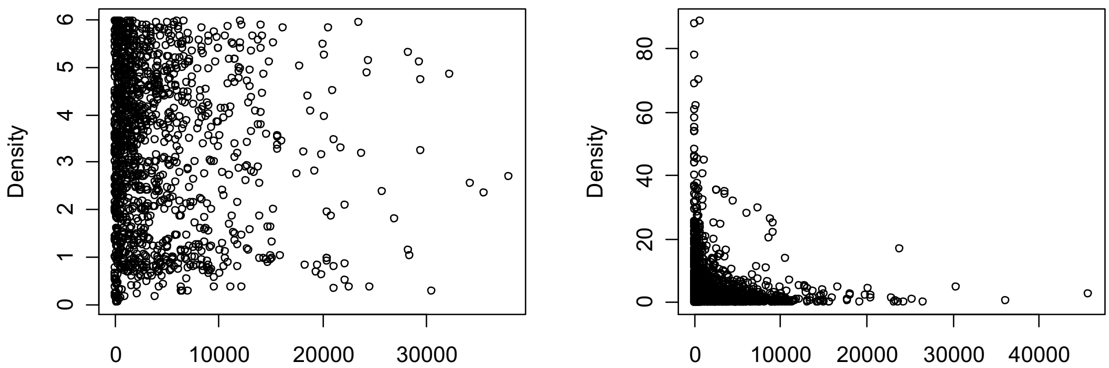


Figure 25 Scatter plots of Actual Weight against Density of Space Requests for Dense Cargo (left) and Volume Cargo (right). The plot on the right shows a clear dependency, in contrast to the plot on the left.

Both the above plots can be sliced in parts, producing a series of small plot parts side-by-side, each part representing a different range of weights. In the plot on the left, each slice would show a distribution of points that is more or less equal across the full vertical range of that slice. This indicates that values for density in the range from 0 to 6 are equally likely to be observed for any dense cargo request with any weight. The fact that most points are in the left part of the graphic is the result of heavy shipments being less frequent.

Slices of the plot on the right, however, would show less and less points in the upper part of each slice. This is in addition to the fact that the upper part of the graphic shows fewer points than the lower part as a whole; the comparison between slices is of importance. Higher density values are increasingly unlikely to be observed for larger volume space requests. If there were no differences between slices, each slice would still indicate higher density values are less frequently observed, but not due to a connection with actual weight.

4.4.2. Numerical Overview

A numerical overview of the available data on density is provided in the table below.

Property	Dense Cargo	Cargo at Pivot	Volume Cargo
Distribution	Uniform	Deterministic	Gamma (shifted)
Distribution Parameters	Min; Max	Constant value	Shape; Scale; Shift
Parameter Values	(0.359; 6.255)	6	(0.354; 15.19; -6)
Sample Size	1,332	552	1,753
Mean	3.31	6	11.38
Variance	2.92	0.0000	81.68
Standard Deviation	1.71	0.0014	9.04
Minimum	0.04	5.99	6.01
25% Quantile	1.78	6	6.80
Median	3.50	6	8.17
75% Quantile	4.81	6	11.86
Maximum	5.99	6.01	94.99

Table 9 A numerical summary of the density for each density classification.

The density classifications are each fitted with a different distribution. The deterministic distribution (*ontaarde verdeling*) is not really a distribution; çargo at pivot has a density of 6 by definition.

4.4.3. The Linear Model for Dense Cargo

The graphic below shows the distributions fitted for dense and volume cargo.

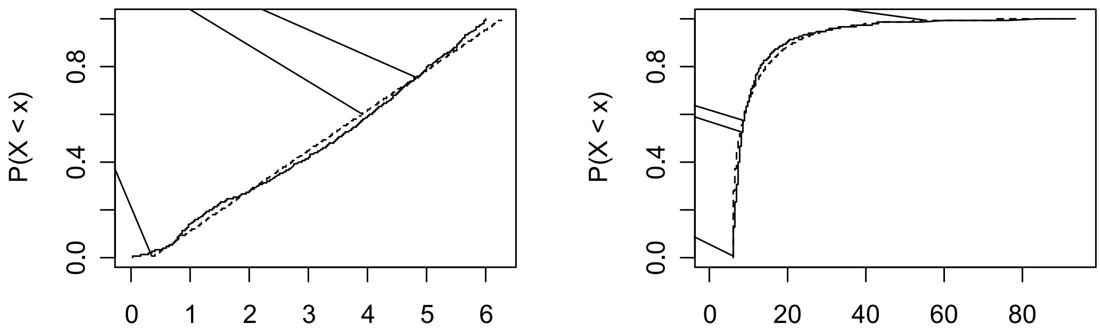


Figure 26 The empirical distribution (solid) and fitted estimate (interrupted) of the density for dense cargo (left) and volume cargo (right).

The density distribution for dense cargo is fitted using a linear model, solved by minimizing the sum of squared residuals. Hence, the fitted line is a statistically optimal fit, because by minimizing the squared residuals, the line obtained is the line whose horizontal distance is closest to the empirical distribution, when the distances are evaluated at the values observed. Please note that the horizontal distance is indeed of interest, rather than the vertical distance, because the data used to fit the linear model is on the horizontal axis. The values estimated are the intercept at the horizontal axis and the slope in horizontal direction. The linear model used fits the values observed given the probability of each observation according to the empirical distribution.

> **Fitting the linear model** A sorted list of all 1,332 density values is the input to the linear model (the response) alongside a list with the empirical probabilities, which are 1…1,332 divided by 1,332 or 1/1,332…1 each a fraction 1/1,332 apart (the effects).
> An initial guess could for example be a vertical line crossing the mean of the response. At each of the effect values, distances are computed from this line to the response corresponding with that effect. Those 1,332 values are then squared and summed up to produce the sum of squared residuals.
> After each such computation, the intercept and slope of the line are changed until the residual sum of squares cannot be reduced further reasonably (big values indicating a linear model does not fit at all) or it reaches zero. Further verification of whether a linear model fits the data can be obtained using *analysis of variance*.

Uniform Distribution	Expression	Remarks
Intercept	0.359	
Slope	5.896	
Range	(0.359, 6.256)	
Sample Size	1332	
Distribution Function	$F(x) = x / 5.896 - 0.0609$	x in (0.359, 6.256)
Quantile Function	$Q(x) = 0.359 + 5.896 \times p$	p in (0, 1)

Table 10 A summary of the uniform distribution of the density of dense cargo fitted with a linear model.

Please note that the fitted density distribution will not produce values exactly matching the definition of dense cargo. Very small values for density will never be produced and some dense cargo requests will be simulated with densities of 6 and slightly higher. Simulation from this distribution, however, produces values matching the data better across the board than when simply simulating from the uniform distribution on 0 to 6.

> **Note** Do not verify simulated data based on the density values of the simulated space requests. The fraction of dense cargo will be lower than expected and the weight, density and revenue distributions that can be obtained from the simulated data will be misrepresenting live data. Use the density classification label simulated instead, for verifying simulated data.

4.4.4. The Conditional Distribution for Volume Cargo

Figure 25, above, shows proof of a dependency between the actual weight and density of volume cargo. As concluded before, higher density values are increasingly unlikely to be observed for larger volume space requests. This must be taken into account when modelling the density of volume cargo.

Several techniques can be applied for modelling dependent data in general. The appendix C. Forecasting Techniques for Dependent Data describes those in some detail. The technique used for the density of volume cargo is Maximum Likelihood Estimation. An approach based on bins containing samples with similar actual weights is used to obtain an initial understanding of the relationship. The latter technique compares well to the intuitive way of partitioning the graphic in vertical slices.

The main assumption made about the conditional distribution is that the density of volume cargo is distributed according to a shifted Gamma distribution, since that distribution fits the data well if its dependency on actual weight is ignored. This is shown in Figure 26 in the plot on the right. The data is shifted to the left, such that its values range from zero to infinity, matching the support of the Gamma distribution, instead of having a minimum value of 6.

Another assumption required is that of the form of the dependency. Figure 25, above, suggests the average density of volume cargo declines very swiftly for small values of actual weight, but that this decline is also stabilizing fairly early on. In addition, the range of values observed is shrinking, indicating a decreasing spread. Hence, both mean and variance of the density distribution are modelled to be a general form polynomial function of the actual weight.

This means that both shape and scale of the density distribution of a single shipment are modelled to be the actual weight of that shipment raised to a power and multiplied with a constant. Inclusion of an intercept tends to decreases the quality of the estimation. Maximum likelihood estimates of the power and multiplier for both shape and scale are determined in a single step, simultaneously.

Gamma Distribution	Shape	Scale
Multiplier	0.56039	48.98255
Exponent	0.03456	-0.29847
Distribution Parameter	0.56039 × *ActualWeight*$^{0.03456}$	48.98255 × *ActualWeight*$^{-0.29847}$

Table 11 A summary of the maximum likelihood estimate of the density of volume cargo.

Example The density of volume cargo shipments with an actual weight of 10 tonnes has a gamma distribution with shape 0.77 and scale 3.13. Recall that the density had been shifted by -6. The density thus has a mean of 6 + 2.4 and a standard deviation of 2.75. Hence, they take up 84 m^3 on average and 95% have volumes between 60 and 160 m^3.

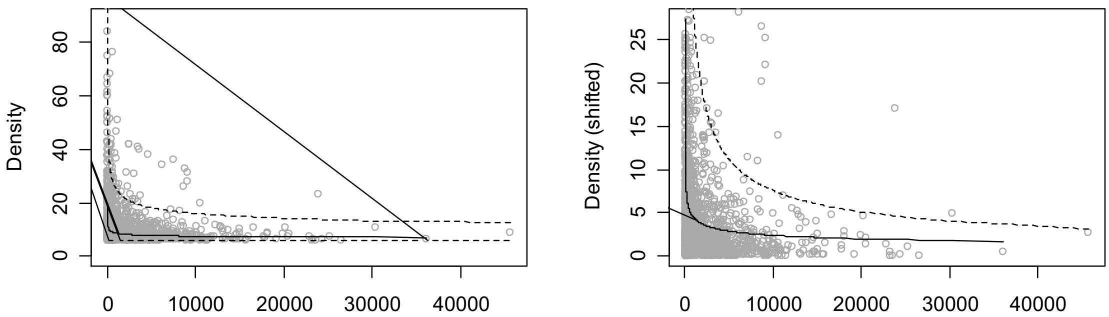


Figure 27 The conditional distribution of the density of volume cargo. On the left median (solid) and a 95% confidence region (interrupted). On the right mean (solid) and variance (interrupted), both on a backdrop of the observations. The graphic on the right displays densities shifted by –6.

4.4.5. Data Quality

The figure below shows a big change over time in the number of requests for cargo at pivot. While the number of requests for cargo at pivot is steady, at the start of Q3 2004, it increases, for around 12 weeks, about as fast as the number of requests for dense cargo decreases. During this period, the total number of requests roughly remains the same. When looking more closely at Figure 23, above, this change is also visible.

Although a reason cannot be identified for certain, this is expected to be due to a change in flight management software. Most likely, this observation does have adverse effect on the forecasting model. Anything else remaining the same, it at least has the effect of making the model more complex. Simply merging the cargo at pivot classification back into the dense cargo classification is not possible.

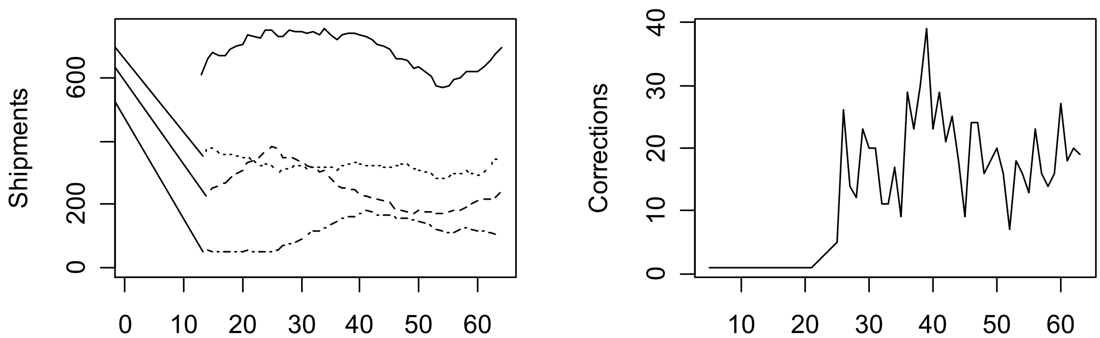


Figure 28 Left the number of shipments per week for dense cargo (interrupted), volume cargo (dotted), cargo at pivot (alternating) and total (solid). These values are 12-week averages. Right the number of shipments for which no volume information is available. The value at the horizontal axis is the number of weeks from January 1, 2004.

4.5. Revenue

The day-to-day practice at polar defines the revenue of a request to be the rate computed over the chargeable weight plus fuel and security surcharge rates multiplied with the actual weight of the request. Other surcharges may apply based on the nature of goods, amongst others. The chargeable weight used in this computation may be different from the chargeable weight based on the pivot rule, computed using the actual weight and the density of the shipment.

Because the revenue payable is due in foreign currency for customers outside the Euro zone, the revenues are expressed in different currencies. They are converted to Euros using a default, weekly-determined exchange rate. This exchange rate is the one also used in the flight revenue report at Polar Air Cargo.

For statistical analysis, the only surcharges that are considered special are the fuel and security surcharges. Any other surcharges applicable are contributing towards the total revenue for the request, but are not separated from it or analyzed by themselves.

Fuel and security surcharges are excluded from revenue in the following analysis. These rates seldom change, and when they do change, this is a result of a company-wide, strategic decision. They are therefore considered constant. The amount of revenue generated due to fuel and security surcharges is deducted from the total revenue generated by a shipment. What remains is the revenue generated by a unified rate per kilogramme.

> **Note** The revenue used in the statistical analysis is the total revenue as paid by the customer, in Euros, excluding fuel and security surcharges. This is true regardless of whether special conditions are negotiated or whether the customer pays additional surcharges for a specific shipment, due to for example the shipment containing dangerous goods.

4.5.1. Graphical Overview

Figure 29 gives an overview of revenue for the three density classifications. The empirical distributions for each density classification in the figure disregard any relationship with actual weight or density that might be present. It shows that the overall revenue distributions are very similar for different density classifications.

The remaining graphics show revenue against both actual weight and density. Hopefully, larger shipments generate more revenue. This shows very well in the middle graphic. For dense and volume cargo, the density provides additional measure of overall size. However, whether or not increasing density implicates higher revenue potential is not obvious from the graphics below.

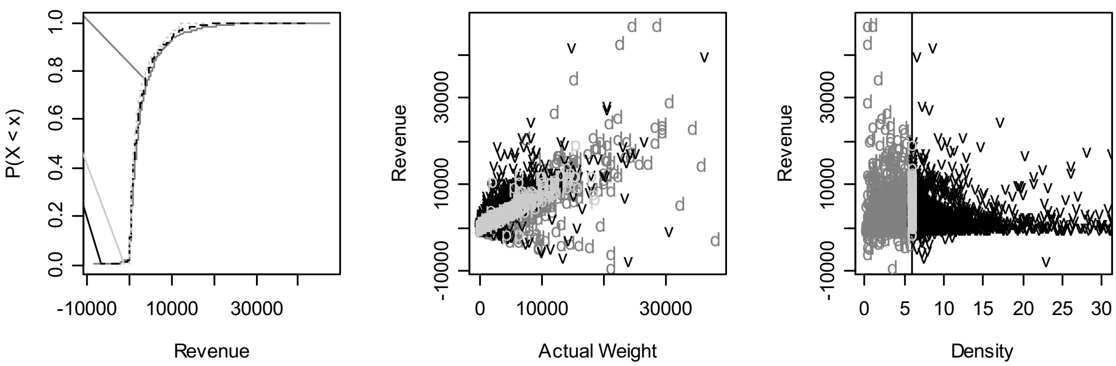


Figure 29 Empirical distribution of revenue (left) and scatter plots of actual weight (middle) and density (right) against revenue. Data for each density classification is differently drawn. The revenue shown is total revenue excluding fuel and security surcharges.

Although the graphic on the right shows generally much smaller revenues for volume cargo shipments with larger densities, this does not mean the relationship with revenue is such that an increasing density results in lower revenue. The dependency between actual weight and density is just displaying here. Shipments with larger densities most often are shipments with very small actual weights and this relationship is most definitely of biggest influence on revenue. Evidently, actual weight really is most useful as the sole measure of the size of a shipment.

4.5.2. Numerical Overview

The table below provides a numerical overview of the revenue of different density classifications as well as of all cargo together.

Property	Dense Cargo	Cargo at Pivot	Volume Cargo	All Cargo
Sample Size	1,332	552	1,753	3,637
No Revenue or Loss	33	17	31	81
Mean	3,350	2,447	2,856	2,966
Variance	22.5E+6	9.3E+6	15.4E+6	17.2E+6
Standard Deviation	4,739	3,055	3,926	4,144
Minimum	-8,562	-1,559	-7,402	-8,562
25% Quantile	671	452	467	523
Median	1,781	1,174	1,473	1,546
75% Quantile	4,246	3,070	3,644	3,756
Maximum	47,459	19,325	42,045	47,459

Table 12 A numerical summary of the revenue of each density classification.

The minimum total revenue including fuel and security surcharges recorded is zero. Shipments that are reported here to have incurred a loss are those that have negative revenue when fuel and security surcharges are deducted from the total revenue. The numbers includes some shipments of company materials. Some 2% of all samples are negative. These are excluded from the investigation.

4.5.3. Determining Factors Affecting Revenue

Possibly, both actual weight and density are of influence on the revenue of a shipment. Because three variables are involved in this relationship, graphical methods become much more difficult. For this reason, segmenting the samples into bins is very useful for analyzing revenue. The figure below is created using this approach.

The samples shown in each of the graphics in the figure are those having densities in a limited range. Due to the densities being close together, the effect of the density on revenue is comparable for all samples in the same plot. However, because the density range of each plot is different, differences between plots indicate a dependency between density and revenue.

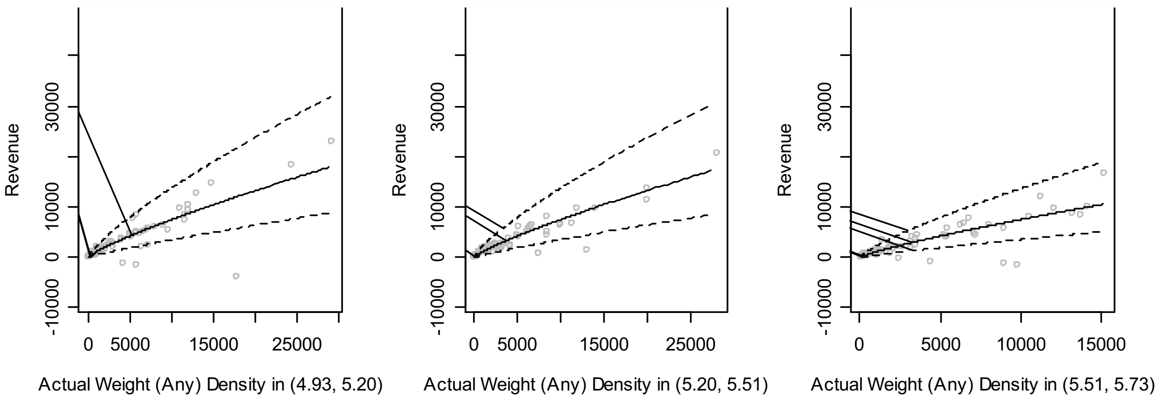


Figure 30 Two scatter plots of revenue against actual weight of samples having densities in two disjoint ranges. Both plots show median (solid) and a 95% confidence range (interrupted) on a backdrop of the observations. The samples shown are dense cargo. The revenue shown is total revenue excluding fuel and security surcharges.

The plots in Figure 30 are for dense shipments only. A similar series of plots can be created for volume shipments. Note that the density ranges in the figure are adjacent and that the mean and confidence range displayed are quite different. This provides proof that the revenue must be modelled to be dependent on density for both dense and volume cargo. Obviously, revenue of cargo at pivot will be modelled differently, since the densities of those shipments are fixed. The confidence range and median displayed are obtained using the model laid out below.

4.5.4. Modelling Revenue

Figure 29 and Figure 30 show proof of a relationship between actual weight and revenue as well as density and revenue. The former is almost linear, as these figures show beautifully. The form of the latter is however not yet established. It is to be expected that a higher density will affect the revenue positively. Moreover, this relationship can be expected to be strict; revenue increases monotonous across the range of possible density values. Given a shipment with known actual weight and certain density, it is always better if it turns out to be of a little bit higher density.

A polynomial dependency of general form is flexible enough to model both dependencies, in a way similar to the model used for the dependency between actual weight and density. Also for revenue, the distribution family is the gamma distribution. Similar to the model used for the density of volume cargo, both shape and scale parameters of this distribution are modelled as polynomial functions, in this case of both actual weight and density.

A polynomial relationship of the form $p = i + m \times (a\text{-}s)^x \times (d\text{-}t)^y$ has shown to work very well. Here p is the distribution parameter (the effect), either shape or scale of the Gamma distribution. a and d are actual weight and density respectively (the factors), both shifted by a different constant value. Inclusion of a vertical shift again has adverse effects, so i is fixed to be zero. Inclusion of horizontal shifts s and t however do improve the distribution fit, especially for small values of actual weight and density. For both shape and scale, the same values for s and t are used. The multiplier and exponents are different for shape and scale. Although revenue is dependent on two factors, only the exponent is necessary twice, because duplicate intercepts or multipliers do not change this function formally.

> **Note** Although using more parameters than necessary is bad practice to boot, using a parameter that can be chosen freely due to another one cancelling its effect introduces a chance the optimization algorithm will choke, return errors, or take a long time to complete.

The table below details the model that is used and the values estimated. The values displayed are estimated using Maximum Likelihood Estimation, both for shape and scale simultaneously for a single density classification.

> **Note** Please note that the extended model presented here can be used for cargo at pivot as well by requiring the shift and exponent of density to be zero. In addition, note that for volume cargo, the density values input to the model already were shifted by -6. The actual shift value is hence -3.398 = 2.602 - 6.

Classification	Parameter	Polynomial Function
Dense Cargo	**Shape**	$1.333 \times (actualWeight + 2.077)^{0.123} \times (density + 0.868)^{0.414}$
Dense Cargo	**Scale**	$3.132 \times (actualWeight + 2.077)^{0.692} \times (density + 0.868)^{-0.404}$
Volume Cargo	**Shape**	$2.509 \times (actualWeight + 2.314)^{0.163} \times (density + 2.602)^{-0.225}$
Volume Cargo	**Scale**	$1.169 \times (actualWeight + 2.314)^{0.641} \times (density + 2.602)^{0.710}$
Cargo at Pivot	**Shape**	$0.468 \times (actualWeight + 26.90)^{0.404}$
Cargo at Pivot	**Scale**	$7.533 \times (actualWeight + 26.90)^{0.432}$

Table 13 Functions describing the parameters of the estimated conditional revenue distribution for each density classification. A value of 6 must be deducted from the density argument for volume cargo before parameters are computed.

> **Example** Consider a series of shipment all of close to 1.5 tonnes and volumes of approximately 4,5 m^3. The shipments are dense with a density of 3. Their revenues excluding fuel and security surcharges have a Gamma distribution with shape 5.753 and scale 287.3. This distribution has a mean of € 1,650. Its 95% confidence range is from € 590 to € 3,250. The fuel and security surcharges are € 0.55 at the end of the data collection period. This adds € 825 to the revenue per shipment. The shipments therefore would generate on an average € 2,475 of total revenue each, with 95% of the shipments generating somewhere between € 1,415 and € 4,075.

The graphics below show mean and 95% confidence range of the conditional revenue distributions for dense and volume cargo. Refer to the appendix titled D. Additional Graphics and Tables for additional graphics and tables regarding the conditional revenue distribution.

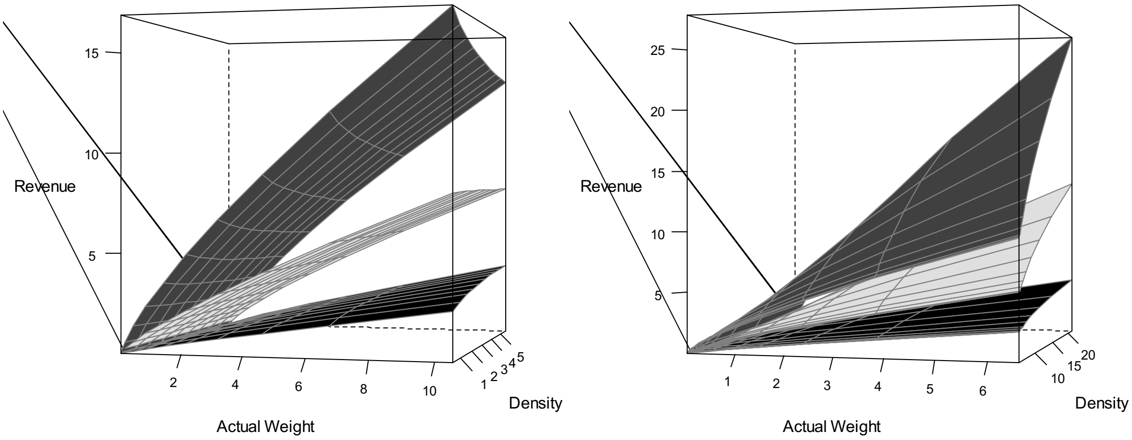


Figure 31 Perspective views of mean (light) and a 95% confidence range (dark) of the revenue distribution for dense cargo (left) and volume cargo (right). Actual weight and revenue are divided by 1000. 10% of the largest actual weights and densities are excluded from the graphics. The revenue shown is total revenue excluding fuel and security surcharges.

The graphic on the left shows that the revenue expectation for a dense shipment is almost indifferent with regard to density, except that a larger density implicates more predictable revenue behaviour. For this reason, dense cargo with higher density is preferred from a business perspective.

An interesting fact showing somewhat obfuscated from the above figure is that the revenue expectation of cargo with densities close to 6 is not at all behaving continuously. The expected revenue of a shipment differs by a considerable ratio depending on whether it is being classified as dense cargo or volume cargo, as Figure 32 shows. With regard to revenue, cargo at pivot hardly differs from dense cargo.

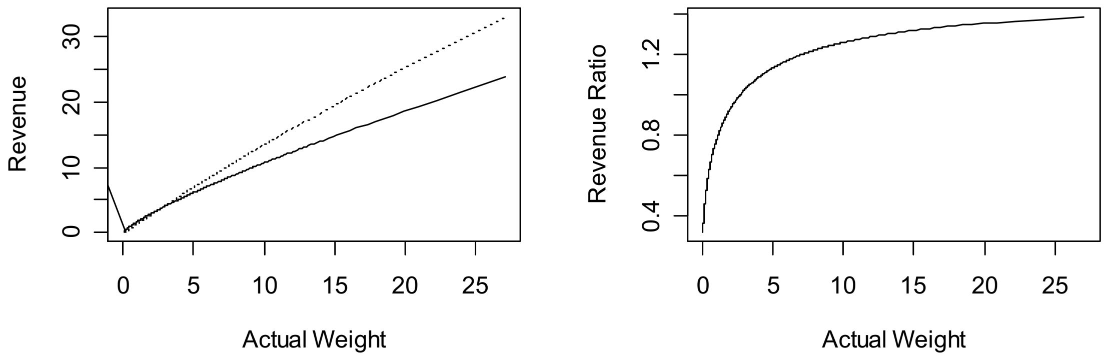


Figure 32 Left expected revenue for dense cargo (solid) and volume cargo (interrupted) against actual weight. Densities are taken to be as close as possible to 6. Right the ratio of these estimates. Actual weight and revenue are divided by 1000. The revenue shown is total revenue excluding fuel and security surcharges.

This behaviour is also visible in the observations. These show that the chargeable weight of all shipments with densities between 6 and 6.25 is 300 kg higher than that of all of those with densities between 5.75 and 6. The actual weight is 260 kg higher.
The revenue of the volume shipments is on average € 530 higher with the revenue of the dense shipments being € 2,650. The rate per kilogramme chargeable weight excluding surcharges is on average € 0.12 higher with the rate for the dense shipments being € 1.07.

4.5.5. Verifying the Estimated Revenue Distribution

The following steps are taken for each density classification to produce a custom quantile-quantile plot that visualizes the quality of the conditional revenue distribution estimate.

- A probability is computed for each observation according to a normal empirical distribution using all observed revenue values. With n samples, these range from 1/(*n*+1) to *n*/(*n*+1).
- For each observation, the conditional revenue distribution is determined. This means shape and scale parameters are computed using the functions in Table 13.
- Using the probability, shape and scale from the previous steps, quantiles of the gamma distribution are plotted for each sample against that samples revenue.

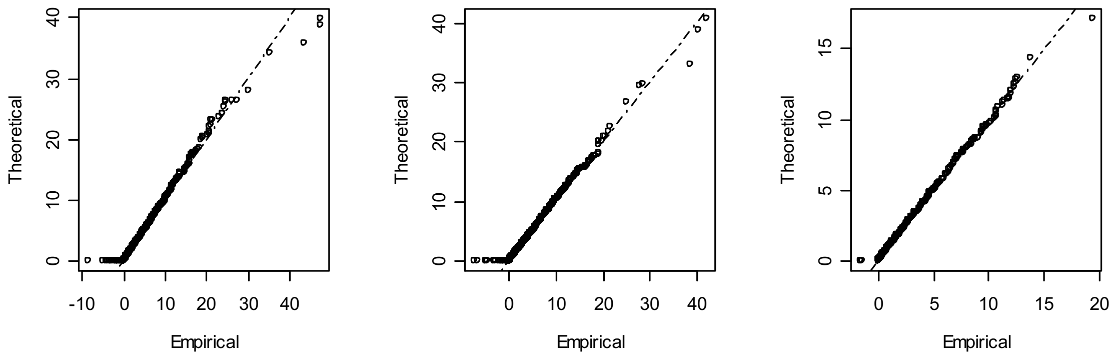


Figure 33 Plots of the revenue observations of dense cargo (left), volume cargo (middle) and cargo at pivot (right) against quantiles of their individual conditional distributions. The interrupted line corresponds with quantiles of a perfect fit. The revenue shown is total revenue excluding fuel and security surcharges, divided by 1000.

Figure 34 shows that the conditional revenue distributions estimated using maximum likelihood perform very well. An exception of some importance is that there are revenues excluding fuel and security surcharges that are below zero. These naturally divert from the estimated distributions, because these support only non-negative values by definition. However, the number of observations this applies to is small.

4.6. Trend and Seasonality

Next to estimating distributions based on all of the samples collected throughout the sampling period, it is interesting to investigate whether trend and seasonality play a role of importance.

If the estimated distributions change linearly with time, this is called a trend. If there is a deviation from the trend throughout the year, this is called seasonality. Both trend and seasonality indicate there is a dependency of the distribution parameters on time.

The main reason to investigate trend and seasonality is to determine whether it is required to simulate using parameters that are estimated based on only a limited period, or whether it suffices to use parameters estimated using all available data.

Trend and seasonality are for the most part investigated using basic methods like moving averages. If significant changes are detected using these methods, the parameters of the model established using all data will be re-estimated, but the model will not change.

4.6.1. Space Request Arrival Rates

Figure 34 shows the development of the daily arrival rates for PO605/6 and PO605/7 week-by-week. The graphics show that the number of space requests fluctuates importantly throughout the year and that the number of sales peaks during the summer season. Both the beginning and end of the year are weaker periods with regard to the number of sales.

The trend line is only included to make the fluctuations throughout the data collection period better visible.

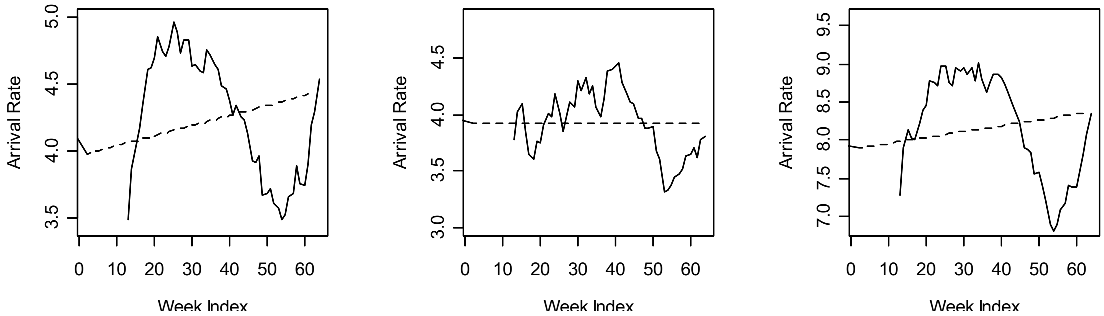


Figure 34 Expected daily arrival rates for PO605/6 (left) and PO605/7 (middle) and total (right) based on a 12-week moving average (solid) and a linear fit showing the trend line (interrupted). The value on the horizontal axis is the number of weeks from January 1, 2004.

The above implicates that the arrival rates for each of the booking periods T-0 to T-7 change from week to week. This is an important fact. The arrival rates are expected to be the most important instrument in tweaking the model for maximum forecasting power.

There are several ways to do this. It is for example possible to specify a change in weekly arrival rate and use daily arrival rates for simulation that are marked up or down by the same fraction. Directly analyzing the arrival rates of the booking periods T-0 to T-7 is also possible. This will uncover that, although the former method will be useful, the arrival rates for different days do not change in concert throughout the data collection period. This might be due partially to sales personnel working weekend shifts during only part of the year.

4.6.2. Actual Weight

The figures below give an overview of the time-dependency of each of the actual weight distributions. Next to mean and standard deviation computed over 12-week periods, the graphic also shows an indication of the relative sample sizes from which the estimates are computed. The horizontal lines are the estimates obtained using all data.

Except for cargo at pivot, shipments are heavier closer to then end of the year. Why cargo at pivot is becoming lighter in contrast may be due to more shipments that should have been recorded to be volume cargo are classified as cargo at pivot instead.

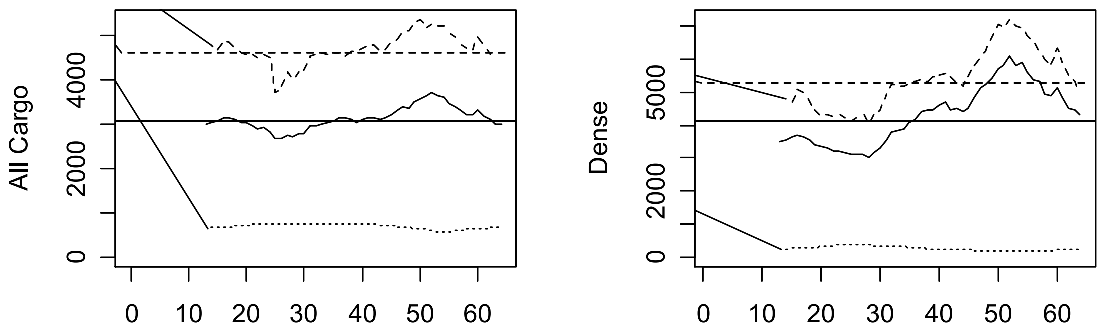


Figure 35 Plot of mean (solid), standard deviation (interrupted) and the sample size (dotted) of the actual weight distribution for all data and for dense cargo on a timeline. Each point is an estimate over the 12 past weeks. The value at the horizontal axis is the number of weeks from January 1, 2004.

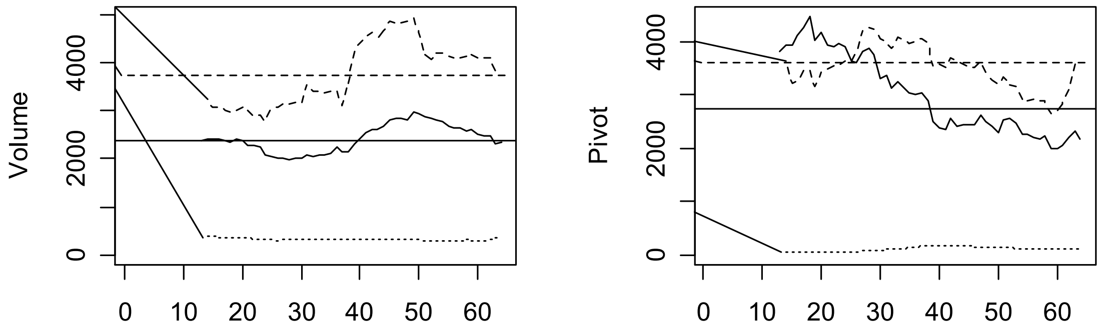


Figure 36 Plot of mean (solid), standard deviation (interrupted) and the sample size (dotted) of the actual weight distribution for volume cargo and cargo at pivot on a timeline. Each point is an estimate over the 12 past weeks. The value at the horizontal axis is the number of weeks from January 1, 2004.

In combination with Figure 34, the graphics show that individual shipment sizes decrease when the number of shipments increases. It seems that, if there are more requests, this is for a larger part due to an increasing number of smaller shipments. This can be attributed to low availability of cargo only aircraft, with carriers having those available attracting large shipments regardless of relative market activity. Hence, an increasing demand will show primarily from an increasing number of smaller shipments. The lower increase of large shipments is however not fully explained stating this. Why is the number of larger shipments not increasing along with the demand? Maybe larger shipments are more frequent as well and split by the customer.

The fact that part-shipments occur somewhat more often during mid 2004 and towards the end of the sampling period is no explanation, because part-shipments are ignored while establishing distribution estimates.

Note that the figures above show a relationship between the mean and variance. A decreasing number of small space requests will have the effect of both an increasing average shipment weight and an increasing standard deviation, because the average shipment weight will remain relatively small.

4.6.3. Density Classification and Density

The figure below shows the fraction of shipments being dense cargo, volume cargo or cargo at pivot as a percentage of the weekly total and the average density of the different density classifications.

Although the latter is steady for cargo at pivot by definition, dense cargo is becoming somewhat heavier towards the end of the data collection period and the density of volume cargo fluctuates quite a bit.

With regard to the fraction of cargo being dense or cargo at pivot, it remains to be seen whether the behaviour showing in Figure 37 will repeat itself.

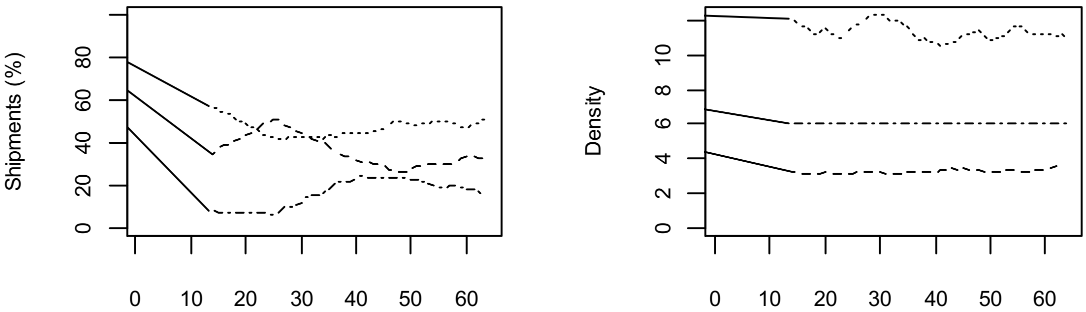


Figure 37 Left the fraction of shipments with certain density classification, right the average density for dense cargo (interrupted), volume cargo (dotted), cargo at pivot (alternating) These values are 12-week averages. The value at the horizontal axis is the number of weeks from January 1, 2004.

4.6.4. Revenue

Figure 38 shows the average revenues per shipment for each density classification as well as averages for all cargo. Due to this value changing so fast for dense cargo around the end of 2005, all revenues of individual dense cargo shipments are plotted in the graphic on the right. The high average seems to be caused by a smaller number of lower revenue dense cargo shipments.

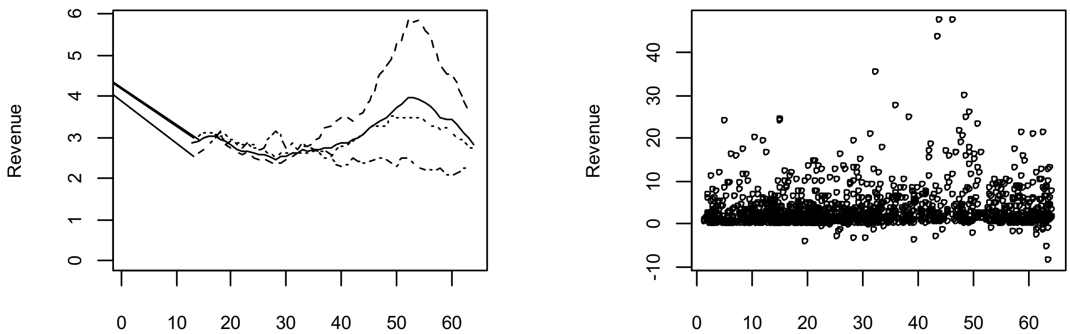


Figure 38 Left, 12-week averages of the average revenue per shipment for dense cargo (interrupted), volume cargo (dotted), cargo at pivot (alternating) and total (solid). Right the revenue for individual dense cargo shipments. The revenue shown is total revenue excluding fuel and security surcharges, divided by 1000. The value at the horizontal axis is the number of weeks from January 1, 2004.

4.6.5. Overall Flight Performance

Figure 39 shows total revenue and details the amount of fuel and security surcharges charged from week to week for both PO605/6 as PO605/7. Although not of direct importance for the forecasting model in the phase of parameter estimation, it is perhaps interesting to see how changes in key statistics like average actual weight per shipment work together to produce flight revenue totals.

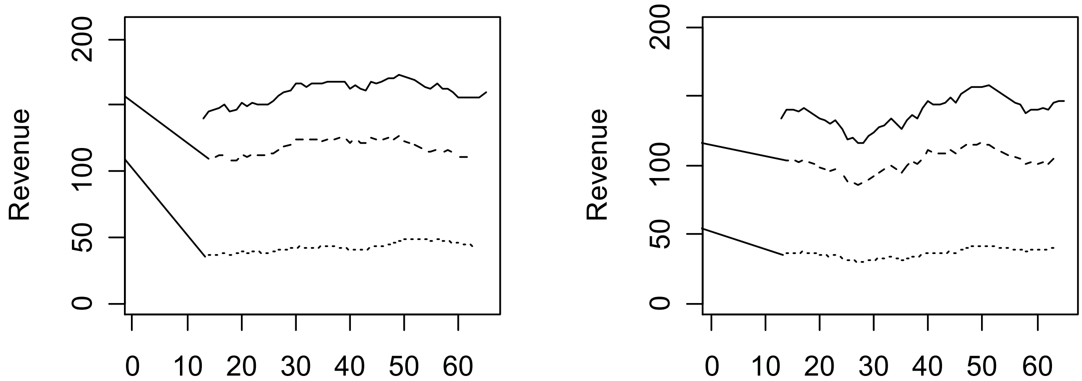


Figure 39 Flight revenue totals (solid) and excluding fuel and security surcharges (interrupted), as well as the total amount of fuel and security surcharges (dotted) for PO605/6 (left) and PO605/7 (right). The values shown are 12-week averages, divided by 1000. The value at the horizontal axis is the number of weeks from January 1, 2004.

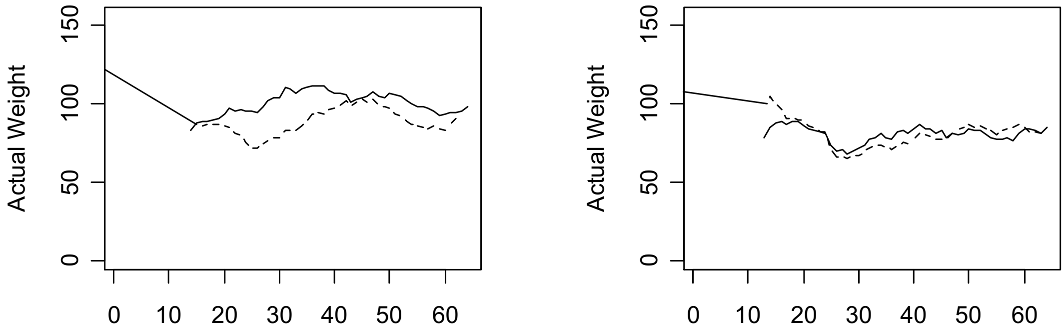


Figure 40 Total actual weight (solid) and volume (interrupted) per flight for PO605/6 (left) and PO605/7 (right). The values shown are 12-week averages. Weights are divided by 1000, volumes by 6000. The value at the horizontal axis is the number of weeks from January 1, 2004.

4.6.6. Comparing PO605/6 with PO605/7

Figure 41 shows empirical distribution functions of the most important properties of space requests, to compare those moved on PO605/6 with those moved on PO605/7. The graphics show that for smaller shipments, below the 80% quantile, the distributions are very much the same, but for larger requests, PO605/6 outperforms PO605/7. Note that this is on a per-shipment level. Large requests are consistently larger for PO605/6 than for PO605/7 and consequently generate more revenue. The difference are standing out the most if only dense cargo shipments are compared.

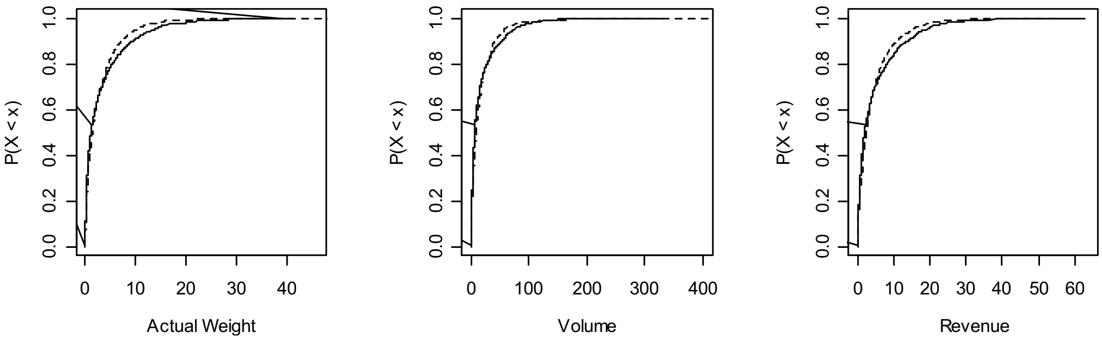


Figure 41 Plots of empirical distribution functions for Actual Weight (left), Volume (middle) and Revenue (right) of all data available for PO605/6 (solid) and PO605/7 (interrupted). The values shown are divided by 1000.

5. Conclusions and Recommendations

5.1. Conclusions

During the internship, a computer system was created that is capable of the following.

- To estimate parameters for the demand model using historical records of cargo airlifted.
- To simulate an arbitrary number of demand sequences, forecasting total demand.
- To compute an optimal Capacity Access Price for any simulated sequence.
- To compute optimal bid-prices using the algorithm Lenstra et al (1982) have provided.
- To perform the actual admission control with both policies, forecasting flight performance.
- To produce single flight performances as if a flight departed with simulated cargo aboard.

Three admission control policies are implemented and flight performance under any of these admission control policies can be forecasted simultaneously.

- Admission control using threshold rates per unit of weight and volume called bid-prices.
- Admission control using the industry standard rate per kilogramme chargeable weight called the Capacity Access Price, add fuel and security surcharges, with a fixed pivot ratio of 6.
- Admission control based on available capacity only, called first-come first-serve.

The software system created is capable of configuring these policies and can be used to assess the flight performances under these policies, as to aid management in selecting the one producing optimal flight performance on a flight-by-flight basis.

For starters, the software system created is capable of supporting in setting an optimal Capacity Access Price for the daily update, which is founded on an accurate forecast. I am sure this alone is a significant improvement to day-to-day flight management already. However, the configuration and enforcement of any of the admission control policies is significantly different from the current day-to-day practice at Polar. The revenue management solution proposed centralizes the admission control. To achieve a similar situation, the Network Revenue & Capacity team is to be the single admission control authority, aided by preferably real-time, on-line booking control software. Moreover, the automated admission control algorithms used to produce flight performances is applying the rules dictated by the admission control policies highly consistently. This currently has its advantages as well as disadvantages. In an ideal situation, the automated admission control algorithm is enhanced and extended, but consequently enforced quite strictly for optimum performance.

The investigation shows a systematic, quantitative approach to revenue management can be expected to make a considerable contribution to operating profits during day-to-day operation at Polar Air Cargo.

The investigation shows applying a bid-price admission control policy may improve flight revenue for PO605/6 in excess of the revenue achievable with admission control based on a Capacity Access Price. Centralizing the admission control and strictly enforcing the rules of an admission control policy based on a Capacity Access Price already has significant positive effects on achievable total flight revenue.

Across the board, an admission control policy based on a bid-price pair has the effect more dense cargo is accepted than under an admission control policy based on a Capacity Access Price. This is the more beneficial the higher the demand is. If the demand is moderate, the fuel and security cost of carrying the extra weight accepted might not exceed the extra revenue generated. Regardless, however, both admission control policies have the effect revenue is improved and total payload airlifted is reduced significantly compared to historical closeout performances. Thus, flight revenue is improved and fuel, security and handling cost are reduced under either admission control policy if they are put to use as proposed in this document.

The way the optimal bid-prices or Capacity Access Prices are computed from a series of simulated arrival streams is not adequate under low-demand situations. Therefore, the bid-prices or Capacity Access Prices computed for PO605/7 are often too high. This has the consequence a first-come first-serve policy produces better flight performances during part of the year than the other policies for this flight.

The software system implements the goals of this investigation.

- It is capable of forecasting total demand, expected flight performance and payload, the demand development over time throughout the pre-flight booking period and can be used to construct confidence ranges for either flight revenue, payload weight, payload volume or the number of requests accepted.
- Three admission control policies are implemented and can be configured, compared and subsequently selected to produce flight performance forecasts.
- Admission control is centralized. Budgets and allocations are ignored by the admission control solution, as they should be.

5.2. Polar Air Cargo Procedures

The total payload capacity available is segmented into allocations and the free space, each with a sales station being responsible for filling that part of the capacity while meeting the budget. A sales station might choose to accept cargo that generates revenue below the opportunity cost under the admission control policy that has been set, in order to meet its budget. This is suboptimal. Take-off payload revenue is maximized by enforcing quite strictly the preset admission control policy that has been shown to maximize expected payload revenue at departure, regardless of whether individual sales stations meet their budget.

Possible solutions for the shared responsibility and contribution rating of sales stations might be to include the number of requests in the budget accepted by admission control and originating from a region. This may be suitable for shipments above a certain threshold weight, only. Another possibility is to depart from per-flight budgets, but to use budgets for a longer period, or at least incorporate a sales station performance rating based on sustained performance as part of a budgeting solution. The natural distribution of request sizes will prove such budgets fair in the long run. An accurate forecasting practice, combined with timely adjustments to the admission control policy, based on realized performance, must then ensure acceptable flight performance on a per-flight basis. Recording what shipments originating from a sales station have been rejected by admission control might be useful to aid performance rating in hindsight.

It is common in the air cargo industry to charge based on a weight called the chargeable weight rather than the actual weight of a shipment. The pivot ratio used to compute chargeable weight based on a shipments weight and volume is considered fixed. However, an admission control policy based on a Capacity Access Price is outperformed by a static bid-price policy. Regardless, a fixed pivot ratio hinders determining an optimal Capacity Access Price.

For communication with customers, the concept of chargeable weight and pivot ratio might prove to be difficult to leave behind. A pivot ratio different from six, or, in general, different from commonly used ratios within the air cargo industry might be difficult to defend. Nevertheless, internally, this is a step to be taken. Communication with customers can be based on chargeable weight and a pivot ratio of six, but this will require the sales force to convert rates and revenue for the customer and will often require negotiating chargeable weights different from the chargeable weight based on the pivot rule with a pivot ratio of six. However, computing the opportunity cost under a bid-price policy is very easy. The method is the same for dense and volume cargo and does already include fuel and security surcharges.

5.3. Data Collection and Data Quality

It should be without question that accurate information about remaining payload capacity on a future flight is very important. Accurate weight, volume and revenue totals have to be fed into the forecasting process as a first step to produce an accurate forecast. This means these totals will have to be up to date at least whenever a forecast is produced or bid-prices are determined.

A good first step is to assure these totals are accurate at the start of each booking period or working day, under the assumption static bid-prices or a fixed Capacity Access Price is used during a whole working day. The ideal situation is that admission control can be done on-line using real-time information from across Europe, enabling determination of optimal bid-prices at any point in time. A centralized, automated solution for this has the added benefit of enabling the recording of request properties regardless of whether a request is accepted.

Recorded volumes are often not accurate, especially for small shipments. Note that most shipments are quite small. One of every four shipments recorded from January 1, 2004 are smaller than 357 kg. In addition, just as many shipments take up less than 2,076 dm^3. Because the number of small shipments is quite large, it is the more important to record reasonable volumes even for very small shipments. If measuring volumes accurately cannot be done cost-effectively, it will without doubt prove to be a big leap forward if awareness was raised amongst the sales force and personnel responsible for data entry, the verification and closeout procedures.

The volume that is recorded in the database along with the moved records is best taken to be that part of the total airway bill volume that actually was on the flight instead of repeatedly recording total airway bill volume. This volume then matches the actual and chargeable weight recorded. It might be useful if these records can be used to prove whether all parts of an airway bill have been moved.

The above is applicable to the manually recorded moved records resulting from the closeout procedure. A wealth of potentially valuable information is also present within the company data centre. However, this information has proven to be very difficult to unlock, due to technical as well as conceptual problems. In general, appropriate means of normalizing and standardizing the flight records may prove to be invaluable. It should be noted Polar was in the process still of the merger with Atlas Air and improvements where under construction in many areas at the time of this writing. In addition, the information technology infrastructure is not solely targeted at scheduled service operations or even the sale of air cargo capacity. The latter can be brought up in defence as well as to identify an opportunity for improvement.

The information technology infrastructure might be developed from a service-oriented point of view, with small, purpose build parts of it providing services to users and other applications in an open manner, improving the exchange of information, rather than aiming to support the building of a single, closed system by an external party. The system supporting the closeout procedure is a good example of such a service, although it is not yet without its own problems. It does provide very accurate records of cargo that moved, with the exception of volume, at the cost of many working hours of manual labour, for a big part due to a challenging user interface.

5.4. Improving Forecasting Quality and Flight Performance

A procedure dubbed the verification and closeout delivers all information that is input into the demand-forecasting model. These procedures focus solely on cargo that has actually been flown. There is no good indication of the number or quantity of cargo requests that have been turned down by sales personnel due to capacity constraints or failed price negotiations.

These price negotiations are mostly done via e-mail, at least as far as the Revenue and Capacity team is concerned. As noted above, segmenting the total payload capacity is suboptimal. A solution has to be formulated to cope with the problem of shared responsibility towards total flight revenue in order to maximize the revenue potential of the existing flights. Having a centralized admission control authority available is highly recommended. The e-mail volume the Revenue and Capacity team already has to cope with is significant. To facilitate a central authority, space requests will have to be formalized and presented to this authority in a way suitable for automated admission control. To which extend the actual admission control will be automated is initially second to aiming to reduce manual processing. An applicable transitional solution might be to put to use a text extraction solution capable of classifying e-mails and recognising for example the dimensions, weight and requested revenues of the parts of cargo forming a single airway bill.

The above, again, would enable the recording of properties of space requests that will never move. Recording reasonably accurate figures of lost sales and cancelled bookings will greatly support forecasting quality. Arrival stream correction using the first-come first-serve solution discussed in this document can only be a temporary solution.

Most information on cost of carrying cargo is unavailable. Ideally, yield management is preferred over revenue management, because revenue management is only an indirect solution, whereas yield management would directly support the company mission. A lot of work in that direction has however already been done and is continuing. It will be important to connect those solutions with the revenue management solution in the future.

Because Polar operates exclusively cargo only aircraft, whereas the bulk of the available cargo capacity on the lanes Polar flies on is so-called belly capacity on passenger aircraft, Polar is in a position to accept cargo that is larger and less common with regard to shipment requirements than most other carriers can accept. This clearly is an advantage. However, a static bid-price policy is only mathematically spoken optimal, optimal under any circumstance and without exception, when arrival rates of space requests would be very high and maximum payload capacity very large. In other words, the smaller individual shipments are, the more the optimality of a static bid-price policy will show in practice.

Marketing Polar as the partner of choice for uncommon cargo has adverse effects, both on total flight revenue as well as handling cost. If market demand were ignored and using theoretical arguments only, it would be better to position Polar as the partner of choice for single letters and postcards. This will hold regardless of whether a static bid-price policy is used or a policy based on a Capacity Access Price. Please note that revenue management is only applicable when different people are willing to pay a different price for a unit of capacity, because only then it matters significantly to discriminate between requests. The foregoing only illustrates Polar is better off carrying a larger quantity of requests rather than carrying larger requests.

Although theoretically a static bid-price policy is optimal, practically it will show to pay off to re-establish bid-prices regularly during the pre-flight booking periods. This is again due to the conditions not being met under which this policy is optimal. With increasing arrival rates and decreasing shipment sizes, the amount of cargo accepted during a period of fixed length, for example a single working day, will converge towards and settle at a fixed average. The spread of the amount of cargo accepted in practice is reason to incorporate realized flight performance into a renewed forecast. The amount of cargo accepted and the revenue already generated is valuable a-priori knowledge or advancing insight, which has to be acted upon regularly. The closer departure comes, the more accurate the forecast will prove to be. This is not an issue if, say, 34,300 shipments arrive during a single week averaging 3,056 grams actual weight.

5.5. Opportunities for Forecasting Model Enhancements

The forecasting model simulates complete requests exclusively. Part-shipments will not be simulated. Although records can be fed into the estimation process that represent part-shipments rather than complete shipments, this has adverse effects on forecasting accuracy. In fact, it is likely that simulation using those distributions generate even larger shipments. The empirical distribution of weights, for example, matches the Gamma distribution very well after parameters are estimated. Including multiple part-shipments instead of a single complete shipment has the consequence that this similarity is partially lost. Especially when Maximum Likelihood Estimation is used, for density and revenue, the estimated fit will still be a Gamma distribution, but with increased probability of observing values in the range corresponding with sizes of part-shipments. These are often still on the large side. With Maximum Likelihood Estimation, the distribution is optimized to fit the bulk of the observations. The tail of the empirical distribution will increase to 1 faster than a normal Gamma distribution. The estimated distribution will not reflect this and large shipments will be simulated with importantly higher probability.

The automated admission control policy being used for forecasting flight performance is straightforward and useful, but can be improved upon in many ways. This also affects the algorithm for computing optimal bid-prices and Capacity Access Prices.

- It ignores the possibility of part-shipping requests.
- It ignores cargo priorities.
- It is only useful with a single flight.
- Hence, it cannot cope with the more complex, multi-lane cargo routing issues.

A request that is subjected to admission control and that does not fit anymore is rejected. Such a request will never be included in any revenue forecast. This is clearly not desired. It will be possible to investigate behaviour with regard to part-shipments quantitatively, in line with the solution direction of the forecasting model as it is presented in this document. Likely, there is a positive probability that a large request cannot be part-shipped, although that would be possible. If the request can be part-shipped, the sizes of the parts are hardly ever arbitrary. A quantitative approach for simulating what parts a shipment may consist of may well prove to be very useful. The admission control policy will have to select the shipment parts moving on the current flight. This is already useful if the parts that do not move on this flight are subsequently discarded.

Assuming a request has normal priority, admission control may decide to move all or part of a shipment on another flight visiting the desired destination directly or indirectly. To this end, it may be useful to acquire and record information indicating what service level is required for each shipment. This information can then later be analyzed and acted on. Many shipments perhaps can be moved without problem on a next flight, being for example PO605/7 instead of PO605/6.

Please note the admission control policy for static bid-prices is capable of dealing with multiple flights with relative ease. How the algorithm of Lenstra et al (1982) for determining optimal bid-prices can be changed is hinted towards in the chapter 2.2. Admission Control. It involves providing the algorithm with vector valued weights, volumes, revenues and capacities rather than scalars. Each coordinate of such a vector concerns a single flight. The slope of the line corresponding with the relative height of the bid-prices is determined just the same way as it is for a single flight. The height of the line, however, is now such that the shipments accepted do not exceed the available capacities of any flight included in the optimization, instead of just the capacity of a single flight.

Implementing the algorithm of Lenstra et al (1982) for multiple flights, and changing the admission control procedures to be able to part-ship, off-load and prioritize cargo, will allow forecasting for flights that are not currently part of the core schedule. Alternatively, as a transitional measure, the existing admission control procedure can be configured to indicate for each shipment subjected to admission control whether it has been accepted or rejected. The requests that have been rejected may be merged with the demand simulated for a next flight.

This may even be useful for forecasting for PO605/7 in the current situation, although special care has to be taken not to over-estimate the demand for the Sunday flight. Demand simulated without this for PO605/7 may already include a fair amount of cargo off-loaded from PO605/6. It is much more applicable for getting a feeling of whether scheduling a third weekly PO605 should be supported.

The flight performance forecast produced by this means for an extra section or a new scheduled flight will depend greatly on the bid-prices used for the existing flight. If these are such that the performance of the existing flight is optimized, the other flight will show at best moderate performance, because requests are accepted on the existing flight greedily. Without additional demand specifically for the new flight, performance may be disappointing. However, especially for extra sections, a forecast produced by this means may support the decision process. It may defend or discourage the decision to schedule an extra landing in the region or provide insight in the profit loss or gain of the extra take-off cycle necessary.

Again, the performance of the existing flight depends on the bid-prices set. It may be protected in practice or may be reported based on the relation between forecasts produced for this flight with and without the new flight and realized performance for all flights. If a new flight performs such that the break-even point with regard to the constant cost is not definitely enough met, the risk of scheduling the new flight may exceed the profit opportunity. Please note that such a forecast should be based on complete arrival streams. The recommendations regarding this should be implemented first. Decisions based on a forecasting model with corrected incomplete arrival streams are the more risky in this situation, because the additional requests simulated due to arrival stream correction can be expected to make up the bulk of the cargo for a new flight.

Appendices

A. Important Probability Distributions

The Uniform Distribution

A randomly chosen point X from an interval (a, b) has a uniform distribution on (a, b). A randomly chosen point U from the interval (0, 1) is said to have a standard uniform distribution. X can be obtained from U by taking $X = a + (b - a) \times U$.

A stochastic variable X is said to have a uniform distribution if its density function is a constant function on its support. Consequently, its distribution function is linear and increasing on its support.

If X has a uniform distribution on (a, b), the following statements hold.

- The density function f of X is $f(x) = 1 / (b - a)$, which equals 1 if $a = 0$ and $b = 1$.
- The distribution function F of X is $P(X < x) = F(x) = (x - a) / (b - a)$, which equals x if $a = 0$ and $b = 1$.
- The expectation of X is $E(X) = (a + b) / 2$, which is 1 / 2 if $a = 0$ and $b = 1$.
- The variance of X is Var(X) = (b − a)□ / 12, which is 1 / 12 if $a = 0$ and $b = 1$.
- R functions `punif`, `qunif`, `dunif` give its cumulative distribution function, quantile function and density function. `runif` can be used to draw random samples.
- Drawings from the standard uniform distribution are obtained with `RAND()` in Excel.

The Exponential Distribution

The exponential distribution is often interpreted as the lifetime of a part experiencing no wear-off, like a light bulb. The time between renewals of a counting process with stationary and independent increments can be proven to have an exponential distribution.

A stochastic variable X is said to have an exponential distribution with rate λ if its density function equals f below. Moreover, if X is distributed with an exponential distribution with rate λ, the following statements hold.

- The density function f of X is $f(x) = \lambda \times e^{-\lambda x}$ for $x > 0$.
- The distribution function F of X is $F(x) = 1 - e^{-\lambda x}$ for $x > 0$.
- Hence, the tail probability of X is $P(X > x) = e^{-\lambda x}$.
- The expectation of X is $E(X) = 1 / \lambda$.
- The variance of X is Var(X) = 1 / λ□.
- X is memory free since $P(X > s + t \mid X > s) = P(X > t)$. A non-negative random variable Y is memory free if and only if Y has the exponential distribution.
- If U is a random variable with a standard uniform distribution, then the random variable $Z = -log(U)$ has a standard exponential distribution.
- R functions `pexp`, `qexp`, `dexp` give its cumulative distribution function, quantile function and density function. `rexp` can be used to draw random samples.
- Excel function `EXPONDIST(x;lambda;cumulative)` returns the exponential distribution.

The Gamma Function

The (complete) gamma function $\Gamma(x)$ is defined to be an extension of the factorial to complex and real number arguments.

It is related to the factorial by $\Gamma(x) = (x - 1)!$ if x is a positive integer. Recall that $n! = n \times (n - 1) \times (n - 2) \times \ldots \times 2 \times 1$ and that 0! is defined to be 1. Hence, for non-negative integer n, the factorial of n is $n! = n \times (n - 1)!$.

If x is not a positive integer, $\Gamma(x)$ is the integral of $t^{x-1} \times e^{-t}$ for $t = 0$ to positive infinity. For non-positive integers including 0, this integral is not finite. Several fast approximation algorithms exist. See for example Mathworld (http://mathworld.wolfram.com/GammaFunction.html).

R function `gamma(x)`.

The Gamma or Erlang Distribution

The exponential distributions with arbitrary rates can be regarded as members of the family of Gamma distributions. A random variable X has a Gamma distribution, also dubbed Erlang distribution, with shape parameter r and rate λ if its density function equals f below. If r is a positive integer, X can be seen as the sum of r random variables having the exponential distribution with rate λ.

- The Gamma density function f is $f(x) = \lambda r / \Gamma(r) \times x^{r-1} \times e^{-\lambda x}$ for x > 0, if r is a positive real number and $\lambda > 0$.
- The Gamma density function f is $f(x) = \lambda r / (r - 1)! \times x^{r-1} \times e^{-\lambda x}$ for x > 0, if r is a positive integer and $\lambda > 0$.
- The Gamma density function f is $f(x) = \lambda \times e^{-\lambda x}$ for $x > 0$, if $r = 1$. This exactly is the density function of the exponential distribution.
- The Gamma distribution function F cannot be given in closed form.
- The parameter r can be seen as a shape parameter.
- The parameter λ is the rate and its inverse, $1 / \lambda$ can be seen as a scale parameter.
- The expectation of X is $E(X) = r / \lambda$ or, alternatively, *shape* × *scale*.
- The variance of X is $Var(X) = r / \lambda$□ or, alternatively, *shape* × *scale*□.
- Shape and scale of a Gamma distribution fit can be obtained from the sample mean (*mean*) and sample variance (*var*) of a sample *s* by taking *scale = var(s)* / *mean(s)* and subsequently taking *shape = mean(s)* / *scale*.
- R functions `pgamma`, `qgamma`, `dgamma` give its cumulative distribution function, quantile function and density function. `rgamma` can be used to draw random samples.
- Excel function `GAMMADIST(x;alpha=shape;beta=scale;cumulative)` returns the gamma distribution, `GAMMAINV(probability;alpha=shape;beta=scale)` returns the quantile function.

The Poisson Distribution

The number of observations of an event that occurs at random during a time interval normalized to be of length 1 is Poisson distributed with a rate matching the average number of observed events per unit of time.

If S is the sum of a series of k independent random variables, each with an exponential distribution with rate λ, the value of k for which S does not exceed 1 is Poisson distributed with rate λ. λ is the number of renewals of the Poisson process with rate λ per unit of time. The number of renewals during the interval from 0 up to t of a Poisson process with rate λ is Poisson distributed with parameter λt.

If X is a random variable with a Poisson distribution with rate λ, the following statements hold.

- The probability density function of X is $P(X = k) = \lambda^k / k! \times e^{-\lambda}$ for $k \geq 0$ and $\lambda > 0$.
- The probability distribution function of X is $F(n)$ is the sum from $k = 0$ to n of $P(X = k)$.
- The expectation of X is $E(X) = \lambda$.
- The variance of X is $Var(X) = \lambda$.
- Suppose the number of events of certain kind is Poisson distributed with rate λ and each such an event is observed with probability p, with p in (0, 1), independent of other events or of the total number of events. The number of observed events is again Poisson distributed with rate λp.
- Due to the above, in addition, if X is Poisson distributed with rate λ_1 and Y is Poisson distributed with rate λ_2, then $Z = X + Y$ is also Poisson distributed with rate $\lambda = \lambda_1 + \lambda_2$.
- Let X_n be a series of random variables with a binomial distribution with probability p_n and n trials. X_n is to be interpreted as the number of successes achieved with n Bernoulli trials with success chance p_n. Suppose $n \times p_n \to \lambda$ when $n \to \infty$ for certain $\lambda > 0$. Then, $P(X_n = k) \to P(X = k)$ if $n \to \infty$ with X a Poisson distributed random variable with rate λ.
- Hence, if Y has a Binomial(n, p) distribution with n large and p small, then Y can be approximated by a Poisson($\lambda=np$) distributed random variable. This proves the binomial process is a suitable discrete analogon of the Poisson process. If a continuous time interval is

split into sufficient subintervals of equal size, such that the probability of observing more than 1 arrival during one such subinterval is sufficiently small, the Poisson process and the Binomial process are practically the same.

- R functions `ppois`, `qpois`, `dpois` give its cumulative distribution function, quantile function and density function. `rpois` can be used to draw random samples.
- Excel function `POISSON(x;mean;cumulative)` returns the Poisson distribution.

B. General Forecasting Techniques

Goodness of fit

The statistical model is an expression of our knowledge of the experiment that led to recording the observations. The model often suggests that the observations are generated from a distribution of some general form. In other words, a family of similar distributions, often parameterized, is assumed to fit the observations. However, this assumption must be validated. This may be done before or after parameters are estimated. To this end, various techniques can be used which test the goodness of fit (*toetsen voor aanpassing*) of the model with the observations.

> **Example** If a random variable defined for an experiment may produce any positive real number, there is a good chance a distribution from the exponential family of distributions may fit the data well. However, because there are many distributions that fit this description, a series of tests are to be conducted to select a particular family of distributions that may be parameterized and estimated more easily than a general form exponential distribution.

Univariate Samples

Suppose a series of numbers are the results of a repeated experiment. From the method by which the experiments are executed, that is, without memory of previous experiments, it may be concluded it is reasonable the numbers are realizations of independent, identically distributed random variables. The numbers together form a univariate sample. The question remains which marginal distribution suits the observations.

Quantile-quantile plots for location-scale families

A very useful technique for testing goodness of fit, which is very useful at early stages, too, is to graphically compare quantiles of the distribution the random variable is assumed to have, called the a-priori distribution, with quantiles of the empirical distribution or the distribution estimated, called the a-posterior distribution. The graphic produced is called a quantile-quantile plot or QQ-plot. Usually, a sequence of probabilities is taken matching the number of realizations and quantiles are computed for both the distributions at these probabilities.

If the QQ-plot shows a straight line, the distributions compared are from the same location-scale family. Distribution functions in a location-scale family are the same save for a horizontal shift or a horizontal scale.

> **Example** Normal distributions are all member of the same location-scale-family regardless of their actual mean or deviation.

Every well-known distribution family like the normal, Cauchy, Gamma and exponential families have their own location-scale family.

Comparing histograms with the probability density function

Another, modestly useful graphical technique to test goodness of fit is to compare the density function of the distribution with a histogram of the observations in the sample. A histogram visualizes the number of realizations falling within a certain range of values.

It is more useful to use the derivative instead of the distribution function itself for this technique, because the shape of its derivative is often much more visually outspoken.

For this technique to work, the bars of the histogram will have to be scaled to match the surface of the density function. This can be done by dividing the bar heights by the total number of samples, because the surface under the density function is always 1 due to the fact that it produces the distribution function. In addition, the bars of the histogram will need to be taken neither too narrow nor too wide. Several techniques exist for selecting the bar width, but experimenting will be needed.

This technique is only useful when a relatively large number of observations are available, that is, more than a hundred or better yet, five hundred.

The Wilcoxon Goodness of Fit Test

The Wilcoxon goodness of fit test can be used to test whether two samples are identically distributed. It is a very useful test, because it does not have any requirements towards what distribution family actually fits the samples. The test is based on the notion that, for two samples to be identically distributed, their observations must be similarly big. It is very easily applied.

Suppose the two samples are X with n observations and Y with m observations. For each of the n observations of X, the number of observations that is smaller is determined. For example if X_1 is the smallest except for 2 observations from Y, the rank number of X_1 is 3. The Wilcoxon test statistic W is the sum of these rank numbers. Big values of this sum W indicate the observations in X are relatively large compared to those in Y.

The null hypothesis that X and Y are identically distributed is rejected for big as well as small values of W, although it is possible to formulate a one-sided null hypothesis. If the hypothesis is rejected, it can be concluded X and Y are differently distributed most definitely. If the hypothesis is not rejected, it can be said it is reasonable to assume X and Y are identically distributed.

The distribution of the rank numbers computed does not depend on the distribution of X or Y, as long as both X and Y have a continuous distribution. It is possible to determine this distribution analytically, or simulation can be used to build the empirical distribution.

> **Example** Suppose two samples, X with n observations and Y with m observations, are tested for whether they are identically distributed. A series of random drawings from the normal distribution with certain fixed mean and standard deviation can be obtained quickly with a variety of tools. Using pairs of such drawings, one with n samples, one with m samples, a series of Wilcoxon test statistics can be computed. These test statistics are distributed the same as the Wilcoxon test statistic for X and Y. An α-confidence range for the Wilcoxon test statistic is obtained by taking the $\alpha/2$ and $1-\alpha/2$ quantiles of this series of test statistics. The null hypothesis that X and Y are identically distributed is rejected at confidence α if the Wilcoxon test statistic for X and Y lies outside this confidence range.

The Kolmogorov-Smirnoff Goodness of Fit Test

The Kolmogorov-Smirnoff goodness of fit test is based on the fact the empirical distribution of a sample and the real distribution must be very similar.

According to the law of large numbers, the empirical distribution function of a sample converges point-wise towards the probability distribution function when the number of observations goes to infinity. This corresponds with the definition of limiting probabilities. It means that, if the number of observations is not too small, the empirical distribution must lie close to the real probability distribution underlying the observations. The Kolmogorov-Smirnoff goodness of fit test is based on the maximum vertical distance between the empirical distribution and the real distribution.

Suppose a sample X, with n observations, is tested for whether it is distributed according to a continuous distribution function F. The empirical distribution function of X, F_n, must resemble F. The Kolmogorov-Smirnoff test statistic T is the maximum vertical distance from F_n to F. The null hypothesis is rejected for large values of T. T is distributed the same, regardless of F or F_n. A confidence range is available in tables or can be simulated.

In practice, however, the above approach is too trivial. This is because F is often not fully known, but rather depends on parameters that have to be estimated from the observations. For example, F is assumed a distribution from the normal distribution family, but its mean and standard deviation have yet to be determined. The Kolmogorov-Smirnoff goodness of fit test may be performed exactly as laid out above in this situation, but the test statistic T no longer distributed independent of F or F_n. Its distribution will have to be simulated in order to obtain a suitable confidence region.

> **Example** Suppose a sample X, with n observations, is assumed distributed according to a Gamma distribution with unknown positive shape h and scale c. The shape and scale are estimated from the observations. The question is whether the estimated Gamma distribution fits the distribution of the observations.
> A series of random drawings of size n may be simulated using the estimated shape and scale. For each such sample, the shape and scale are again estimated. For each random drawing, the test statistic described above is computed using these shapes and scales. This series of test statistics together form the empirical distribution of the test statistic T for X.

The null hypothesis that X is distributed according to a Gamma distribution with shape h and scale c is rejected at confidence a if the Kolmogorov-Smirnoff test statistic T for X is larger than the $(1 - a)$-quantile of the empirical distribution simulated. Please note the confidence region in this case is one-sided, because the null hypothesis is rejected only for large values of T.

The Kolmogorov-Smirnoff goodness of fit test can only be applied if the distribution of the observations is continuous. However, it might prove to be useful for discrete distributions with a small adjustment. Recall that a random variable with a discrete distribution only takes on a limited number of values. The distribution function of a random variable with a discrete distribution is a stair step function, as is the empirical distribution. This will have the consequence that small differences in the distribution functions immediately produce large values for the Kolmogorov-Smirnoff test statistics, being the maximum vertical distance.

The Kolmogorov-Smirnoff goodness of fit test may however be applied for discrete distribution functions that are linearly interpolated. That is, the distribution function and the empirical distribution are first modified, such that they become part wise continuous functions. For values the underlying distribution does not take on, the distribution function returns a linear interpolation between the largest value that is smaller and the smallest value that is larger than the value being evaluated. The Kolmogorov-Smirnoff test statistic is subsequently taken to be the maximum vertical distance between these continuous functions.

Example The unmodified Kolmogorov-Smirnoff goodness of fit test cannot be used to determine whether a Poisson distribution with certain mean is a suitable fit for the number of space request arrivals during a certain day. This number of arrivals is a non-negative integer and hence, both the empirical distribution function as well as the fitted distribution function only takes on integer values. The Poisson distribution fitted may be modified by taking $P(X \leq k + x) = P(X \leq k) + x \times P(X = k + 1)$ for x in $(0, 1)$ and k any non-negative integer. The empirical distribution function is modified similarly. These modified distribution functions can subsequently used to determine the Kolmogorov-Smirnoff test statistic.

C. Forecasting Techniques for Dependent Data

Testing for Consistency

In many cases, observations are not just numbers, but vectors of related information. It is often important to determine whether different coordinates of such a vector are statistically related, next to being logically related. A statistical relationship might be useful to explain a certain coordinate based on the value of the remaining coordinates of the vector of related information.

Suppose a vector consists of coordinates x and y, that is, each observation can be written as (x, y). A plot of these coordinates on the 2-dimensional plane is called a *scatter plot*. If there is an outspoken relationship between these coordinates, this will be immediately visible from the scatter plot.

Such a relationship might not be obvious, however. A numerical expression of the power of the linear relationship is the *sample correlation coefficient*. The value of the correlation coefficient will be between -1 and 1. If it is 1, there is a perfect positive linear relationship. An increasing value of the x-coordinate corresponds with a proportional increase in the y-coordinate. Likewise, a value of -1 proves a perfect negative linear relationship. An increasing value of the x-coordinate now corresponds with a proportional decrease in the y-coordinate. If the x-coordinates and y-coordinates are realizations of independent experiments, that is, if the x and y-coordinates are unrelated, the sample correlation coefficient will be 0. This statement cannot be inverted, however. A correlation coefficient of 0 does not prove the coordinates are unrelated. It does only prove the relationship is not linear.

Using Maximum Likelihood Estimation to Model Space Request Density

The general idea is to choose two pairs of powers and multipliers as parameters. For each shipment, a shape and scale is computed using the actual weight and the parameters chosen. Next, for each shipment, the density function of the Gamma distribution with the shape and scale computed is evaluated at the cargo density of the shipment.
These values are the likelihood of observing the density of the shipment given its actual weight. The sum of these values is the likelihood of observing all the samples assuming the densities are distributed according to a Gamma distribution with the shape and scale computed with the current parameter estimates.
The parameter values that maximize the likelihood of observing the actual weight and density pairs in the dataset provide a maximum likelihood estimate of the conditional density distribution.

> **Note** Although likelihood can be seen as a measure of probability, likelihood values are not probabilities. This does not matter, however, because only the parameter estimates obtained using maximization are of interest, not the size of the maximum.

Often, the logarithm of the likelihood is computed and summed to produce the maximum log likelihood, because small changes in likelihood values between zero and one are stretched over all of the negative y-axis by the logarithm function, providing expressive input to the optimization algorithm used. This is a valid method because the logarithm function is strictly increasing, so it only maps the likelihood values onto a different position on their axis. Again, the actual likelihood value is of no interest.

> **Note** The optimization algorithm is generic and will search for improvements of likelihood at the edges of the bounding box defined by the upper and lower limits of the parameters. Evaluation of any of the functions at the edges of the bounding box must return strictly positive values, because a negative shape or scale parameter is not allowed for the Gamma distribution. The upper and lower bounds for a single parameter must not be equal or the bounding box will have zero size and optimization will not take place. You can however specify values that are practically the same to have the algorithm effectively ignore a parameter, keeping it at its starting value.

Using Bins to Analyze Data Based on Similarity of the Factors

Dividing the graphic in slices can be done by creating so called bins, with each bin containing only that part of the data having factors that are close to each other. Those bins may be partitioning the data, such that each sample is in only one bin, but the bins may also overlap, such that each sample is in a number of consecutive bins.

Fitting the distribution of the response can then be done for each bin separately, independent of the other bins. Simulating responses is then done by simulating the factors first, selecting the most appropriate bin based on the factors simulated and using the distribution estimate of that bin for simulating the response.

D. Additional Graphics and Tables

Conditional Revenue Distribution

		Dense Cargo		Volume Cargo		Cargo at Pivot	
Factor	**Parameter**	**Shape**	**Scale**	**Shape**	**Scale**	**Shape**	**Scale**
Actual Weight	**Shift**	2.077	2.077	2.314	2.314	26.895	26.895
Actual Weight	**Exponent**	0.123	0.692	0.163	0.641	0.404	0.432
Density	**Shift**	0.868	0.868	2.602	2.602	0	0
Density	**Exponent**	0.414	-0.404	-0.225	0.710	0	0
(combined)	**Multiplier**	1.333	3.132	2.509	1.169	0.468	7.533

Table 14 Maximum likelihood estimates of the function parameters for the revenue distribution of each density classification.

Arrival Stream Correction Using First-Come First-Serve

The columns MA and FCFS have been removed from this document before publication.

	PO605/6				PO605/7			
Week	**MA**	**%**	**FCFS**	**Delta**	**MA**	**%**	**FCFS**	**Delta**
14		20		-470		15		108
15		20		-671		15		411
16		15		-1,854		15		-1,859
17		30		-524		15		1,173
18		20		-1,161		10		-36
19		15		581		10		-2,446
20		10		-256		10		-72
21		10		-2,058		10		-2,618
22		10		-1,295		15		148
23		15		206		10		580
24		10		-1,010		10		-245
25		10		-3,870		0		-4,987
26		10		-626		0		4,568
27		15		769		0		-1,776
28		10		-845		0		2,488
29		20		605		0		835
30		20		-338		0		4,398
31		20		602		0		-3,141
32		35		448		0		-649
33		25		650		0		-1,674
34		25		-839		0		2,430
35		30		322		0		2,747

	PO605/6				PO605/7			
Week	MA	%	FCFS	Delta	MA	%	FCFS	Delta
36		25		43		0		-406
37		40		199		0		-4,634
38		40		-278		0		-1,332
39		50		-6,625		0		-1,633
40		60		-5,423		0		-5,585
41		60		-458		10		2,383
42		60		-2,690		10		9,178
43		45		197		10		7,599
44		50		-410		10		2,394
45		50		-1,797		0		-3,435
46		60		1,590		10		-120
47		60		-5,033		10		-234
48		60		-1,034		0		-2,014
49		60		-3,080		15		1,627
50		60		281		15		848
51		60		-2,290		20		-8
52		50		-40		10		-2,165
53		60		796		15		585
54		45		-418		25		176
55		40		1,247		15		629
56		35		726		15		3,252
57		25		-616		10		4,751
58		40		336		10		-1,195
59		35		-604		10		2,169
60		25		-800		10		-487
61		25		410		0		-1,237
62		15		135		0		-3,467
63		25		1,559		0		-1,049
64		10		-1,919		0		24
65		25		496		0		663
66		25		127		0		521

Table 15 The percentages used to increase arrival rates while performing arrival stream correction for PO605/6 and PO605/7 using first-come first-serve. The weekly arrival rates are increased such that a weeks flight performance under first-come first-serve matches the performance estimate obtained by computing 12-week moving averages closest.

Space Request Arrival Rates – PO605/6

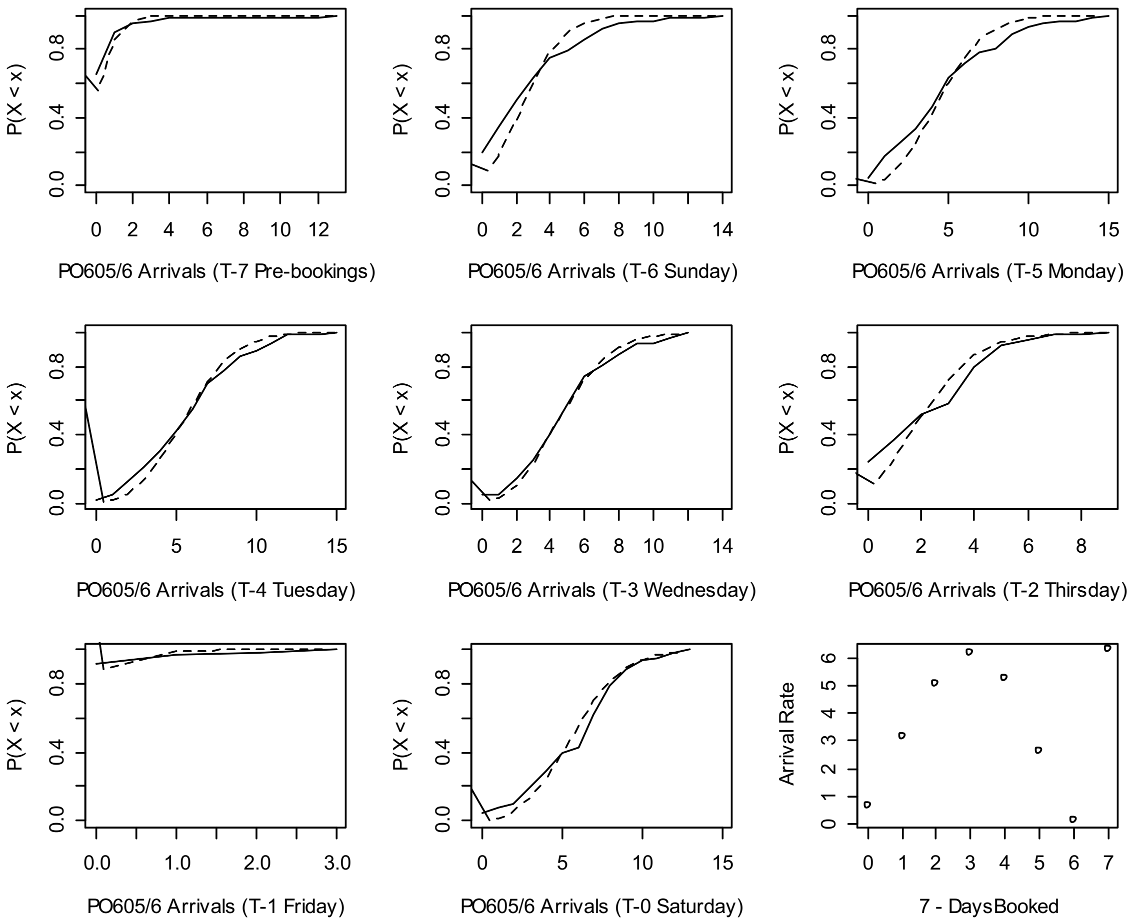


Figure 42 Daily arrival rates for PO605/6. There is one plot for each booking period and a plot summarizing the expected arrival rates. The expected arrival rates are plotted for booking periods in the order of booking period label.

Day of Week	**Pre**	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**	**Sun**	**Total**
Booking Period	**T-7**	**T-5**	**T-4**	**T-3**	**T-2**	**T-1**	**T-0**	**T-6**	**(all)**
Sample Size	63	63	63	63	63	63	63	63	504
Mean	6.30	2.65	5.25	6.21	5.10	3.16	0.67	0.13	3.68
Variance	8.99	4.84	7.39	9.84	1.86	8.81	3.13	0.24	11.90
Standard Deviation	3.00	2.20	2.72	3.14	1.36	2.97	1.77	0.49	3.45
Minimum	0	0	0	0	0	0	0	0	0
25% Quantile	4	1	3,5	4	2,5	1	0	0	0
Median	7	2	5	6	5	2	0	0	3
75% Quantile	8	4	6,5	8	7	4,5	1	0	6
Maximum	13	9	12	15	15	14	13	3	15

Table 16 Numerical summaries of the daily arrival rates for PO605/6. The column Pre concerns pre-bookings.

Space Request Arrival Rates – PO605/7

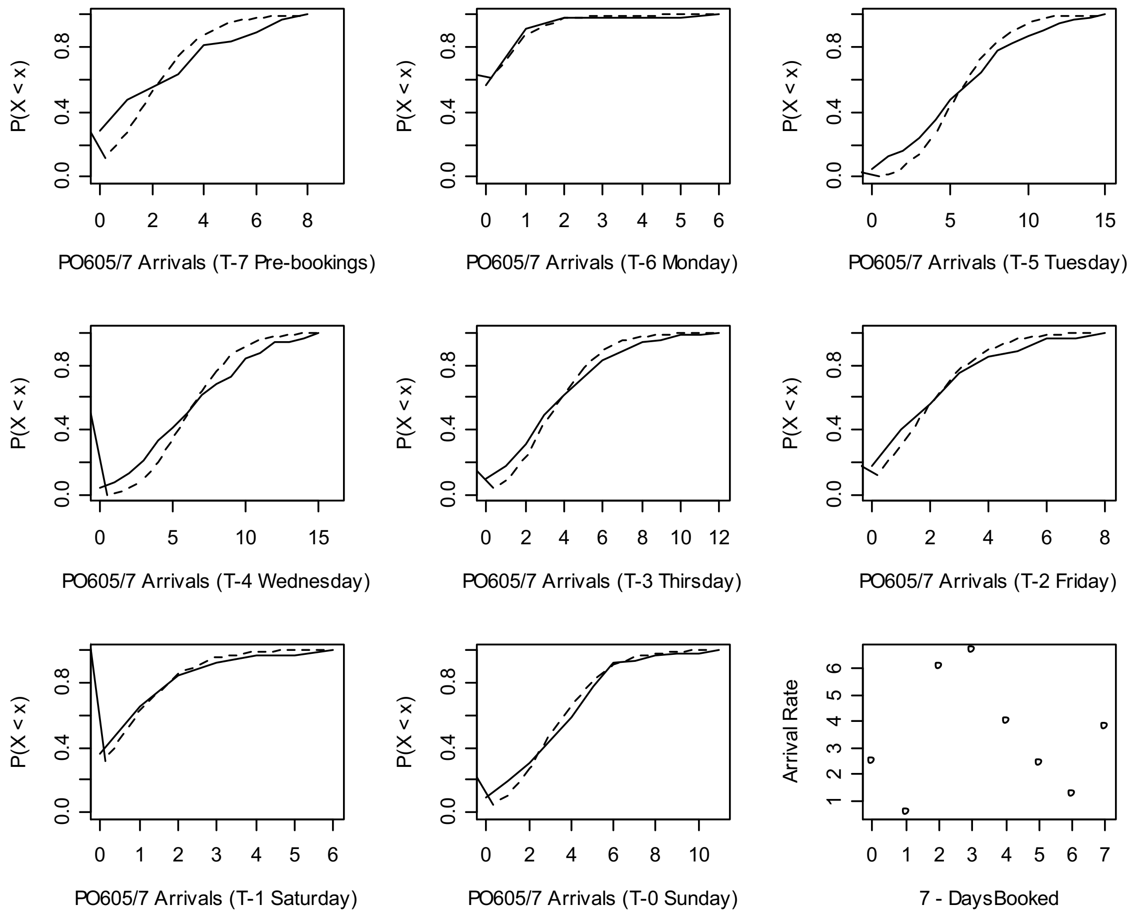


Figure 43 Daily arrival rates for PO605/7. There is one plot for each booking period and a plot summarizing the expected arrival rates. The The expected arrival rates are plotted for booking periods in the order of booking period label.

Day of Week	**Pre**	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**	**Sun**	**Total**
Booking Period	**T-7**	**T-6**	**T-5**	**T-4**	**T-3**	**T-2**	**T-1**	**T-0**	**(all)**
Sample Size	63	63	63	63	63	63	63	63	504
Mean	3.81	1.29	2.44	4.00	6.70	6.10	0.57	2.54	3.43
Variance	5.61	2.05	4.12	7.48	4.57	3.06	0.86	6.09	10.74
Standard Deviation	2.37	1.43	2.03	2.74	2.14	1.75	0.93	2.47	3.28
Minimum	0	0	0	0	0	0	0	0	0
25% Quantile	2	0	1	2	4	4	0	0	1
Median	4	1	2	4	6	6	0	2	3
75% Quantile	5	2	3,5	6	0	8	1	4	5
Maximum	11	6	8	12	15	15	6	8	15

Table 17 Numerical summaries of the daily arrival rates for PO605/7. The column Pre concerns pre-bookings.

Admission Control Policy Performance with Corrected Arrival Streams – PO605/6

Historical flight performance of PO605/6 was € 156,032 over 66 flights from January 10, 2004, with on average 103,270 kg payload weight and 567,507 dm^3 payload volume from 30.79 requests.

Simulated flight performance of PO605/6 was € 2,067 higher on an average, with 6,105 kg and 55,869 dm^3 less payload.

Over 52 flights, this equals € 107,484 more revenue, 317.5 tonnes and 2,905 m^3 less payload with 155 more accepted requests.

Performance Measure	All Arrivals	Accepted			
All Simulated Flights	100%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	211,356	157,140	156,413	158,099	1,686
Payload Weight	138,982	102,061	94,581	97,165	2,584
Payload Volume	758,051	550,506	516,539	511,638	-4,901
Requests	40.22	34.11	33.48	33.78	0.30
Unconstrained Flights	28.1%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	139,584	139,623	129,655	131,515	1,861
Payload Weight	88,738	88,751	78,906	80,791	1,886
Payload Volume	474,443	474,247	427,201	427,789	588
Requests	34.33	34.32	32.16	32.59	0.44
Constrained Flights	71.9%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	239,503	163,985	166,907	168,524	1,617
Payload Weight	158,686	107,261	100,728	103,586	2,858
Payload Volume	869,275	580,301	551,575	544,521	-7,054
Requests	42.53	34.02	34.00	34.24	0.24

Table 18 Summary of the effects of different admission control policies on the flight performances of 53 simulated PO605/6 flights, based on arrival streams corrected conservatively using first-come first-serve.

Admission Control Policy Performance with Corrected Arrival Streams – PO605/7

Historical flight performance of PO605/7 was € 125,910 over 66 flights from January 11, 2004, with on average 80,609 kg payload weight and 496,116 dm^3 payload volume from 28.03 requests.

Simulated flight performance of PO605/7 was € 1,761 higher on an average, with on average 702 kg and 30,764 dm^3 less payload.

Over 52 flights, this equals € 91,572 more revenue, 36.5 tonnes and 1.600 m^3 less payload with 26 more accepted requests.

Performance Measure	**All Arrivals**	**Accepted**			
All Simulated Flights	100%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	139,666	130,305	127,339	127,671	332
Payload Weight	89,161	82,689	79,539	79,907	368
Payload Volume	527,236	490,623	465,069	465,352	283
Requests	30.21	29.26	28.49	28.53	0.04
Unconstrained Flights	71.9%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	119,419	120,333	116,029	116,397	368
Payload Weight	75,422	75,746	72,034	72,411	377
Payload Volume	442,987	447,402	419,363	419,933	570
Requests	28.67	28.87	27.89	27.95	0.05
Constrained Flights	28.1%	**FCFS**	**CAP**	**BP**	**Difference**
Flight Revenue	193,323	155,803	157,312	157,547	235
Payload Weight	125,572	100,441	99,429	99,772	343
Payload Volume	750,510	601,139	586,197	585,718	-479
Requests	34.30	30.26	30.08	30.07	-0.01

Table 19 Summary of the effects of different admission control policies on the flight performances of 53 simulated PO605/7 flights, based on arrival streams corrected conservatively using first-come first-serve.

Indexes

Literature Index

Harn, K. van and P.J. Holewijn (1997), *Inleiding in de Waarschijnlijkheidsrekening*, Faculteit der Wiskunde en Informatica, Vrije Universiteit, Amsterdam

Hendricks, Gregory L. and Terry Elliott, Implementing Revenue Management Techniques in an Air Cargo Environment
http://www.unisys.com/transportation/insights/white__papers/papers.htm?insightsID=23715

Kan, Rinnooy A.H.G., L. Stougie and C. Vercellis (1993) A Class of Generalized Greedy Algorithms for the Multi-Knapsack Problem, *Discrete Applied Mathematics*, 42, 147-167

Koole, Ger (January 2004) Optimization of Business Processes: Applications and Theory of Mathematical Modelling, *http://www.math.vu.nl/~koole/obp/*

Lenstra, A.K., J.K. Lenstra, A.H.G. Rinnooy Kan, and T.J. Wansbeek (1982) Two Lines Least Squares, *Annals of Discrete Mathematics*, 16, 201-211

Oosterhoff, J. and A.W. van der Vaart (January 2003) *Algemene Statistiek*, Faculteit der Exacte Wetenschappen, Vrije Universiteit, Amsterdam

Pak, Kevin and Rommert Dekker (June 2004) Bid-prices for a 0-1 Multi Knapsack Problem, *ERIM Report Series Research in Management*, ERS-2004-055-LIS

Ross, Sheldon M. (1972) Introduction to Probability Models, Academic Press, *ISBN 0-12-598475-8*

Talluri, K.T. and G.J. van Ryzin (1998) An Analysis of Bid-Price Controls for Network Revenue Management, *Management Science*, 44, 1577-1593

Venrooy, L.P. van (January 2002) Yield Management: Toepassing door luchtvaartmaatschappijen, *Business Mathematics and Informatics paper*, Vrije Universiteit, Amsterdam

Tables

Figures