INTEGRATING THE FLEET ASSIGNMENT MODEL WITH UNCERTAIN DEMAND

by

Anwar Hood Ahmed

Submitted in partial fulfillment of the
requirements for the degree of
Master of Philosophy

at

Brunel University
Uxbridge, London
January 2010

BRUNEL UNIVERSITY

MATHEMATICAL SCIENCES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "INTEGRATING THE FLEET ASSIGNMENT MODEL WITH UNCERTAIN DEMAND" by Anwar Hood Ahmed in partial fulfillment of the requirements for the degree of Master of Philosophy.

Dated: January 21st, 2010

External Examiner: ______________________________
Dr. Elia El-Darzi

Research Supervisor: ______________________________
Dr. Cormac Lucas

Examining Committee: ______________________________
Prof. John Beasley

____
Dr. Keming Yu

____

BRUNEL UNIVERSITY

DATE: January 21st, 2010

AUTHOR: Anwar Hood Ahmed

TITLE: INTEGRATING THE FLEET ASSIGNMENT MODEL WITH UNCERTAIN DEMAND

DEPARTMENT OR SCHOOL: Mathematical Sciences

DEGREE: MPhil. CONVOCATION: July YEAR: 2010

To Nasra, Barka, Hood, Hafswa and Muhammad.

Table of Contents

Table of Contents . . . . **v**

List of Tables . . . . **ix**

List of Figures . . . . **x**

Acknowledgements . . . . **xii**

Abstract . . . . **xiii**

Chapter 1 Introduction **1**

1.1 OR Application in the Airline Industry 1

1.2 OR Application within Revenue Management 3

1.3 OR Application within Network Planning 6

1.3.1 Fleet Planning 6

1.3.2 Route Planning 7

1.4 OR Application within Schedule Planning 9

1.4.1 Schedule Generation 9

1.4.2 Fleet Assignment 10

1.4.3 Through and Maintenance Routing 13

1.4.4 Crew Scheduling 14

1.5 Conclusion 16

Chapter 2 Airline Schedule Robustness **17**

2.1 Defining Robustness 17

2.2 Measuring Robustness 18

2.3 Modelling Robustness Techniques 19

2.4 Heuristic Robustness Techniques 21

2.5 Robustness within FAM 25

2.5.1 Demand Robustness 25
2.5.2 Operational robustness 26
2.6 Conclusion 26

Chapter 3 Airline Fleet Assignment Modelling 28
3.1 FAM Development 28
3.2 Basic FAM Formulation 29
3.3 Solution Method 33
3.4 FAM Extensions 36
3.5 FAM with Aircraft Utilization 38
3.6 Cost and Revenue Data 43
3.7 Conclusion 45

Chapter 4 Stochastic Programming (SP) Modelling Approaches . 46
4.1 Background 46
4.2 The Recourse Problems 46
4.3 Distribution Problems 50
4.4 Chance-Constrained Problems 52
4.5 The Inter-relationships and Bounds 52
4.6 Solution Approaches 53
4.6.1 Universe 53
4.6.2 Decomposition 54
4.6.3 Statistically Based Method 55
4.6.4 Decomposition Methods with Sampling 56
4.7 Examples of SP Solvers 57
4.8 Conclusion 58

Chapter 5 The SP-FAM 59
5.1 Problem Description 59
5.2 Previous Attempts 59
5.3 Mathematical Formulation 60

5.4 Model Features 64
5.4.1 The Standby Aircraft 64
5.4.2 The Objective Function 65
5.4.3 The Artificial Variable 66
5.4.4 Demand-Spill Constraint 66
5.4.5 Aircraft Utilization Constraint 70
5.4.6 Aircraft Utilization Cost 72
5.4.7 The Standby Aircraft Penalty Cost 72
5.5 Conclusion 73

Chapter 6 Scenario Generator 74
6.1 Integrating Scenario Generator and SP Model 74
6.1.1 Scenario Trees 75
6.1.2 Scenario Generation 76
6.1.3 Measure of Quality of a Scenario Tree 77
6.2 Scenario Generation for FAM 77
6.3 Calibration Process and Network-Simulation Model 79
6.4 The Scenario Generator 81
6.4.1 Holt-Winter's Additive Time Series Model 82
6.4.2 Scenario Generator Assumptions 86
6.5 Conclusion 88

Chapter 7 SP-FAM Solution Process 90

Chapter 8 Case Study 96
8.1 SP-FAM 96
8.1.1 SP-FAM Statistics 98
8.1.2 SP-FAM Stochastic Measures 98
8.2 Stability Measure 100
8.2.1 In-Sample Stability Measure 101
8.2.2 Out-of-Sample Stability Measure 102

8.3 Conclusion 104

Chapter 9 Conclusion and Future Work **105**

9.1 Conclusion 105

9.2 Future Work 107

9.2.1 Solution Algorithm 107

9.2.2 Revenue Effects 107

9.2.3 Improving Demand Forecast 108

9.2.4 System Integration 109

Bibliography **110**

List of Tables

Table 2.1 Typical KPIs. 20

Table 5.1 Passenger demand scenarios. 67

Table 5.2 Demand-spill constraint. 68

Table 5.3 Constraint 5.3.6 and function C_{U1}. 71

Table 5.4 Constraint 5.3.12 and function C_{U2}. 71

Table 6.1 Techniques used in scenario generation 78

Table 6.2 Marketing Information Data Tape 83

Table 6.3 Computation of seasonal index 83

Table 6.4 Seasonal index 84

Table 6.5 Identifying the trend component. 85

Table 8.1 Here-and-now model statistics 98

Table 8.2 Z_{TP}: Objective values and SP measures 98

Table 8.3 Z_{RP}: Objective values and SP measures 99

Table 8.4 Z_{AP}: Objective values and SP measures 99

Table 8.5 Z_{VP}: Objective values and SP measures 100

Table 8.6 In-sample stability. 101

List of Figures

Figure 1.1 An overview of planning in the airline industry. 2
Figure 1.2 The revenue management process. 3
Figure 1.3 Aircraft selection criteria. 6
Figure 1.4 Network evolution. 8
Figure 1.5 The schedule planning process. 9
Figure 1.6 Schedule construction steps. 10
Figure 1.7 Feasible connections in one station. 13

Figure 3.1 A two-type, two station fleet-flow time-space network. 31
Figure 3.2 A one-type, one station fleet-flow time-space network. 34
Figure 3.3 Network reduction for a time-space network. 35
Figure 3.4 The combined fleet assignment model. 38
Figure 3.5 The relationship between price and demand. 43

Figure 6.1 Multi-stage tree. 76
Figure 6.2 Scenario tree for two-stage SP. 76
Figure 6.3 Calibration process and network-simulation model. 80
Figure 6.4 TSM flow chart. 82
Figure 6.5 Overview of the solution approach. 86

Figure 7.1 SP-FAM solution approach. 90
Figure 7.2 Before and after schedule aggregation. 91
Figure 7.3 Before and after defining fleet homogeneity. 92
Figure 7.4 Aggregated swap enumeration. 92
Figure 7.5 AMPL Shell. 93

Figure 8.1 In-sample stability. 102

Figure 8.2 Out-of-sample stability. 103

Figure 9.1 O-D pair. 109

Acknowledgements

I would like to thank Dr. Chandra Poojari, my first supervisor, for his many insights during the first year of research that set the precedence for the remaining years. I would also like to express my gratitude to Dr. Cormac Lucas, for agreeing to continue supervision of the work after Dr. Chandra Poojari left Brunel University.

I am grateful to my two brothers and their families, Abdulrahman and Khalil, who supported me financially especially during the first year of my study. They gave me love and morale to pursue my goal non-interrupted. I would also like to thank my uncle Idha and his family, who gave me a home during my stay in London. My mother has always remained a pillar of my success.

Finally, I would like to thank my family, whom without their support this research would never have been possible. I sincerely thank my neglected wife, Nasra, who was always ready to stand by me, I sincerely love you for this. My other children, Barka and Hood, I owe a great compensation for the time that I was meant to be with them.

Abstract

One of the main challenges facing the airline industry is planning under uncertainty, especially in the context of schedule disruptions. The robust models and solution algorithms that have been proposed and developed to handle the uncertain parameters will be discussed. Fleet assignment models (FAM) are used by many airlines to assign aircraft to flights in a schedule to maximize profit. In the context of FAM, the goal of robustness is to produce solutions that perform well relative to uncertainties in demand and operation.

In this thesis, we introduce new FAMs (i.e. DFAM1 and DFAM2) that tackles the common problem associated with aircraft utilization. Subsequently, stochastic programming (SP) is presented as a method of choice for the research. Through the use of a *two-stage SP with recourse* technique, the DFAMs are extended to SP-FAMs (SP-FAM1 and SP-FAM2). The main distinction of the SP-FAM compared with other FAMs is that, given a stochastic passenger demand, it gives a strategic fleet assignment solution that hedges against all possible tactical solutions. In addition, we have a tactical solution for every scenario. In generating the demand scenarios, we use a network-simulation model embedded with a *time-series* engine that gives a snapshot of one week that is representative of any other week of the scheduling season.

We later outline the approach of solving the SP-FAMs where the schedule is compacted through several preprocessing steps before inputting it into SAS-AMPL converter. The SAS-AMPL converter prepares all the data into readable AMPL format. Finally, we execute the optimizer using a FortMP solver (integrated in AMPL) that invokes branch-and-bound algorithm. We give a proof of concept using real data from a Middle East airline. Our investigations establish clear benefits of the recourse FAM compared to alternative models. Finally, we propose areas of future research to improve SP-FAM robustness through solution algorithms, revenue management (RM) effects, calibration of network-simulation models and system integration.

Chapter 1

Introduction

"Management scientists (Operation researchers) have wrecked the airline industry" (Coates, 2005).

In this chapter, an overview of *Operations Research* (OR) in the airline industry and major areas of its application will be discussed. Specifically, the illustrations of the applications will be made with a simple OR model as used in *Revenue Management* (RM), *Network Planning* (NP) and *Schedule Planning* (SP).

1.1 OR Application in the Airline Industry

The field of *Operations Research* (OR) and specifically optimization, has had a profound impact on the airline industry (Yu and Yang, 1998). Although airline optimization models have been used since the 1950s, airline planning and operation has become increasingly complex and dynamic (Klabjan, 2005). The need for effective and proper planning has increased the use of optimization models in the airline industry (Sylla, 2000). This need has also come about due to increasing pressure on profitability in a fiercely competitive environment.

The unprecedented revolution in airline operation optimization has been contributed to: (i) the evolution of technology and rapid advancement of optimization, (ii) increasing focus on the *bottom-line* in the industry, and (iii) better understanding (by researchers) of the issues in the planning of airline operations (Yu, 1998). In order to optimize profit, an airline must continually evaluate performance of existing routes and pro actively undertake appropriate remedial action in terms of capacity rationalization, route optimization, network optimization, hub optimization and code-share optimization. Today major airlines have fully fledged OR divisions that cater across the company on optimization related issues. Similarly, small airlines buy optimization applications *off-the-shelf* and employ OR personnel in vital areas that have significant elements of optimization. The major contributors in the application of airline optimization is depicted in Figure 1.1 that reflects the core processes typical for many airlines.

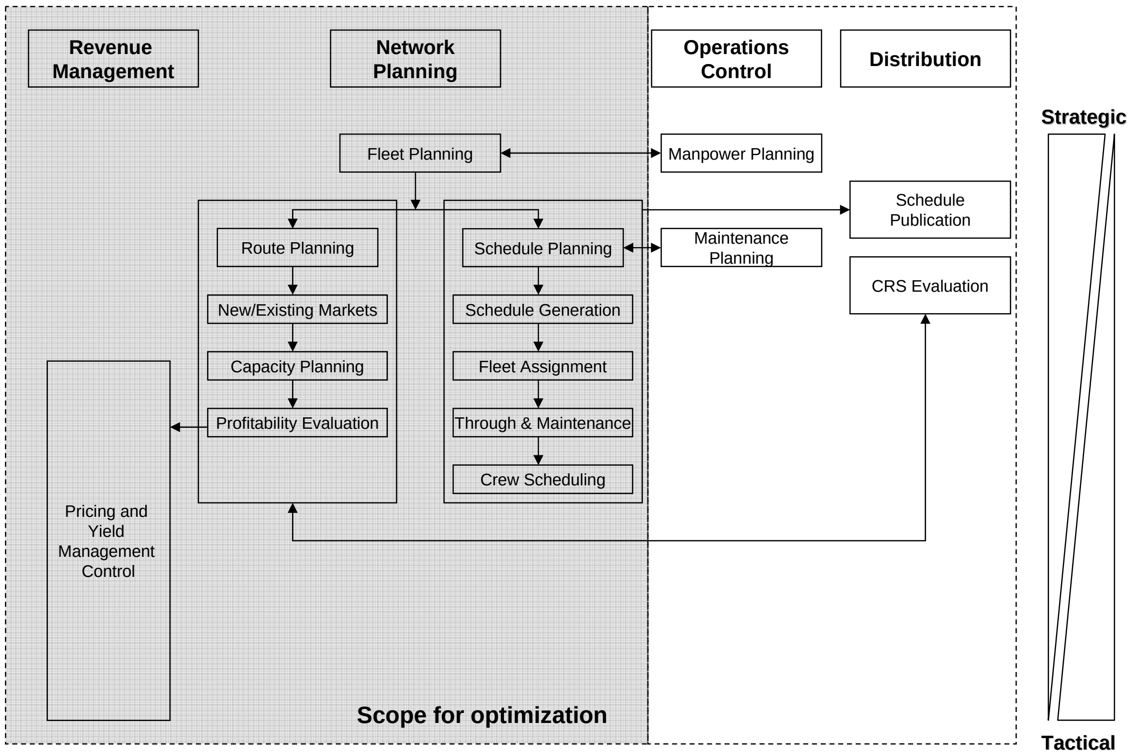


Figure 1.1: An overview of planning in the airline industry.

The airline planning process is inter related and can broadly be classified into three phases; *strategic planning*, *tactical planning* and *operational planning*. Some of the phases have an element of both operational and tactical or tactical and strategic approaches. Strategic planning is long-term planning and attempts to position an organization within its environment and typically spans a period of over five years. On the other hand, tactical planning is a series of specific actions necessary to support the accomplishment of the strategic, overall plan and covers a period of 1 - 5 years. Operational planning is setting out clearly the implementation of the strategic plan against specific objectives and covers a period of six months to one year.

1.2 OR Application within Revenue Management

Airlines sell seats in the same plane compartment with different restrictions on purchase and price. Pricing of seats in different categories has an important bearing on revenues (fill all seats and do not turn away high-fare customers (Busutulli, 1999)). RM is the integrated control of capacity and price whose sole objective is to maximize revenue; it means selling the right seat to the right customer at the right price and the right time (Klophaus and Polt, 2007).

There is a common belief that RM originated in airline industry with the first true model accredited to Littlewood. The basic model has been through numerous refinements; SP formulation, nonlinear programming formulations, dynamic programming formulations, along with various algorithms for solving these models, including decomposition methods and some quite elaborate schema (Boyd, 2006). An extensive review on the evolution and current research on RM has been elaborated (McGill and Ryzin, 1999). The biggest refinement was the introduction of origin and destination control, or simply O-D control. O-D control optimizes over an entire flight network as opposed to individual flight legs. The current setting of many RM departments within airlines is a de-linking of the pricing role from the RM role; however, the future, especially in the context of O-D control, is the dynamic integration of the two (pricing and RM) through the entire booking period. The new paradigm will determine what is the most appropriate fare to display to the consumer prior to the departure to ensure that the highest probability of sale is attained. The paradigm differs from the current setting where the RM department alone determines the allocation of seats based on the forecast passenger demand for predefined products (Dunleavy and Westermann, 2005).

As an illustration, the RM process normally follows a five-stage process as shown in Figure 1.2.

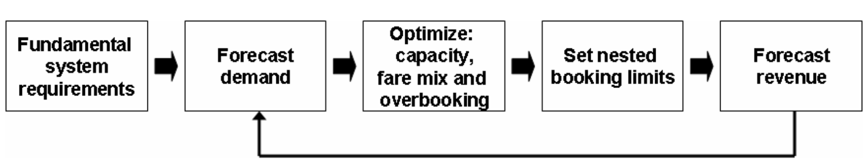


Figure 1.2: The revenue management process.

Fundamental system requirements: This entails establishing fare rules to gain the best price from different customers. Such a pricing strategy will depend on several factors such as demand elasticity, competitive issues, marketing factors and cost factors.

Forecast demand: As demand varies frequently, capacity needs to be rationalized to accommodate the variations. When demand is high, select the high-yield traffic and spill the rest to the low-demand flights. Similarly, when demand is low, attracting additional traffic either by recapturing own spill, capturing demand from competitors by intensifying sales and marketing efforts.

Optimize capacity, fare mix, overbooking: When there is excess demand in a lower compartment while empty seats (that are not expected to sell) are available in a higher compartment, the capacity in the higher compartment is rationalized to accommodate the excess demand in the lower compartment. In *fare-mix* optimization the bookings limits are calculated using a popular heuristic approach known as the *Expected Marginal Seat Revenue*. This technique takes three different forecast values by booking class; mean demand, demand variability, and expected revenue or fare to optimize to give the best fare mix. *Overbooking* is done deliberately as some already booked passengers do not show up at the time of departure. The no-show could be as a result of missed connections, double booking or fake bookings that have not been cancelled. An overbooking optimization takes the no-show forecast and overbooks to offset the no-shows.

Set nested booking limits: Seat availabilities in the *computer reservation systems* can be displayed in several ways. For example, if a booking class has reached the limit of its own allocation, the booking class would have sold out; this is known as *partitioned control* or *discrete nesting*. The main disadvantage of this approach is a situation where a high booking class is sold out while a lower booking class is still available. *Serial* (or linear) nesting overcomes such situations by creating a linear order of the booking classes such that the lower booking class will still be available once the high booking class is sold out.

Forecast Revenue: Forecasting revenue provides the airline with an early warning of future revenue shortfall. Specifically, the source of revenue weakness pertaining to origin or *point of sale* and the day of the week is identified. Once identified, the impact of competitors fares are evaluated and strategies to combat the situation are devised. At this point, the airline can match, cut down or maintain the fares.

A simple RM model based on marginal values or *bid-price* (Higle and Sen (2001)) is associated with seat availability and the willingness of the customer to accept or reject the offered price. Suppose customers request flight services for a particular *itinerary* comprising flight legs with different class (based on fare) composition.

Sets and indices

i=1..I denotes itineraries (i $\in$ I).

l=1..L denotes flight legs (l $\in$ L).

Decision variables

x_i = number of seats allocated to itinerary i.

Parameters

$a_{i,l}$ = 1 if itinerary i uses flight leg l, otherwise it is 0;

v_i = unit revenue for itinerary i;

c_l = number of seats available on leg l; and

$\bar{d}_i$ = expected passenger demand for itinerary i.

Constraints and objective function

$$P_{DLP} \qquad Max \sum_{i \in I} v_i x_i \tag{1.2.1}$$

$$subject \;\; to:$$

$$\sum_{i \in I} a_{il} x_i \leq c_l \qquad \forall l \in L \tag{1.2.2}$$

$$0 \leq x_i \leq \bar{d}_i \qquad \forall i \in I \tag{1.2.3}$$

The *objective* function (1.2.1) maximizes the revenue for each itinerary; the *allocation* constraint (1.2.2) ensures that the number of seats to be allocated is less or equal to the available seats (capacity) of the itinerary and the *demand* constraint

(1.2.3) ensures the number of seats to be allocated in the itinerary is less or equal to the passenger demand for the itinerary.

1.3 OR Application within Network Planning

1.3.1 Fleet Planning

Fleet planning encompasses acquiring the right type of aircraft with the right number of given seats for the airline (Stone, 1998). In a nutshell, fleet planning defines an airline's structural build-up and is characterized by aircraft type, number of aircraft, deployment of aircraft, attainment of financial goals, timing of purchase and retirement of aircraft. Acquiring the right number of seats, entails analyzing the route network for existing and new destinations that is driven by demand forecast. The simulated route network is embedded in a profitability model that is optimized to determine the best mix of aircraft (Boeing, 2003). On the other hand, consideration for buying the right type include, among other factors, operating cost, price, performance, comfort, range, commonality, ease of maintenance, operational flexibility, stability of value, external noise, emissions and design of cargo hold (Clark, 2001). The above fleet selection criteria are outlined and shown in Figure 1.3.

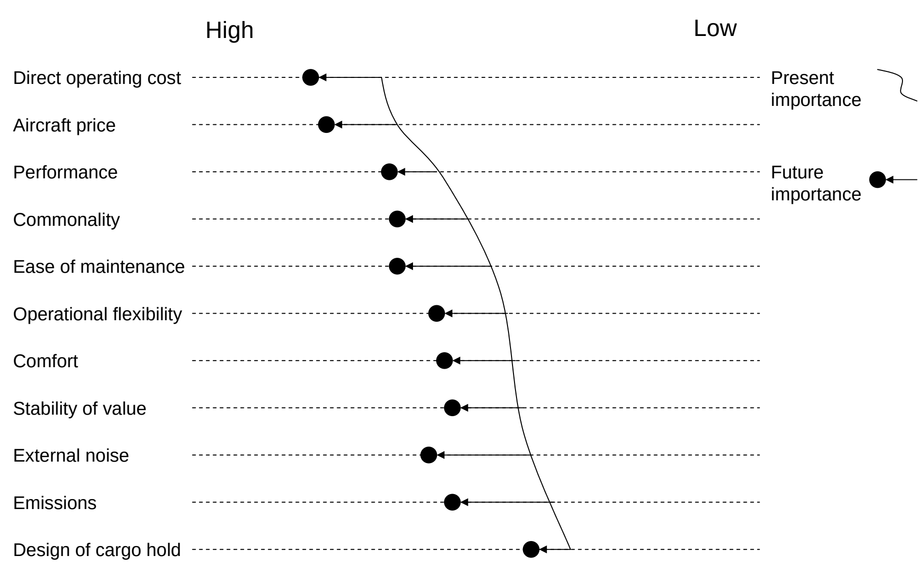


Figure 1.3: Aircraft selection criteria.

Belobaba (2006) outlined two approaches of fleet planning methodologies, the *top-down or macro* approach and the *bottom-up or micro* approach. The macro approach allows the rapid evaluation of new aircraft types, given high level of assumptions, at a defined sub-system, region, or route level. In contrast, the micro approach is the most commonly used approach as it entails detailed evaluation of routes and allows scenario analysis. Mathaisel (2008) confers with Belobaba on the approaches and distinguishes two optimization models for fleet planning problems, the *Cell* (Mathaisel, 1981) model and *FA-4* model. The *Cell* model is an economic-based macro model with financial constraints that reduces the size of the multiple-period problem by clustering individual routes into classes or *cells*, such as short-haul, medium-haul, or long-haul. The model finds the optimum combination of aircraft types and frequencies for all of the *cells* in this aggregate network. The reduced size of the problem permits a rapid evaluation of the fleet requirement for a number of years into the future. On the other hand, *FA-4* is a micro approach and uses the *fleet assignment model* to make optimal decisions on routes, aircraft types and frequencies for a single period in time, over all periods in the planning horizon. The *FA-4* approach assists the airline firm in optimizing: the route network design for the airline; the assignment of aircraft for each route; and the frequency of service on each route. The results also provide valuable insight into the profit impact of a specific route upon a network, and into how the addition of new stations or the deletion of existing ones affect optimal routings. These decisions are made for a single period in time.

1.3.2 Route Planning

Airlines have developed different appraisal matrices that rank routes based on different performance attributes such as, *load-factor*, market-share analysis, network value, average fare paid by passenger per miles flown (*yield*), unit revenue per available seat, unit cost per available seat, etc. The ranking pin-points both star-performers and under-achievers, thus prompting detailed scrutiny. The specific route analysis is then done using different models (e.g., *network-simulation* model) and practical decisions undertaken. Such decisions can lead to: frequency adjustment, schedule synchronization, capacity rationalization and code-share evaluation.

Similarly, new routes have to be planned with the right entry timings, frequencies and appropriate aircraft type to be deployed. Using a multi-criteria technique, Ahmed and Taskila (2007) devised a scoreboard where potential routes are prioritized on a launching strategy. Such a technique entails a process of network analysis through a *network-simulation* model and other micro-economic indicators.

A start-up airline will adopt a point-to-point network structure and evolve into a single hub structure as the market matures. Once there is enough traffic between the two non-hub points, the network evolves further into a hub by-pass and eventually a multi-hub fragmented structure. The network evolution stages, are depicted in Figure 1.4.

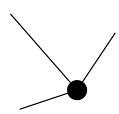
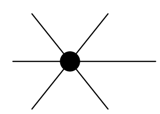
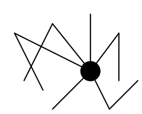
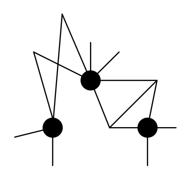
Stage1	**Stage 2**	**Stage 3**	**Stage 4**
A point-to-point structure	Single hub structure	Increased direct non-hub offering (hub by-pass)	Multi-hub, fragmented structures
			
This is a starting point for most airlines	Once a network has been established, connections can be organized	Once there is enough traffic between two non-hub points, direct flights are offered	Connections are organized in other airports

Figure 1.4: Network evolution.

The other enormous task involves the development of scheduling *wave* or *bank* systems to optimize on connectivity, especially for an airline that embraces the hub-and-spoke network. Lufthansa Systems (Jeschke, 2006) have developed a hub optimizer that maximizes unconstrained network revenue by optimizing the arrival and departure times for flight legs that operate via a hub. This is done by shifting services in time, splitting services and swapping services. Here service means a set of flights operating on different days of the week and all the flights fulfil several conditions such as; operated by the same airline, have the same flight number, are connected to the same hub, departs or arrives in the same hub, have the same routing, have the same slot requirements, etc. The hub optimizer takes slots and rotational constraints into account.

The *bank* system comes under criticism, especially for a large network, as it normally leads to over capacity with a high-cost structure that leads to lower aircraft and airport utilization (both employee and gate utilization). The emerging popular concept, *rolling* hub or *breaking* the bank is a system that spreads flights out during the day instead of arranging them in the previously peaked connecting banks. The main distinction of the *rolling* concept is that the *waves* are split and overlapped, so that the inbound *bank* and outbound *bank* occur simultaneously in the same direction.

1.4 OR Application within Schedule Planning

The *schedule planning process* follows a four-stage process as shown in Figure 1.5.

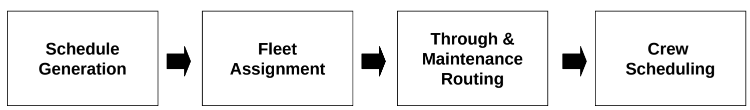


Figure 1.5: The schedule planning process.

1.4.1 Schedule Generation

The four-stage process for schedule generation (Sylla, 2000), comprises demand estimation, schedule construction (refer to Figure 1.6), slot management and schedule distribution. Schedules are developed from a previous schedule by estimating demand for the season under consideration. The estimated demand is optimized taking several parameters such as connecting flights, desired turn-around times, aircraft availability, customer preference and competitors schedule. Slot refers to timing rights given to an airline for landing and take-off; mainly applicable to international markets and other slot-controlled airports; they are normally negotiated twice in a year, during summer and winter. Finally, the schedule is filed into the global distribution channels to ensure that it is correctly displayed.

Erdmann et al. (2001) described the schedule generation problem as the problem of determining aircraft rotations observing operational constraints and fleet sizes, and

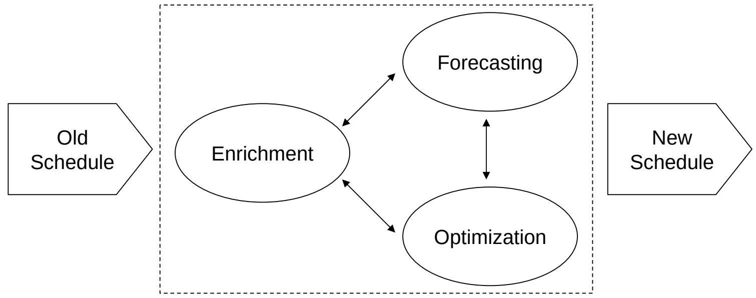


Figure 1.6: Schedule construction steps.

of re-routing passengers taking seat capacities into account, such that the combined aircraft and passenger costs are minimized. He proposed a model for charter operation that generates the schedule from scratch as opposed to the conventional process of generating a new schedule from the previous one.

1.4.2 Fleet Assignment

Once the schedule has been generated, the next stage is to assign a specific type of plane to each leg. Different planes have different operating costs, ranges and capacities. Furthermore, competitiveness also dictates which aircraft can be assigned on a specific leg. The generated schedule, together with the projected demand (potential revenue) are the inputs to the FAM that assigns aircraft to the appropriate legs. The basic FAM can be stated in two ways, using the arcs to represent connection (*connection* networks) or using the arcs to represent flight legs (*time-space* networks). In this section, the connection-based formulation (Abara, 1989) will be described and the time-space formulation will be described in Chapter 3.

The connection network has nodes that represent the points of time when flights arrive or depart. In addition, an imaginary master source node and a master sink node are conceptualized (not actually created) in the network to account for the beginning-of-the-day and the end-of-the-day effects. There are three types of arcs representing the different types of connections: the flight connection arcs link the arrival nodes to the departure nodes, the terminating (connection) arcs link the arrival nodes to the master sink node to represent aircraft arriving and remaining at the station for the rest of the day, and the originating (connection) arcs link the master source node to the departure nodes to represent the aircraft that are present at the station at the beginning of the day. All flight connections have to be feasible with respect to flight arrival and departure times; that is, the minimum turn-time has to be observed between the arrival flight and the following departure flight to allow for the connection.

Sets and indices

L is the set of flight legs indexed by i, l and j; K is fleet type ($k \in K$); S is a set of stations ($s \in S$); A_s and D_s is the sets of arrivals and departure legs for station S, ($s \in S$), respectively. (The indices $i = 0$ and $j = 0$ denote originating and terminating arcs, respectively).

Decision variables

f_{ijk} is a binary decision variable that takes on a value of 1 if the feasible connection between flight leg $i \in L$ to flight leg $j \in L$ is covered by aircraft type k.

Parameters

p_{jk} is the benefit of departing flight j of the connection and is a combination of profit, aircraft utilization, etc; c is the unit operating cost of assigning the type f_{0jk} that initiates the use of a fleet type k for some flight leg j; N_k is the number of aircraft available for type k.

Constraints and objective function

$$Max \sum_{i \in L \cup \{0\}} \sum_{j \in L} \sum_{k \in K} p_{jk} f_{ijk} - c \sum_{j \in L} \sum_{k \in K} f_{0jk} \tag{1.4.1}$$

$$subject \ \ to:$$

$$\sum_{i \in L \cup \{0\}} \sum_{k \in K} f_{ijk} = 1 \quad \forall j \in J \tag{1.4.2}$$

$$\sum_{i \in L \cup \{0\}} f_{ilk} - \sum_{j \in L \cup \{0\}} f_{ljk} = 0 \quad \forall l \in L, k \in K \tag{1.4.3}$$

$$\sum_{i \in D_s} f_{0ik} - \sum_{i \in A_s} f_{i0k} = 0 \quad \forall s \in S, k \in K \tag{1.4.4}$$

$$\sum_{i \in L} f_{0ik} \leq N_k \quad \forall k \in K \tag{1.4.5}$$

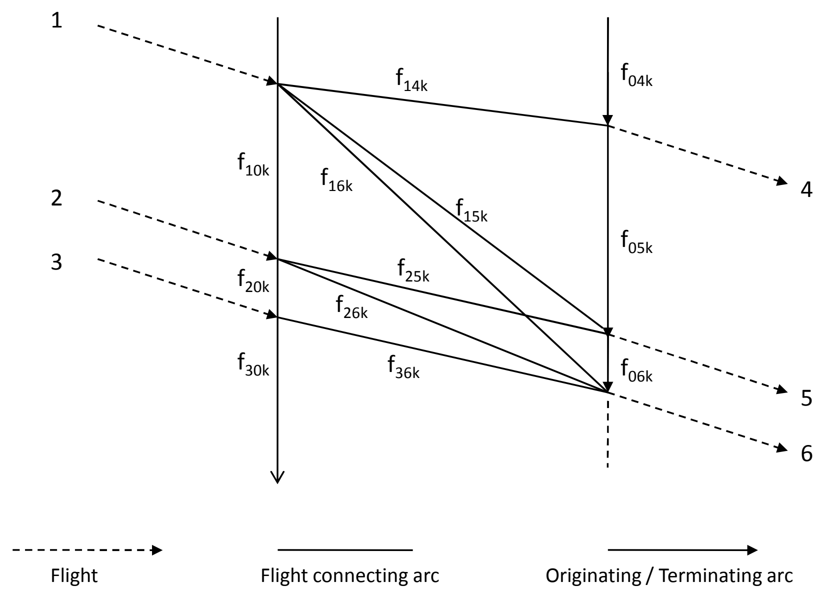
The *objective* (1.4.1) is to maximize profit by subtracting costs ($c \sum_{j \in L} \sum_{k \in K} f_{0jk}$) from revenue ($\sum_{i \in L \cup \{0\}} \sum_{j \in L} \sum_{k \in K} p_{jk} f_{ijk}$). The *cover* constraint (1.4.2) requires that each flight is preceded by an arrival or an originating *arc* that is covered by a fleet type. To ensure the integrity of the network, the *equality* or *continuity of equipment* constraint (1.4.3) ensures that each flight served begin (sequence originating or continued from another flight) and end (sequence termination or turn into another flight) on the same aircraft type. The *schedule balance aircraft* constraint (1.4.4) ensures that the same number of aircraft of each type remain at each station every night so that the same assignment can be repeated daily. In the case of an unbalanced schedule, Abara made a provision to automatically balance the schedule by introducing an origination shortage variable (P_{sk}) and a termination shortage variable (Q_{sk}) for each station s and aircraft type k combination. Using this provision, constraint (1.4.4) can be restated as:

$$\textstyle\sum_{i \in D_s} f_{0ik} + P_{sk} \; = \; \textstyle\sum_{i \in A_s} f_{i0k} + Q_{sk} \qquad \forall s \in S, k \in K$$

The *availability* constraint (1.4.5) limits the number of available aircraft to use in the assignment, that is, the total number of aircrafts assigned to flight legs ($\sum_{i \in L} f_{0ik}$) should not be more than the available of that type (N_k).

Figure 1.7 shows twelve connections: six (feasible) flight connections, three connections from arrival flights to termination, and three connections from originations to departure flights. Since Abara defined *turn* as the successive assignments of an aircraft to two consecutive flights, this become a major limitation as non-connecting flights (i.e. 1-6) are not included in the model. Further, since all feasible connections had to be specified, the model expanded into unmanageable size due to the large number of possible connections. Abara deals with this problem by specifying a limit on the number of connection variables that are considered for each flight.

In the estimation of cost, Abara uses a nominal unit operating cost, denoted by c, for each assignment of type f_{0jk} that initiates the use of fleet type k for some flight leg j. However, the industry standard specifies that the cost factor (c) should account for both the flight leg and fleet type (i.e. c_{ik}). Similarly, revenue is based on a stochastic model of demand, but Abara assumes that each leg is independent (i.e. p_{jk} is assigned arbitrarily to the turns in which the flight is the departing segment).



Arrow arcs 1-3 are arriving flights and 4-6 are departing flights. They are not included in the decision variable. The decision variables are the binary connections variables f_{ijk} for fleet type k to cover connection i to j, where f_{i0k} represent the terminating connections and f_{0jk} represent the originating connections.

Figure 1.7: Feasible connections in one station.

1.4.3 Through and Maintenance Routing

When a plane passes into a hub by combining two flights without a change of aircraft, it is called a *through flight*. Through flights are displayed first in the passenger reservation system, have more revenue and are preferred by passengers to connecting flights. The Through-Assignment Model (TAM) takes input from FAM and determines through connections by identifying inbound and outbound flights at each city flown by the same fleet. Ahuja (2006) stated the following TAM formulation.

Sets and indices

L is a flight leg ($i,j \in L$); K is fleet type ($k \in K$); O is a station ($o \in O$); *IN(o,k)* is the set of inbound flight legs to node *(o,k)* and *OUT(o,k)* is the set of outbound flight legs from node *(o,k)*.

Decision variables

$$x_{i,j} = \begin{cases} 1 & \text{if flight with leg } (i,j) \text{ is selected;} \\ 0 & \text{otherwise.} \end{cases}$$

Parameters

$p_{i,j}$ is the through benefit of assigning flight leg *(i,j)*.

Constraints and objective function

$$Max \sum_{i \in IN(o,k)} \sum_{j \in OUT(o,k)} p_{i,j} x_{i,j} \tag{1.4.6}$$

$$subject \;\; to:$$

$$\sum_{j \in OUT(o,k)} x_{i,j} = 1 \;\; \forall i \in IN(o,k) \tag{1.4.7}$$

$$\sum_{i \in IN(o,k)} x_{i,j} = 1 \;\; \forall j \in OUT(o,k) \tag{1.4.8}$$

The *objective* function (1.4.6) maximizes through benefits that represent the set of most profitable through connections; the *inbound* constraint (1.4.7) ensures that only one flight is covered by a fleet type in the inbound flight; and the *outbound* constraint (1.4.8) ensures only one flight type is covered by a fleet type in the outbound flight. Note that this formulation is a generalization of the assignment problem, in practice, there are some additional constraints that must also be satisfied.

In addition, the airlines have to adhere to regulatory requirements. One such requirement involves the maintenance of the aircraft, therefore on a regular basis each plane stays overnight at a maintenance base for a minimum period of time. Given a flight schedule with aircraft assigned to it, the aircraft maintenance-scheduling problem is to determine which aircraft should fly which segment and when and where each aircraft should undergo different levels of maintenance check. Sriram and Haghani (2003) proposed a model with objective function that minimizes maintenance cost plus any costs incurred during the re-assignment of aircraft to the flight segments.

1.4.4 Crew Scheduling

The first task in the crew scheduling process is the process of matching pilots and flight attendants for a series of flights that start and end at the hub, this is known as *crew pairing*. In describing the problem (Barnhart et al., 1999), stated as follows: Given a set of flights (corresponding to an individual fleet type or fleet family), choose a minimum cost set of pairing such that every flight is covered exactly once (i.e. every flight is contained in exactly one pairing). The crew pairing problem is just the set partitioning problem, partitioning the set of flight legs into disjoint pairings, each containing a valid sequence of flights, to make the total cost the minimum:

Sets and indices

F^k is a set of daily flights assigned to fleet type k ($f \in F^k$); P^k is a set of feasible pairings for fleet type k ($p \in P^k$).

Decision variable

$$y_p = \begin{cases} 1 & \text{if pairing } p \text{ is in the solution;} \\ 0 & \text{otherwise.} \end{cases}$$

Parameters

c_p = is the cost of pairing p.

δ_{fp} = is 1 if flight f is included in pairing p, otherwise it is 0;

Constraints and objective function

$$Min \sum_{p \in P^k} c_p y_p \qquad (1.4.9)$$

$$subject\ \ to: \quad \sum_{p \in P^k} \delta_{fp} y_p = 1, \quad \forall f \in F^k \qquad (1.4.10)$$

The *objective* function (equation 1.4.9) minimizes the penalty cost of the pairing that partitions the flights; equation 1.4.10 ensures that every flight is included in exactly one pairing. This formulation requires explicit formulation of all pairings which can be difficult to achieve because of the numerous work rules that must be checked to ensure legality and also because of the huge number of potential pairings.

Next, the resultant solution (that generated possible pairings) needs to assign specific individuals to those pairings; we call this the *crew assignment* problem. When crew are allowed to bid on their preferred work schedule, usually based on seniority, the model gets transformed into a *bidding* problem. Otherwise, individualized schedules, lead to a *rostering* problem, where, unlike the generic *crew assignment* problem, the focus is both on crew needs and cost. The crew assignment model finds the cost-minimizing *bidlines* where the crew can bid for their top choices.

1.5 Conclusion

In this chapter we gave an overview of OR in the airline industry and narrowed down to four major areas of application. In the area of RM, we traced the origin, outlined the process and discussed the contemporary thinking in the field. We illustrated the concept through a simple *bid-price* model.

Next, we discussed the application within network planning focusing on fleet planning and route planning. In fleet planning, two major phases of aircraft acquisition were mentioned, the *type* and the *size* processes. As for the first phase, different selection criteria (such as, aircraft price, range, comfort, etc) are explained. As for the aircraft *capacity* selection, two approaches are outlined, that is, the *cell* approach and the *FA-4* model. In route planning, different process are elaborated ranging from route evaluations, selection of new markets, wave development and optimization, among others.

In the third area of OR application, schedule planning with the four inter-twined phases are mentioned. In the schedule generation phase,a further four-stage process is outlined, that is, demand estimation, schedule construction, slot management and schedule distribution. The second phase of schedule planning is the fleet assignment process that assigns aircraft to flight legs. Under fleet assignment, we illustrate the connection-based model followed by a description of the generic through-assignment model for the third phase. In the last phase (crew scheduling), we discussed the crew-pairing model.

Chapter 2

Airline Schedule Robustness

"The hard-dollar costs of disruption are relatively easy to calculate. The softer costs, such as lost passenger goodwill, are almost impossible to measure. Both are important" (Cook, 2002).

In this chapter, the definition of *robustness* in the context of airline schedules will be outlined, followed by different approaches and interpretations of *schedule robustness* as an important element during irregular operation. The popular *key performance indicators* that are used to measure *schedule robustness* are outlined. Later, the robust models and solution algorithm that have been proposed and developed to handle schedule robustness will be discussed. Finally, the differentiation between FAM robustness and operational robustness is made; with the former defining the scope of this research.

2.1 Defining Robustness

There is no universally scientific accepted definition of robustness (Vinke, 2003); it is context-driven and many analogies can be given. Robustness analysis (Rosenhead, 2002) embraces two important elements; the uncertainty factor that obstructs confident decision-making and that the decision must be or can be staged. Rosenhead has defined *robustness analysis* as a way of supporting decision-making when there is radical uncertainty about the future. In an analogous approach, robustness analysis has been equated as *sensitivity analysis*; where an optimal solution responds to changes in input that might vary in the future. Not surprising, even in the context of airline scheduling that there is no systematic way to define robustness (Lan et al., 2003). The ultimate aim of robustness is to reduce downstream impact and overall cost disruptions; (Clarke, 2004) outlined the following characteristic for airline robustness:-

- Robustness is the ability to perform as desired in as many operating conditions or situations as possible and return to normal operations as soon as possible after disruption

- Robustness requires both flexibility to respond quickly and damping
- Robustness is created by first understanding the dynamics of the system and then figuring out the right combination of flexibility and damping
- Robustness can only be measured in terms of performance in the operating environment

In today's fiercely competitive operating environment, any form of schedule inefficiency is very costly (Lohatepanout and Jacobs, 2003). Robustness can be created in two ways; one is to have a flexible schedule that will easily cope with demand fluctuation. The alternative approach is the damping effect where robustness is achieved by an independent schedule component such as having a spare aircraft to operate on disrupted flight. Ultimately, robustness is implemented by either integrating it into the schedules during the planning stage or re-optimizing the schedule after disruptions occur.

Schedule disruptions are caused by two major factors: (i) weather unpredictability that has a consequential effect of reducing or increasing airport capacity through diversions and, (ii) flight delays and cancellations that accounts for the larger part and are brought about by demand uncertainties, technical reasons such as aircraft maintenance and overall poor planning. Disruptions are highly inconvenient for passengers and connecting flights that cascade the delays across the network. The disruptions interfere with the fleet and crew assignment that needs to be re-scheduled to re-accommodate passengers. As the problem recurs, airline strength is assessed on the ability to improve robustness, as it can make or break the airline.

2.2 Measuring Robustness

A popular method of controlling schedule robustness is through a management process of measurement by *key performance indicators* (KPIs) and accountability. Rather than using a KPI trend, a combination of KPIs is used to set SMART targets, that is, targets which are specific, measurable, attainable, realistic and time-bounded. The

KPIs prompt for further analysis to identify the root cause of the disruptive component; appropriate remedial action is then taken to gauge against future recurrence. The root cause can be mitigated by many factors such as inefficient schedule design (e.g. short time allocation on the aircraft turn-around), poor and inadequate airport infrastructure, air traffic control delays and congestion, passenger behaviour, congestion, engineering delays, poor resource planning, interdependencies of connecting flights, unpredictable weather and rostering of both technical and cabin crew (Lee and Moore, 2003), (Wu and Reid, 2003) and (Bratu and Barnhart, 2005). The KPIs measures the quality of the schedule in two ways; either by its *integrity* and (or) *cost effectiveness*. Integrity refers to the schedule reliability and is a major indicator of operational excellence, typical KPIs are shown in Table 2.1.

Although the KPIs outlined above are the most popular, different airlines come up with varying KPIs that suit their business model. An example worth mentioning is a statistical approach developed by a team of researchers for KLM (Bian et al., 2005). In their approach, the number of *Aircraft On Ground* (ACOG) were analyzed on different schedule waves at KLM. Then a number of input variables were defined and correlated with schedule *performance indicators* (PI). The six or seven most highly correlated input variables were then regressed to determine the best balance between parsimony and explanatory power. The analysis on the expected number of aircraft on the ground has shown to provide a good prediction of the robustness of a given schedule. A summary of the major approaches used by researchers and practitioners in order to incorporate robustness in the schedule will be discussed in the next section.

2.3 Modelling Robustness Techniques

Researchers have tackled the schedule robustness issue from a mathematical modelling perspective, where at least two models get integrated or a re-modelling is done to incorporate a robust component.

Integrating schedule design and crew scheduling: As a way to enhance robustness, American Airlines (Lohatepanout and Jacobs, 2003) developed a more pragmatic approach by introducing a feedback loop between the integrated schedule design and

Key Performance Indicators	*Explanation*
% of flights cancelled	Number of cancelled flights divided by total number of flights.
% of late arrivals	Number of late arrivals divided by total number of flights.
Rate of flight cancellation	Total number of flights divided by number of cancelled flights.
Distribution of delays	Number of flights delays distributed over the week.
On-time performance	Number of flights that arrive or depart on time.
Disrupted passengers	Number of passengers who missed their flights.
Number of violation of crew rules that is *Flight Time Credit* (FTC)	FTC is the difference between the number of minutes paid and the number of minutes flown as a percentage of the number of minutes flown.
Number of upgrades	Frequency of change in the assignment to larger aircraft.
Average passenger delay	Mean number of passengers delayed in a given period.
Average delay per flight	Mean number of flights delayed for a given period.
Average delay of disrupted passengers	Number of passengers that have been affected by cancellation or missed connection in a given itinerary for a specific period.
% of disrupted passengers	Number of disrupted passengers divided by total number of passengers in a given itinerary for a specific period.
Average delay of non disrupted passengers	Mean number of passengers that were not disrupted but had their flight delayed for a specific period.

Table 2.1: Typical KPIs.

crew scheduling. This approach differs from the common practice where crew scheduling is explicitly done after FAM results have been obtained. Unlike the explicit integration, the feedback process is more flexible and provides a good mechanism to capture much of the impact associated with the other process. Although the model is data thirsty and sometimes difficult to quantify resource availability (due to complex union contracts and regulations); it offers the most efficient crew reduction plan that leads to significant profit improvement.

Degradable airline schedule: In the aftermath of disruption, the concept of delaying propagation (Kang, 2004) was advanced where, in order to minimize the impact of the whole schedule, the disruptive part of the schedule is isolated. The schedule is degraded into several smaller and independent schedules called layers (sub-schedules). The priorities for each layer are based on revenue where the information is available to passengers before they buy their tickets. In formulating the problem, integer programming was applied to find a feasible schedule that minimizes the total penalty. The penalty was incurred if an itinerary was not assigned to its desired layer (or higher priority layer). The research has now taken another dimension, instead of propagation delay we could have independent delay that is caused by other factors - not independent or not a function of the routing (Lan et al., 2003).

Flight schedule re-timing: In a re-timing (connection-based) model (Lan et al., 2003), the number of disrupted and mis-connected passengers are minimized. Passengers are disrupted when there is a flight delay (that is, slack time is *0*); hence adding more slack can be a solution for connecting passengers. By moving flight departure times in a small time window, better allocated slack is obtained. The models input included historical distributed data for a number of disrupted passengers for each connection with the model constraints defined as follows: a) for each flight, exactly one copy will be selected; b) for each connection, exactly one copy will be selected and this selected copy must connect the selected flight-leg copies; and c) the current fleeting and routing solution cannot be altered.

Integrating FAM with other models: FAM integration is one of the most vibrant research areas with a track record of the most successful implemented robust models. The FAM integrated models include crew scheduling, schedule design, aircraft routing, through-assignment, RM, etc. The integration of such models are discussed in the next chapter.

2.4 Heuristic Robustness Techniques

Similarly, the schedule robustness issue has also been addressed through a development and application of specific heuristic techniques; either applied as a solution algorithm or as a robust technique.

Neighbourhood search: A neighbourhood search algorithm has enjoyed application in the solving of FAM (Gotz et al., 1999), and is attributed to faster solutions compared to both *linear programming* and *integer programming* algorithms. An example worth mentioning is an algorithm developed (Love et al., 2002) with an objective function that minimises *real cost* (associated with direct operating cost) and *virtual cost* (associated with disruptive situations). Based on the *swap* neighbourhood, they implemented four heuristics; *Iterative Local Search* (ILS), *Revised ILS*, *Steepest Ascent Local Search (SALS)* and *Revised SALS* (RSALS). Another two-phase heuristic approach for the aircraft rotation problem was adopted by Air France (Ambrosini et al., 2003). Under phase one, the network flow problem was solved by the construction of an initial solution covering all the flight legs and optimizing flight cost connections. In the second phase, constraints were considered in order of priority by improvement of swaps between aircraft to respect the *Maximum Take-Off Weight* (MTOW) taking into account the overhaul.

Multi-objective genetic algorithm: Many integrated problems entail multiple objectives that can be solved as one instance as opposed to sequentially. For example, (Ahuja et al., 2003) modelled FAM with multi-criteria objectives where ground manpower and crew costs were added to the *combined fleet assignment* model (ct-FAM) (Ahuja et al., 2007). Similarly, a genetic algorithm (Lee et al., 2003) has been used to fine-tune the departure times of each of the flight legs in the schedule while preserving the aircraft rotation and crew assignment. The *model objective function* (multi-objective) was to optimize flight departure time with constraints on; crew connection (minimum turn time), crew maximum duty time, crew rest time, aircraft connection (minimum turn time), aircraft scheduled maintenance and station curfew. The decision variables were defined as the adjustments made to the original scheduled departure times. The idea was to find the *right* adjustments to all the flight legs so as to optimize robustness without violating the constraints.

Simulation approach: American Airlines (Green, 2002) developed a three-phase simulation approach where the schedule attributes that influence dependability were measured. The impact of changes in those attributes were then evaluated and optimized on an explicit account of dependability. On measuring the attributes into a dependability relationship, simulations were used to classify the attributes into four broad categories. Metrics for each schedule attribute were then defined and calculated in a prototype tool developed in SAS. The prototype provides the source data required for querying or reporting capabilities. At the next evaluation stage, operational dependability of alternative schedules were simulated while varying the schedule attributes to establish quantitative relationships to dependability. The dependability impact of broader issues (such as crew pairing strategies, system operating policies, schedule protection strategies, etc) were also evaluated. A typical existing tool in the market using the above approach is SIMAIR, that evaluates operational performance of schedules and recovery procedures. Delta (Zhao, 2003*b*) has also applied simulation in identifying potential flight delays.

In a stochastic model (Mederer and Frank, 2002) robustness has been defined as a schedule whose deterministic parameters are as insensitive to real life variables as possible. *Blocktimes* and *overall ground time* were considered as the major influences of the stochastic model. The *blocktime* comprised aircraft flying time, taxiing and holding time while *overall ground time* comprised the normal ground time and delay time. Under this approach, schedule was integrated with the stochastic Monte Carlo simulation model; the overall result was an impressive improvement on the schedule punctuality, regularity and connection quality.

A Multi-criteria approach: Delta (Zhao, 2003*a*) built a rotation model using a multi-criteria technique with the following objectives; maintaining schedule consistency, locking rotations in terms of specific aircraft assigned to specific legs, checking on ground time violations, checking on aircraft gate availability, consecutive red-eye and ensuring maintenance check for major, line and spread maintenance. A decision support system, *Operations Planning Model* (OPM), was developed that could achieve multiple business tasks. The OPM followed a two sequential step; the *global approach* and the *heuristic approach*. The global optimization model provides a good initial rotation and looks at the number of aircraft used, major maintenance, gate availability, fixed ground time violations and hard-forced connection. The heuristic approach looks into the multi-criteria optimization nature of an aircraft schedule and the trade-off between business rules and constraints. Furthermore, the approach looks at improving the robustness of an aircraft rotation, such as, line maintenance, maintenance spread, schedule consistency, crew connections, etc.

Delay perturbation: In formulating the basic minor delay perturbation model (Huag, 1998), the consequences of a flight delay whenever there is an incident was considered.

To achieve robustness, one or more of the following five strategies are used.

1. Speed up strategic model: Under this strategy, the node-arc formulation of FAM built extra arcs that link the origin node; consequently, the travel time between two points was minimized.

2. Cancellation of strategic model: This was achieved by changing the objective (cost) function to allow cancellation.

3. Swap an aircraft before departure time with an alternative aircraft.

4. Delay flight strategic model: The model is similar to the basic delay model but with an additional delay bundle arc constraint.

5. Ferry flights: This considers a spare aircraft in support of a disruptive schedule operation.

A similar model that not only includes options for delaying and cancelling flights, but also incorporates a measure of deviation from the original aircraft routings has also been proposed (Thengvall et al., 2000). A number of studies addressing the schedule perturbation problem with reference to a ground delay program as it occurs have also been presented (Luo and Yu, 1998).

Just-in-time (JIT) approach: Gershkoff (1998) advocated a six-stage approach for enhancing robustness. The first one is to cancel flights with zero booking. Although flights with zero booking hardly exist in real life situations, the concept of a zero booking flight is created in a high frequency market where customers have multiple equivalent flight times to choose from. The analysis done shows that by forcing the demand onto the non-JIT flights, profitability can be increased on off-peak days by idling the JIT aircraft. Although the concept of idling a perfectly serviceable aircraft contradicts scheduling managers objective of maximizing aircraft utilization, the analysis shows otherwise. The second is to substitute old aircrafts on JIT flights. Old aircraft tend to have a high operating variable cost and a low fixed cost, so it makes sense to allocate them to JIT. The third is to encompass all high-frequency routes; since as the number of flights in a high-frequency market decreases, the average time gap of somewhere between 1-2 hours, the assumption that traffic will spill to the nearest flight begins to break down, even for a monopoly route. The fourth strategy is to use marginal aircraft to chase highest transient revenue opportunities. Under this option, the aircraft is routed to a different city pair experiencing transient demand. When a particular route has high traffic, transient demand can suffice to fill the aircraft in both directions. The fifth strategy is to rationalize resources other than aircraft and crews. Although a difficult approach, matching ground resources to demand can generate high savings. It is argued that the ground resources needed by the airline to operate variable flights must be sensitive to demand fluctuations and therefore must become variable as well. The final strategy is to dynamically substitute aircraft to meet fine-grain demand variations. This refers to capacity rationalization process which, where demand is high, leads to increase of aircraft size instead of frequencies. Likewise, reduction of capacity size instead of decreasing frequencies when demand is low.

2.5 Robustness within FAM

The objective of robustness is to produce FAM solutions that perform well relative to uncertainty in demand and operations. A schedule planned in advance (more than three months to departure) is hardly ever flown as planned, especially for a network carrier. During these periods, many unforseen disruptive events would have occurred; such events could be categorized as either demand fluctuation or operational issues. Demand fluctuations necessitate capacity readjustments while operational issues, such as weather unpredictability and aircraft maintenance delays, lead to the schedule adjustment. While it is impossible to have a perfect schedule (Lee et al., 2003), the airline strength is measured on how robust its schedule is in coping-up with both demand uncertainties and operational performance.

2.5.1 Demand Robustness

The reliability of FAM solutions is pegged to the accuracy of cost and revenue estimates. The standard computation of *aircraft operating costs* is well defined, relatively stable and does not pose modelling difficulties. The major challenge is on estimating revenue through fare and demand forecasts. The total demand for a particular flight fluctuates for a given time of day, by day of week and season of the year. In addition to these more predictable or *cyclical* fluctuations in demand, there are also less predictable variations in demand around the mean or expected value for a flight. Similarly, passenger class allocation based on revenue management effects, adds to the complexity of revenue estimation. For this research, we concentrate on the demand uncertainty and how it can be generated and integrated into FAM.

2.5.2 Operational robustness

It has been noted that existing planning systems do not have effective methods to manage disruptions (Lan et al., 2003); hence airlines adopted a centralized process called Operations Control Center (OCC) or Station Control (Pbuero, 2001). Such a centre centrally manages operations of aircraft, crews and passengers. The centre monitors flight irregularities and comes up with a recovery plan that returns to the original schedule. The OCC calls for sophisticated systems as the challenges involved are enormous (Barnhart, 2003). Following a disruption, there is a little slack time to recover operations with a cost minimization model that involves a sophisticated plan for aircraft scheduling, crew scheduling and passenger routing. Furthermore, the time-frame of coming up with the best operational decisions is small and hard, any delay in coming up with a solution will make the decision obsolete. Additionally, the complexities involved could have severe consequences, as a small delay may propagate through the entire network, last for several days and affect both passenger goodwill and overall profits.

Serio is a typical decision support system that is used in a Station Control (Alvarez, 2002). Several studies and surveys for the irregular airline operations have been presented (Bratu and Barnhart (2006), Ball et al. (2007), Clarke (1997) and Seth et al. (1998)).

2.6 Conclusion

We have now seen that there is no universally scientific accepted definition of robustness; it is context-driven and many analogies can be given. In the context of airline scheduling, the ultimate aim of robustness is to reduce downstream impact and overall cost disruptions. We also noted that robustness can be created in two ways; one is to have a flexible schedule that will easily cope with demand fluctuation. The alternative approach is the damping effect where robustness is achieved by an independent schedule component such as having a spare aircraft to operate on disrupted flight.

The different interpretations of schedule robustness given lays a foundation in defining the scope of our research. In particular, the FAM robustness relates to variation or uncertainty in demand. The total demand for a particular flight fluctuates for a given time of day, by day of week and season of the year. In addition to these more predictable or cyclical fluctuations in demand, there are also less predictable variations in demand around the mean or expected value for a flight. For this research, we concentrate on the demand uncertainty and how it can be generated and integrated into FAM.

Chapter 3

Airline Fleet Assignment Modelling

The FAM is enjoying increasingly wide application and is one of the most vibrant research areas. In this chapter a review of the development of *basic* FAM is presented. The chapter also provides a detail description using a *time-line* structure, and describes its solution algorithm. Also the various FAM integrated models are discussed. Later, we propose two alternate ways of capturing aircraft utilization that represents the deterministic equivalent of the SP-FAM (to be described in *Chapter 5*). Finally, the challenges associated in estimating cost and revenues data as input requirements to the FAM are outlined.

3.1 FAM Development

The airline fleet assignment problem has undergone a major metamorphosis since inception (Ferguson and Dantzig, 1955). They formulated a combined fleet assignment and aircraft routing model that maximizes operating profit for a fixed schedule with known deterministic demand. During that period until early 90's, several other FAMs came into existence, but none enjoyed wide application. The first true FAM, that lead to the *basic* FAM discussed in the next section, is credited to Abara (1989) who used the *connection-network* formulation discussed in Chapter 1. In Abara's solution approach, the *schedule balance* and the *aircraft availability* constraints are relaxed with appropriate penalty functions added to the objective function. The structure of this *connection-network* necessitated specification of all feasible connections that lead to a very large scale model. To solve this problem, Abara had to put a threshold on the number of possible connections for each flight.

Using a similar formulation, (Rushmeier and Kontogiorgis, 1997) came up with some preprocessing techniques that solved the problem without the need to specify feasible flight connections. They came up with the idea of partitioning stations into *complexes* that represented feasible connections with schedule balance i.e., where there is an equal number of incoming and outgoing legs. In consideration of additional crew-based constraints, they designed a heuristic to solve the problem. In the heuristic, LP relaxation is first solved and the resulting solution is rounded to obtain an initial solution that is solved by branch-and-bound. This solution used a considerable computing time.

(Berge and Hopperstad, 1993) and (Hane et al., 1995), were among the first researchers to use the *time-line* representation that has largely become the method of choice in formulating subsequent fleet assignment problems. The *time-line* network structure is essentially a *multi-commodity network flow problem.* The commodities are the different types of aircraft with flights as a common resource to be used by only one commodity. Unlike the *connection* network, the *time-line* network gives more flexibility in establishing feasible connections that fall within the time and space consideration. The flexibility manifests itself in the reduction of decision variables where the number of flight legs is far less than the number of possible connections. Hane's formulation is similar to Abaras' but with significant computational improvements. The model includes a *ground-arc* network that tracks aircraft on the ground, but unlike Abaras, does not indicate aircraft turns. The model also proposed the use of interior-point programming and branching, that have since become a standard method of solution algorithm.

After the period of the basic FAM development, a new era of tackling robustness within FAM was addressed. This was marked by extensions; mainly focusing on integrating FAM with other models and using FAM to address irregular operational issues.

3.2 Basic FAM Formulation

In describing the *basic* FAM (Hane et al., 1995), we summarized the tutorial given by Sherali et al. (2006). Unlike the *connection network* described earlier, the *timeline* network is cyclic with time-lines for each airport. Between time-lines there are arcs for all the flights, (where *flight arcs* represent flight legs), and each time-line has a node for each arriving or departing flight. The *ground arcs* correspond to aircraft waiting times between the nodes (or staying at the same station for a given time). We also have *wrap-around arcs* that ensure overnight connectivity linking the last events of the day with the first events of the next day. As each airport station is tied to one fleet type, the airports are duplicated into multiple copies making independent networks that allows only one fleet type to be assigned to each flight leg. Thus, given any fleet type in an independent network (or sub-network) for each station, a network time-line consists of a series of event nodes that occur sequentially with respect to time, along with the ground and wrap-around arcs that link these event nodes. Note that the arrival node is placed at the *block* time which is the actual arrival time plus the aircraft turn-around time ready for the next departure.

Figure 3.1 shows a time-line network for two stations (A and B) with two fleet types (*Type 1* and *Type 2*). The nodes represent an event (arrival or departure) according to its time of occurrence. To make a distinction between the fleet types, the full arrows denote the fleet of *Type 1* and the broken arrow to the fleet *Type 2*. The slanted arrows represent *flight arcs*, the vertical arrows are the *ground arcs*, and the curved arrows are the *wrap-around arcs*. The wrap-around arcs are shown only for *Type 1* and are suppressed for *Type 2*. The arc pairs (A_1, A_2) to (F_1, F_2) represents flights *A* - *F* flown by *Types 1* and *Types 2*, respectively. Note that *A1* stays on the ground for a shorter time before the next departure on flight *B1* using *Type 1*, compared to *A2* that requires relatively longer turn-times before the next departure on flight *B2* using *Type 2*.

The fleet assignment problem can be described as: given a flight schedule with fixed departure times and aircraft operating cost on each flight leg, find the least cost assignment of aircraft types to flights, such that (1) each flight is covered by exactly one fleet type, (2) flow of aircraft by type is balanced at each airport, and (3) only the available number of each type are used.

Sets and indices

L: the set of flight legs indexed by i;

K: set of different aircraft types indexed by k;

O: set of stations (airports) indexed by o;

T: the sorted set of all event (arrival or departure) times at all airports, indexed by t. The last node in the counting line is denoted by t_n. The index t^+ denotes an event after time t while t^- denotes an event before time t ($t \in T$).

R: the set of nodes in the network indexed by $\{k,o,t\}$;

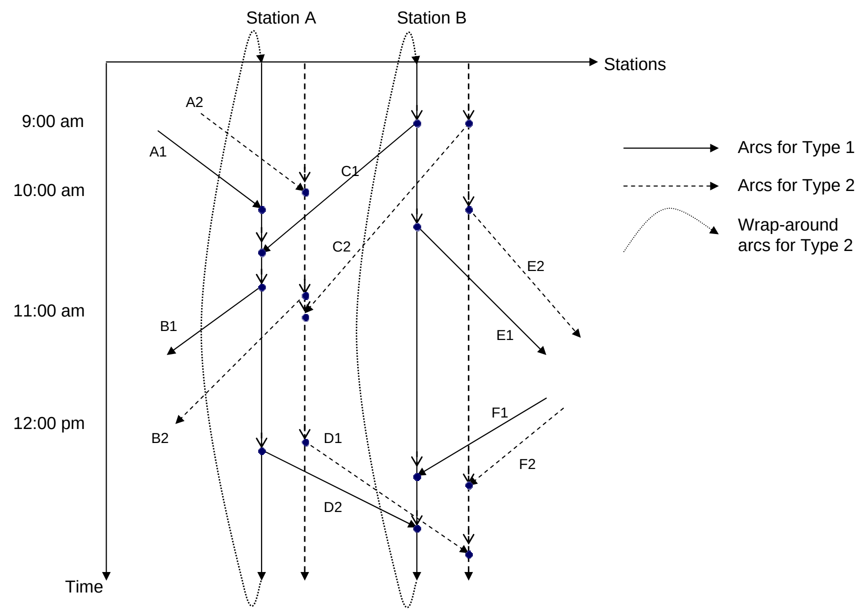


Figure 3.1: A two-type, two station fleet-flow time-space network.

CL(k) denotes the set of flight legs crossing the count time (i.e. t_n) flown by k i.e. the set of flight legs where an aircraft of type k may be in the air at time t_n;
I(k,o,t) is the set of inbound flight legs to node $\{k,o,t\}$; and
O(k,o,t) is the set of outbound flight legs from node $\{k,o,t\}$.

Decision variables

$$f_{k,i} = \begin{cases} 1 & \text{if fleet type } k \text{ is assigned to flight leg } i \text{ ;} \\ 0 & \text{otherwise.} \end{cases}$$

$Y_{k,o,t} \geq 0$ is the number of aircraft of fleet type k, on the ground at station o, and time t; Y_{k,o,t^+} is the number of aircraft of fleet type k, on the ground at station o, just following time t ($t \in T$); and Y_{k,o,t^-} is the number of aircraft of fleet type k, on the ground at station o, just prior to time t ($t \in T$).

Parameters

$c_{k,i}$ denotes aircraft *direct operating cost* of assigning aircraft k on flight leg i (£ per aircraft flight leg); and
N_k denotes the number of available aircraft of type k.

Constraints and objective function

$$Min \sum_{k\in K}\sum_{i\in L} c_{k,i} f_{k,i} \qquad (3.2.1)$$

$$subject\ \ to:$$

$$\sum_{k\in K} f_{k,i} = 1 \quad \forall i \in L \qquad (3.2.2)$$

$$Y_{k,o,t^-} + \sum_{i\in I(k,o,t)} f_{k,i} - Y_{k,o,t^+} - \sum_{i\in O(k,o,t)} f_{k,i} = 0 \quad \forall\{k,o,t\} \in R \qquad (3.2.3)$$

$$\sum_{o\in O} Y_{k,o,t_n} + \sum_{i\in CL(k)} f_{k,i} \leq N_k \quad \forall k \in K \qquad (3.2.4)$$

The objective function (3.2.1) is to minimize the cost of the fleet assignment; the *cover* constraint (3.2.2) ensures that each flight is covered once and only once by a fleet type. Constraint 3.2.3 and 3.2.4 can be explained using Figure 3.2 that consider events happening in station A using fleet type k in the time-line network. In the diagram, we have three arriving (inbound) flights A2, A6 and A8 occurring at times $t = 1$, $t - 1$ and t, respectively. Similarly, we have three departing (outbound) flights B1, C1 and Z1 occurring at times t, $t+1$ and t_n, respectively. Note that when arrivals and departures occur simultaneously (e.g. A8 arrives at t and B1 departs at the same time), arrivals precedes departures in the time-line. The *conservation flow* constraint (3.2.3) ensures *aircraft balance*, that is, the number of aircraft of type k on ground in city o just before time t (i.e. Y_{k,o,t^-} or Y_{k,A,t^-} in our case) plus the number (sum) of inbound flights on aircraft of type k arriving in station o at time t (i.e. $\sum_{i\in I(k,o,t)} f_{k,i}$ or flight A8 in our case) must be equal to the number (sum) of outbound flights on aircraft of type k departing from station o at time t (i.e. $\sum_{i\in O(k,o,t)} f_{k,i}$ or flight B1 in our case) plus the number of aircraft of the same type on ground in city o just after the departure time t (i.e. Y_{k,o,t^+} or Y_{k,A,t^+} in our case). The conservation equation makes up the bulk of the model constraints. There is one constraint for each aircraft at each node, where each flight results in two nodes. Hence, if there are 10 aircraft and 2,500 flights per day, and if every aircraft could fly every flight leg (which is not true due to operational limitations), and if all departure and arrival times for a given aircraft at a given city were unique (also not true), then we would have 2,500 $\times$ 2 $\times$ 10 = 50,000 conservation equations.

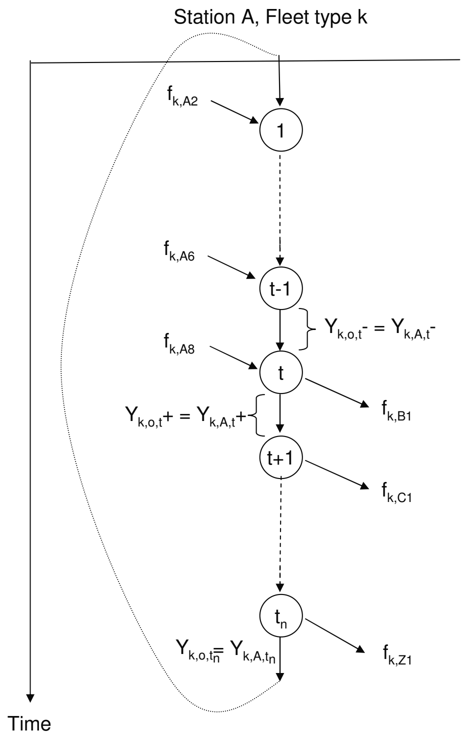
The *count* constraint (3.2.4) ensures that for each aircraft type, the total number of aircraft on the ground or in the air at any point in time cannot exceed the total available. It is enough to ensure that the count constraint is satisfied at one particular point in time i.e. the conservation flow constraint ensure that if the count constraint is satisfied at some point of time, it is also satisfied at any other point in time. If we take a snapshot of time at the event in the last node (t_n), the number (sum) of aircraft of type k on ground at station o (i.e. $\sum_{o \in O} Y_{k,o,t_n}$) plus the number (sum) of flights in air using aircraft k (i.e. $\sum_{i \in CL(k)} f_{k,i}$) should not exceed the total available (i.e. N_k). As a simple illustration and for simplicity, assume the event times for flight A2-Z1 are as shown in Figure 3.2.

$$\{t = 1 = 3\ am,\ t - 1 = 6\ am,\ t = 12\ pm,\ t + 1 = 3\ pm,\ t_n\ = 9\ pm\} \in T$$

In this case, the count time at the last node (t_n) is set to 9 pm. Note that the number of aircraft on ground count (Y_{k,o,t_n}) at station station A will remain the same between 9 pm until the following day at 3 am (as there are no other events happening). Since the available number of aircraft (i.e. N_k) is known and that the number of assigned flights (on air) using aircraft k at 9 pm (i.e. $\sum_{i \in CL(k)} f_{k,i}$) can be measured, if the constraint holds at 9 pm, then the conservation flow constrain ensures that it will hold at any other given time.

3.3 Solution Method

The sheer size of the FAM problem and the required computational time necessitates a series of preprocessing steps to make the representation more compact. In the first instance, the *node aggregation* step is done where flights that arrive before a departing flight in some pre-defined timing are consolidated. In this case, the order of a node's event occurring does not matter as long as all the connections are intact. Figure 3.3 illustrates the concept where *Part (a)* shows a network time-line representation of a station for one type having flights A-H. The downward vertical arc representing the time occurrence in increasing order of each event. *Part (b)* demonstrates the node aggregation where flights B-C and D-G are consolidated into two nodes representing a total reduction of four nodes from the initial eight. This preprocessing step has the



Flight, i	Time	Event	Description
A2	t=1	Arrival	From city M to A
.	.	.	.
.	.	.	.
A6	t-1	Arrival	From city N to A
A8	t	Arrival	From city P to A
B1	t	Departure	From city A to B
C1	t+1	Departure	From city A to C
.	.	.	.
.	.	.	.
Z1	t_n	Departure	From city A to Z

Note:
1. t_n is the last node in the counting line; its successor is t=1. Between the last node (t_n) and the first node (t=1) there are no events happening; the aircraft ground count remains the same i.e. Y_{k,o,t_n}
2. A2, A6 and A8 are the inbound flights.
3. B1, C1 and z1 are the outbound flights.
4. At the current time, denoted by t, two events are taking place i.e. an arrival (flight A8) and a departure (flight B1).

Figure 3.2: A one-type, one station fleet-flow time-space network.

effect of removing thousands of rows and columns from the initial steps, making it easier to solve without infeasibility.

The second preprocessing step is based on heuristic observations of the network structure. For example, in a hub-and-spoke structure, flights between spokes can be reduced significantly. Normally, the activity at the spokes have sporadic flights and if one makes an assumption that no aircraft should be on the ground at some time, the ground arcs can be removed to form what are called *islands* as shown in *Part (c)*. Furthermore, since there are no ground arcs from one island to another, some of the flights within the island can be aggregated as one.

The third preprocessing technique *eliminates missed connections* by creating *aircraft balance*. If two flights that must be flown consecutively result in a missed connection, an aircraft balance will not be maintained because of the longer turn-times for that fleet type. This pair of flights can either be removed from the circulation or create a dummy flight that will maintain the aircraft schedule balance.

After the above preprocessing steps, solved the LP relaxation problem. Thereafter,

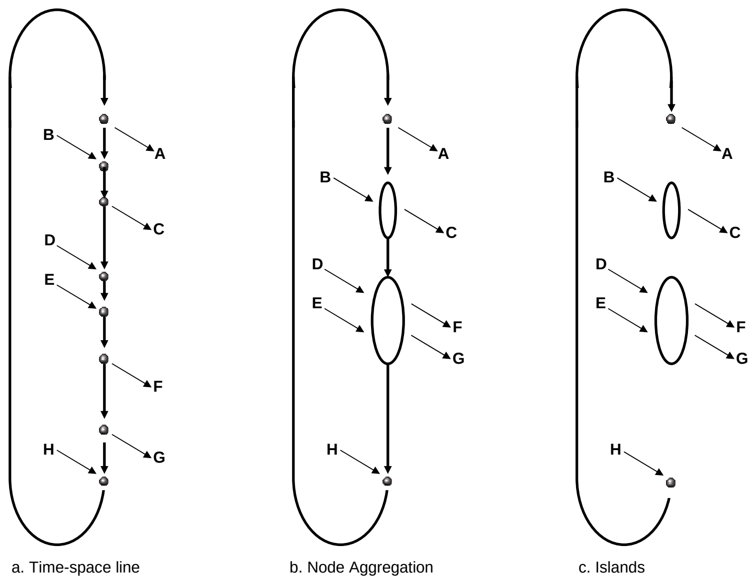


Figure 3.3: Network reduction for a time-space network.

a rounding heuristic is invoked to fix the variables resulting in a new problem. The new problem is then solved using branch-and-bound, or a combination of algorithmic strategies such as interior-point methods, or dual steepest-edge simplex approach, among others. As the solution method involves a heuristic phase, it is not guaranteed to be optimal. Nevertheless, it is reported that the IP objective after the rounding is very much closer to the LP objective.

A further milestone in the reduction of the computational effort, is a heuristic algorithm, *simulated annealing*, which has now become popular in substituting LP and IP algorithms. Specifically, introducing side constraints to the basic FAM, such as *homogeneity* and *time windows* makes the use of LP and IP very difficult. The homogeneity constraint forces the flight to be flown by the same sub-fleet on different days while time windows allow departure times to vary slightly around a preferred time. Gotz presented an algorithm that proceeds in three phases: a preprocessing, a simulated annealing and a postprocessing phase (Gotz et al., 1999). The first phase aims at reducing the number of legs, as in the aggregation step discussed earlier. In this case, the algorithm, initially allows explicit specification of sequential flight legs that can be grouped together. Thereafter, the algorithm provides the option to combine legs on low-frequented airports. In the simulating annealing phase, a neighbourhood search algorithm is defined that allows for some transition stages, *change* and *swap*, that have the effect of altering and changing sub-fleets during the FAM execution. Finally, the postprocessing phase follows suit with a *hill-climbing* algorithm that uses the same neighbourhood function as the simulation annealing and ensures termination in a local optima.

3.4 FAM Extensions

This section describes the various forms of FAM integration that contribute to robustness.

Integrating FAM and crew scheduling: FAM has been integrated with a *crew pairing* model (Barnhart, Lu and Shenoi, 1998) known as the *duty-based* model (or DPP). The DPP model was not only much easier to solve compared to the generic *crew pairing* but had a relatively good approximation solution. The integrated model comprised various sub problems; the fleet assignment sub problem and a number of crew scheduling sub problems, one for each fleet family. The objective function minimizes the total fleet assignment and total time-away-from-base crew costs. In solving the integrated model, they developed *an advanced sequential solution approach* that replaced the basic FAM with an approximated model. In this approach, the fleeting decisions are first deduced, thereafter, crew pairing problem is solved for each fleet family.

Integrating FAM with schedule design: (Desaulniers et al., 1997) integrated FAM with the scheduling design process by allowing departure times to vary within certain time-windows, thus allowing different connection possibilities for the FAM. Similarly, Rexin et al. (2000) followed suit with a similar model but with a set of discrete time windows for each leg that represent possible departure times.

Integrating FAM with aircraft routing problem: The *aircraft routing* problem determines a sequence of flights, or routes, that must be flown by individual aircraft such that the assigned flights are included in exactly one route and all aircraft can be maintained as necessary. This process normally follows after the FAM has been executed. The combination of the FAM with the *aircraft routing* problem in a *string* based model was also presented (Barnhart, Boland, Clarke, Johnson, Nemhauser and Shenoi, 1998). Where a *string* refers to a sequence of connecting flights that begins and ends at maintenance stations, satisfies flow balance, and is maintenance feasible.

The objective of the *string* model is to select the set of strings that have the minimum time necessary to perform maintenance attached to the end of the last flights in the string; such that, each flight segment is assigned to exactly one fleet, and, for any fleet, its assigned flights are partitioned into a set of rotations with each aircraft in the fleet assigned to one rotation at most and the total cost is minimized.

The robustness problem could also be formulated where flight cancellation and delays are considered simultaneously. Such a model (Clarke, 1997) allows for multiple fleet aircraft swapping in flight scheduling, provided the candidate aircraft is capable of flying a given flight segment. The overall framework of the model is represented on a time-space network called the *schedule map* and uses efficient tree-searching algorithms as a solution algorithm. The model has been described as a hybrid of the traditionally defined fleet assignment and the aircraft routing or rotation problem.

Integrating FAM with through-assignment model: The *Through-Assignment Model* (TAM) has been integrated (Ahuja et al., 2007) with FAM into a single model known as the *combined fleet assignment model* (ctFAM). The models are solved sequentially with optimal fleeting decisions from FAM used to solve TAM. The objective of the TAM model is to determine through connections. The ctFAM is then solved heuristically using the neighbourhood search algorithm with the optimal solutions from both FAM and TAM used as a starting point for the neighbourhood search. Figure 3.4 (Ahuja, 2006) shows the sequential steps for solving the ctFAM. In *Step 1*, FAM is first solved without consideration of through revenue. Thereafter, TAM is solved without changing the initial generated FAM solution. In *Step 2*, the model maximizes the total contribution of both the fleet assignment and through assignment.

Integrating FAM with station purity: Similarly, there have been advances in the concept of *station decomposition* (Smith, 2004) where *station purity* was integrated into FAM. *Station purity* ensures that the number of fleet types serving a given station does not exceed a specified limit and is classified into maximum purity (where spoke stations use only one fleet type) and moderate purity (when small spokes use one fleet type, larger spokes use one or two). The decomposition concept is made in reference to the hub-and-spoke model for a typical airline flight network. If a hub is removed from such a network, the network decomposes into a set of spoke stations (station purity), with its own set of flights.

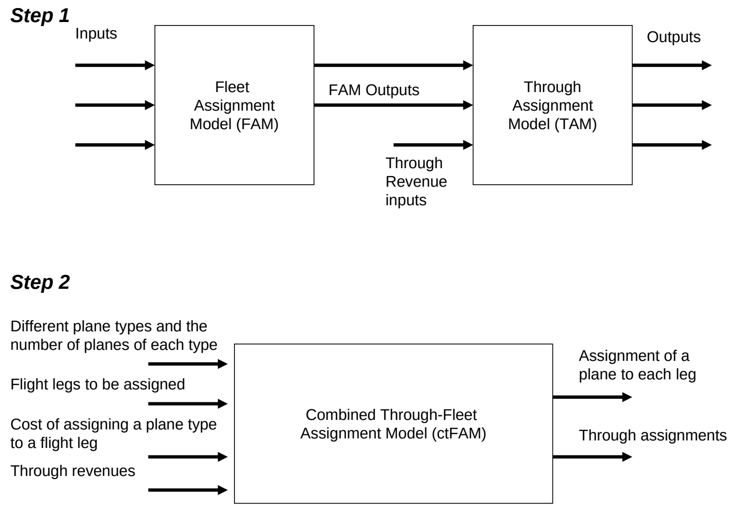


Figure 3.4: The combined fleet assignment model.

Integrating FAM with RM: Many approaches have been investigated on incorporating RM into FAM. Smith (2004) gives a detailed survey of the integrated aspects and shows that FAM solution quality is sensitive to revenue assumptions. In particular, incorporating RM into FAM gives a superior solution than using average passenger revenues. One of these approaches, is known as the *tactical* FAM which represents the deterministic equivalent of SP-FAM. In the next section, the *tactical* FAM will be discussed and in *Chapter 5* a detailed description of the SP-FAM will be presented.

3.5 FAM with Aircraft Utilization

We adopted the Lufthansa System (Lufthansa, 2006) distinction made on their NetLine-Plan and NetLine-Sced planning tools between *strategic FAM* and *tactical FAM*, respectively. Unlike the *tactical FAM*, the *strategic FAM* is integrated with the *network-simulation model* that represents an ideal market under competitive forces where passengers are captured, spill, and are sometimes lost to a competitor. A typical *strategic FAM* is the *itinerary based FAM* (Barnhart et al., 2002) that has the salient features mentioned above. On the other hand, the *tactical FAM* is integrated with the airline's own projected demand, has spill but does not consider the recapture effect.

The *tactical FAM* is identical to the *basic FAM* discussed earlier except for the modification of the objective function (equation 3.5.1) that includes both the aircraft assignment cost (i.e. $\sum_{k\in K}\sum_{i\in L} c_{k,i} f_{k,i}$) and an additional passenger revenue (i.e. $\sum_{k\in K}\sum_{i\in L} X_{k,i} \times q_i \times f_{k,i}$). The passenger revenue computation considers passenger demand, aircraft spill and average ticket price. Note that spill occurs when passenger demand exceed aircraft capacity.

$$Min \sum_{k\in K}\sum_{i\in L}(c_{k,i} - X_{k,i} \times q_i) \times f_{k,i} \tag{3.5.1}$$

where,

d_i denotes passenger demand on flight leg i;
u_k denotes a maximum spill factor on aircraft k;
$seats_k$ denotes number of seats on aircraft k;
q_i denotes average ticket price on flight leg i (£ per aircraft); and
$X_{k,i}$ denotes the minimum of the number of aircraft seats of type k or demand after spill on aircraft k for flight i (customers), that is, $X_{k,i} := \min\ (seats_k,\, d_i \times (1 - u_k))$. If passenger spill on aircraft k with 150 seats has a factor of say 0.1 and the demand for the flight is 200, then the $X_{k,i}$:= min(150,180)=150.

Some of the major challenges associated with FAM optimized results are the aircraft over-utilization and under-utilization that makes the resultant fleeting decision non-implemental. In such a case, the optimization expert will seek to trade-off aircraft by imposing swap restrictions and re-optimize until a realistic utilization balanced is attained. Here, the objective of FAM is not only to minimize cost (or maximize profit) but to ensure that there is aircraft balance in terms of utilization. A wide-body aircraft has a higher *direct operating cost* and a consequent opportunity (or idle) cost compared to a narrow-body aircraft. Although the network planner's goal is to ensure a higher utilization for all aircraft, given a conflicting utilization result, a wide-body should have a higher utilization preference during the re-optimization process.

Over-utilization is good from financial perspective and does not pose a major concern, since it depicts efficiency and a high return on investment. However, it could also mean that the schedule is not robust to disruption. An over-utilized aircraft has little slack time to account for the flight delays, maintenance delays, or simply, the aircraft cannot be used as a standby. Nevertheless, an average standard maximum aircraft threshold block-time value (MX_k) can easily be defined to account for the robust slack-time. Conversely, under-utilization is of much greater concern as it depicts an aircraft not efficiently being utilized with a low return on investment. As such, an average minimum aircraft threshold value (MN_k) is usually defined based on an industry average or more specific tailored as KPI to the airline.

On the foregoing basis, it would appear trivial to include a constraint that restricts the aircraft utilization within the stipulated range (i.e. MX_k and MN_k). In this case, we would extend the tactical FAM and include constraint 3.5.2.

$$N_k \times MN_k \leq \sum_{i \in L} f_{k,i} \times B_{k,i} \leq N_k \times MX_k \quad \forall k \tag{3.5.2}$$

The constraint ensures that the total number of block-time flown on aircraft k on all flight legs (i.e. $\sum_{i \in L} f_{k,i} \times B_{k,i}$) is less or equal to the total maximum threshold value (i.e. $N_k \times MX_k$) on aircraft k and more or equal to the total minimum threshold value (i.e. $N_k \times MN_k$) on the same aircraft. $B_{k,i}$ denotes the block-time on flight leg i using aircraft type k. Block-time refers to the number of hours (minutes) incurred by an aircraft from the moment it first moves for a flight until it comes to rest at its intended blocks at the next point of landing, or returns to its departure point prior to take-off.

Unfortunately, the use of such a *hard* constraint highly limits the aircraft assignment and leads to infeasible results. In reality, some aircraft will always fall below the MN_k value and can only be controlled to a limited extent as opposed to getting confined within the threshold boundary. If the MN_k value is made smaller (or zero), the infeasible results would be eliminated and optimality attained, however, the constraint would become *redundant* or insensitive to smaller incremental values. To counter the inherent deficiency, we offer two approaches that integrate the aircraft utilization to the tactical FAM.

Option 1: DFAM1

The most ideal way is to impose an aircraft utilization cost (w_k) whenever the utilization falls below the MN_k value. In this case, we are using a *softer* restriction where under-utilization is still acceptable below the MN_k value but at a cost. The *aircraft utilization* constraint (denoted by equation 3.5.7) computes the number of under-utilized block-time on aircraft of type k by subtracting the total number of block-time flown on aircraft k on all flight legs (i.e. $\sum_{i \in L} f_{k,i} \times B_{k,i}$) from the available minimum that the aircraft can operate (i.e. $N_k \times MN_k$). The utilization variable is split into two parts (i.e. $r_k = rp_k$ - rm_k), the positive part ($rp_k \geq 0$) denotes the under-utilized variable and the negative part ($rm_k \geq 0$) denotes the expected-utilized variable (i.e. within the acceptable or expected range). The new formulation of the deterministic equivalent of the SP-FAM is identical to the *tactical* FAM with modification of the objective function and an addition of the utilization constraint.

$$Min \sum_{k \in K} \sum_{i \in L} (c_{k,i} - X_{k,i} \times q_i) \times f_{k,i} + \sum_{k \in K} rp_k \times w_k \tag{3.5.3}$$

$$subject \;\; to:$$

$$\sum_{k \in K} f_{k,i} = 1 \quad \forall i \in L \tag{3.5.4}$$

$$Y_{k,o,t^-} + \sum_{i \in I(k,o,t)} f_{k,i} - Y_{k,o,t^+} - \sum_{i \in O(k,o,t)} f_{k,i} = 0 \quad \forall \{k,o,t\} \in R \tag{3.5.5}$$

$$\sum_{o \in O} Y_{k,o,t_n} + \sum_{i \in CL(k)} f_{k,i} \leq N_k \quad \forall k \in K \tag{3.5.6}$$

$$rp_k - rm_k = N_k \times MN_k - \sum_{i \in L} f_{k,i} \times B_{k,i} \quad \forall k \in K \tag{3.5.7}$$

Note that the variable splitting technique depicted in equation 3.5.7 and the linear objective function ($\sum_{k \in K} rp_k \times w_k$) is a special case of a convex separable piecewise-linear objective function, with a slope of zero where the variable is negative (rm_k) and a slope of w_k where the variable is positive (i.e rp_k).

An alternate way of representing the objective function and aircraft utilization constraint 3.5.7 is by changing the equation sign (from = to $\geq$) and not splitting the utilization variable (note that in this case $r_k \geq 0$). Using this representation, constraints 3.5.4 - 3.5.6 will remain intact while the objective function (3.5.3) will be substituted with equation 3.5.8 and constraint 3.5.7 change to equation 3.5.9.

$$Min \sum_{k\in K}\sum_{i\in L}(c_{k,i} - X_{k,i} \times q_i) \times f_{k,i} + \sum_{k\in K} r_k \times w_k \tag{3.5.8}$$

$$r_k \geq N_k \times MN_k - \sum_{i\in L} f_{k,i} \times B_{k,i} \quad \forall k \in K \tag{3.5.9}$$

In this formulation, the solver will automatically look for the lowest feasible value of r_k given all other variables. If the right-hand side (rhs) of the constraint is non-negative (implying a case of under-utilization), r_k will be set equal to the rhs (i.e. the computed under-utilized value); if the rhs is negative (implying a case of the expected-utilization), the sign restriction on utilization (i.e. since $r_k \geq 0$) will result in r_k automatically being set to zero.

Option 2: DFAM2

Another way of modelling the *aircraft utilization* is to introduce a constraint that will compute the utilization variable ($r_k \geq 0$) as the difference between the total maximum threshold value on aircraft k (i.e. $N_k \times MX_k$) and the total number of block-time flown on the same aircraft on all flight legs (i.e. $\sum_{i\in L} f_{k,i} \times B_{k,i}$). The r_k value is then penalized with a utilization cost (w_k) in the objective function. This approach aims at gaining high utilization rate for each aircraft irrespective of the MN_k value. Using this representation, constraints 3.5.4 - 3.5.6 will still remain intact as in *tactical* FAM, while the objective function and the utilization constraint will change. The model will be represented as:

$$Min \sum_{k\in K}\sum_{i\in L}(c_{k,i} - X_{k,i} \times q_i) \times f_{k,i} + \sum_{k\in K} r_k \times w_k \tag{3.5.10}$$

$$subject \ \ to:$$

$$\sum_{k\in K} f_{k,i} = 1 \quad \forall i \in L \tag{3.5.11}$$

$$Y_{k,o,t^-} + \sum_{i\in I(k,o,t)} f_{k,i} - Y_{k,o,t^+} - \sum_{i\in O(k,o,t)} f_{k,i} = 0 \quad \forall \{k,o,t\} \in R \tag{3.5.12}$$

$$\sum_{o\in O} Y_{k,o,t_n} + \sum_{i\in CL(k)} f_{k,i} \leq N_k \quad \forall k \in K \tag{3.5.13}$$

$$r_k = N_k \times MX_k - \sum_{i\in L} f_{k,i} \times B_{k,i} \quad \forall k \in K \tag{3.5.14}$$

3.6 Cost and Revenue Data

In order to produce the fleeting decisions, FAM requires estimates of both aircraft operating costs and revenues per flight leg. Airlines use different conventions in the computation of aircraft operating costs that are fairly stable and reliable. As such, cost does not pose any modelling challenges (Smith, 2004). The problem comes in the estimation of revenue, that is a function of the aircraft capacity, the uncertain passenger demand and fare variation.

An average fare is derived from a price that has comprehensive elements which include levels, rules, routings, booking class, and distribution channel. The problem is further compounded as the fare paid needs to be apportioned to the specific flown sectors and later weighted against the passenger demand to arrive at an average unit fare. While it is easy to analyze a historical average unit fare, it is not easy to forecast the same at flight level given the high seasonal variation.

Similarly, demand for any flight varies by class type for each day of week and season. Estimating actual demand, especially close to flight departure, is almost impossible. A good forecasting model will always strive to project close to reality. The relationship between price and demand is shown at Figure 3.5. *Part (a)* shows the relationship before price segmentation with a distinction between those who are prepared to pay more and those who would paid less. *Part (b)* is the scenario of multiple prices and the maximization of revenue.

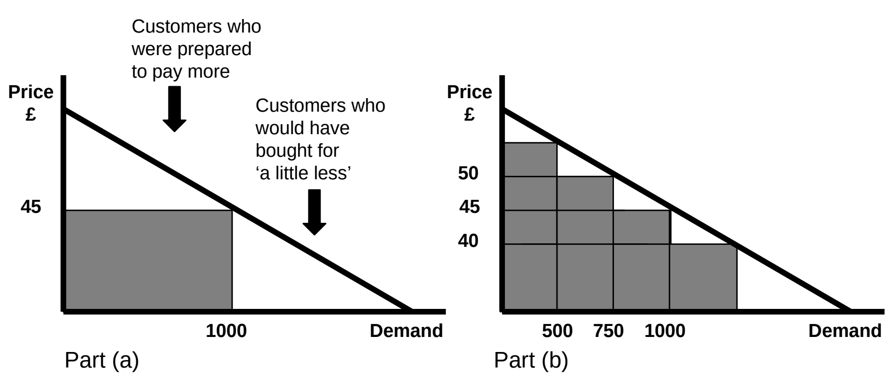


Figure 3.5: The relationship between price and demand.

Essentially, the challenges of combining price and demand has been captured in the *origin-destination yield management* model (Smith, 2004) that can be stated as:

Sets and indices

L is a flight leg ($i \in L$); P is a set of all passenger types, defined by their itinerary and fare class ($p \in P$); and p_i is the subset of passenger type p on leg i.

Decision variables

$alloc_p$ is the number of seats allocated to passenger type $p \in P$, $traf_p$ is the number of passengers carried (traffic) by passengers of type $p \in P$.

Parameters

Dmd_p is the demand for passenger type $p \in P$; cap_i is the seating capacity of flight $i \in L$; and rev_p is the average revenue per passenger for passenger type $p \in P$.

Constraints and objective function

$$\text{Max} \sum_{i \in L} \sum_{p \in p_i} rev_p E(traf_p | Dmd_p, alloc_p) \tag{3.6.1}$$

$$subject \;\; to: \quad \sum_{p \in p_i} alloc_p \leq cap_i \quad \forall i \in L \tag{3.6.2}$$

Where the *objective* function *3.6.1* maximizes total expected revenue across all flight legs by finding allocations for each path subject to the *capacity* constraint 3.6.2 that ensures the sum of allocation on each leg is less than the leg capacity. Note that $E(traf_p | Dmd_p, alloc_p)$ is the expected traffic for passenger type $p \in P$ given the demand distribution for passenger p, Dmd_p and seating allocated to passenger type p, $alloc_p$, that is:

$$\begin{aligned} E(traf_p | Dmd_p, alloc_p) &= P(Dmd_p \leq alloc_p) * Dmd_p \\ &+ (1 - P(Dmd_p \leq alloc_p)) * alloc_p \end{aligned} \tag{3.6.3}$$

The major challenge is to forecast revenue before integrating such a model into the basic FAM. The tactical FAM described earlier, resembles the integration of the basic FAM and the above RM model.

3.7 Conclusion

We have now described the basic FAM and the inherent problem associated with aircraft utilization. We then explored different FAM extensions before coming-up with our new extensions (i.e. DFAM1 and DFAM2) that effectively tackles the aircraft utilization problem. In addition, the DFAMs lays a concrete foundation for a further extension into the SP-FAMs (to be described in Chapter 5).

Chapter 4

Stochastic Programming (SP) Modelling Approaches

This chapter describes the SP as a method of choice for modelling optimization problems that involve uncertainty. We start with a background information and discuss the major classes of SP problems. In particular, the recourse-based approach would eventually lead to an extension of the DFAM (described in section 3.6) to SP-FAM (in Chapter 5). The other modelling approaches and their inter relationship could be linked to Chapter 8 where the proof of concept is made. The different solution algorithms discussed could be linked to Chapter 7 where the SP-FAM is solved.

4.1 Background

A mathematical programming problem in which some of the data are unknown, that is, they are subject to uncertainty, random influences or statistical variations is called a stochastic programming (SP) problem. In general, SP problems can be classified into three (Mitra, 2001) major categories i.e. recourse problems, distribution problems and chance-constrained problems.

4.2 The Recourse Problems

The classical linear program with recourse partitions the problem variables into two stages, those that have to be decided *here-and-now* (the first-stage decisions), and those that can be decided after the uncertain parameters reveal themselves (the second-stage recourse decisions). In this approach the key underlying decisions must be made currently in the face of future uncertainties. At a later time, the uncertainties are resolved by observing a joint realization of the values of all uncertain parameters. At that time, corrective (recourse) actions are taken in response to the outcomes that materialize. The objective is to minimize the expected total cost, which includes the direct cost of the first-stage decisions and the expected cost of the second-stage corrective actions.

Let

ω denote the random event,

$x \in \Re^{n_1}$ denote the first-stage decisions,

$c \in \Re^{n_1}$ denote the cost associated with the first-stage decisions,

$b \in \Re^{m_1}$ denote the right-hand side of the first-stage system,

$A \in \Re^{m_1 \times n_1}$ denote the constraint matrix of the first-stage decisions,

$y(\omega) \in \Re^{n_2}$ denote the second-stage decisions,

$q(\omega) \in \Re^{n_2}$ denote the cost of the second-stage system,

$h(\omega) \in \Re^{m_2}$ denote the right-hand side of the second-stage system,

$B(\omega) \in \Re^{m_2 \times n_1}$ denote the linking matrix corresponding to the first-stage decisions in the second-stage system,

$D(\omega) \in \Re^{m_2 \times n_2}$ denote the matrix corresponding to the second-stage decisions in the second-stage system.

The two-stage SP problem with recourse is expressed as:

$$
\begin{aligned}
Z_{hn} \qquad\qquad & \qquad min\ cx + E_\omega(q(\omega)y(\omega)) && (4.2.1)\\
& subject\ \ to: \\
& \qquad\qquad Ax \ \geq\ b \\
& B(\omega)x + D(\omega)y(\omega) \ \geq\ h(\omega) \\
& \quad x \geq 0, y(\omega) \geq 0.
\end{aligned}
$$

The set of constraints $B(\omega)x + D(\omega)y(\omega) \geq h(\omega)$ describe the links between the first-stage decisions x and the second-stage recourse actions $y(\omega)$. Note that we require that this constraint holds with probability 1, or for each possible $\omega \in \Omega$. The objective function of Z_{hn} contains a deterministic term cx and the expectation of the second-stage objective $q(\omega)y(\omega)$ taken over all realizations of the random event ω, $E_\omega(q(\omega)y(\omega))$. The second-stage term is difficult to evaluate because, for each ω, the value $y(\omega)$ is the solution of a linear program. As each component of q,h,B, and D is a possible random variable. Let $B_i(\omega)/D_i(\omega)$ be the i^{th} row of B(ω)/D(ω). Piecing together the stochastic components of the second-stage data, we obtain a vector $\xi(\omega) = (q(\omega), h(\omega), B_{1.}(\omega), \ldots, B_{m_2}(\omega), D_{1.}(\omega), \ldots, D_{m_2}(\omega))$, with potentially up to N $= n_2 + m_2 + (m_2 \times (n_1 + n_2))$ components. A single random event ω (or scenario) can influence several random vectors, $\xi(\omega)$. Therefore, model Z_{hn} can be written as

$$
\begin{aligned}
Z_{hn} \qquad & \min\ cx + E_{\xi}Q(x,\xi(\omega)) \\
& subject\ \ to: \\
& Ax \geq b \\
& x \geq 0
\end{aligned}
\tag{4.2.2}
$$

$$
\begin{aligned}
where \quad Q(x,\xi(\omega)) = \min\ \ & q(\omega)y(\omega) \\
subject\ \ to: & \\
B(\omega)x + D(\omega)y(\omega) \ \ & \geq \ \ h(\omega) \\
y(\omega) & \geq 0
\end{aligned}
\tag{4.2.3}
$$

Recourse problems are extended in a number of ways. One of the most common is to include more stages. A general multi-stage recourse problem with T stages can be written as (Dempster (1980) and Birge (1988)):

$$
\begin{aligned}
& Min_{x1}\{q_1x_1 + E_{\xi(\omega_2)}\{min_{x2}(q_2x_2 + E_{\xi(\omega_3)|\xi(\omega_2)}\{min_{x3}(q_3x_3 + ... \\
& \qquad E_{\xi(\omega_T)|\xi(\omega_1)...\xi(\omega_{T-1})}\{min_{x_T}(q_Tx_T\}....]\})\}\} \\
& subject\ \ to:
\end{aligned}
$$

$$
\begin{array}{lcl}
A_1x_1 & \geq & b_1 \\
B_2x_1 + A_2x_2 & \geq & b_2 \\
\quad B_3x_2 + A_3x_3 & \geq & b_3 \\
\qquad . & & . \\
\qquad\ . & & . \\
\qquad\ \ . & & . \\
\qquad\qquad B_Tx_{T-1} + A_Tx_T & \geq & b_T
\end{array}
$$

$l_1 \leq x_1 \leq u_1$

$l_t \leq x_t \leq u_t \quad t = 1, ..., T$

where $\xi(\omega_t) = (b_t, q_t, A_{t,1}, ..., A_{t,n_t}, B_{t,n_t})$, t=2,..T, are random vectors in some canonical probability space $(\Omega, \mathcal{F}, Pr)$. The sub-indices of the matrices **B**, **A**, and the vectors **q**, **x**, **b** refer to the stages of the problem. Once the realized values are observed at a stage *t*, the information required to decide the actions at stage *t+1* is known.

The interpretation of the above model is as follows: first, decision vector x_1 is chosen to satisfy period 1 constraints, and then in period 2, after having observed the realization of random vectors $\xi(\omega_1)$, a decision vector x_2 is chosen. In general, period t, decisions x_t have to be adapted to the information arrival process $\{\xi(\omega_1), ...\xi(\omega_t)\}$, as well as to the decision sequence $\{x_1, ...x_{t-1}\}$; namely, decisions should be taken before the outcome of future realizations of random events, a requirement known as *non-anticipativity* of decisions. Multi-stage SP models with recourse corresponds to the sequence of anticipatory decisions taken prior to observing future random events, as well as adaptive decisions that must be made to compensate for deviations from prescribed targets as random events unfold.

Letting S be the total number of scenario, model Z_{hn} can also be re-written in discrete representation as:

$$
\begin{aligned}
Z_{hn} \qquad & \qquad min\ cx + \sum_{s=1}^{|S|} p_s q_s y_s \qquad (4.2.4)\\
& subject\ \ to:\\
& \qquad Ax \ \geq\ b\\
& B_s x + D_s y_s \ \geq\ h_s \ \ \forall\ s \in S\\
& x \in \Re^{n_1}, y_s \in \Re^{n_2}, y_s \geq 0
\end{aligned}
$$

As presented, without assuming any additional properties or structure on Z_{hn}, we would describe the problem as having *general recourse.* In many cases, there is specific structure in the recourse subproblem that can be exploited for computational advantage. Some of the more commonly types are described as follows.

Simple Recourse: A special case of the recourse model, known as the simple recourse model, arises when the constraint coefficient matrix in the second-stage problem, D, form an identity matrix. That is, D=[I -I], and $y(\omega)$ is divided correspondingly as $(y^+(\omega),y^-(\omega))$ and $q(\omega)$=$(q^+(\omega),q^-(\omega))$. Note that, in this case, the optimal values of $y^+(\omega)$,$y^-(\omega)$ are determined purely by the sign of $h(\omega)$-$B(\omega)x$ provided that $q^+(\omega)$+$q^-(\omega) \geq 0$ holds with probability one.

Fixed Recourse: A fixed recourse problem is one in which the constraint matrix in the recourse subproblem is not subject to uncertainty (i.e., it is fixed or not random). The recourse subproblem with $Q(x, \xi(\omega))$ representation is similar to *equation 4.2.3* but without the random variable in the matrix D. However, when the second stage objective coefficients are also fixed, the dual representation of the recourse subproblem (with fixed set of dual feasible solutions) is given by:

$$\begin{aligned} Q(x, \xi(\omega)) = Max \ \ \pi^T(h(\omega) - B(\omega)x) \\ subject \ \ to: \quad \quad \quad \quad \\ \pi^T D \leq q^T \\ \pi \geq 0. \end{aligned} \tag{4.2.5}$$

Complete Recourse: This property holds when there exist $y(\omega) \geq 0$ such that $D(\omega)y(\omega)$=p for all $p \in \Re^{m_2}$. A slightly less strenuous property, which leads to the same result, is known as *relatively complete recourse*. Suppose we define two sets for the model Z_{hn} as: $K_1 = \{x|Ax = b, x \geq 0\}$ and $K_2 = \{x|Q(x, \xi(\omega)) < \infty\}$. A stochastic program has *relatively complete recourse* if $K_1 \subset K_2$, that is, every solution x that satisfies the first-stage constraints, $Ax = b$, has a feasible completion in the second-stage.

4.3 Distribution Problems

The optimization problems which provide the distribution of the objective function value for different realizations of the random parameters and also for the expected value of such parameters are broadly known as the distribution problems.

The Expected Value Problems: A more quantifiable approach is to solve the original linear program where all the random data have been replaced with the expected values. Let $\bar{f} = \sum_{s=1}^{|s|} p_s f_s$, $\bar{B} = \sum_{s=1}^{|s|} p_s B_s$, $\bar{D} = \sum_{s=1}^{|s|} p_s D_s$ and $\bar{h} = \sum_{s=1}^{|s|} p_s h_s$. We define the expected value model as

$$
\begin{aligned}
P_{ev} \; &= \; min \;\; cx + \bar{f}y \\
subject \;\; to: & \\
Ax &\geq b \\
\bar{B}x + \bar{D}y &\geq \bar{h} \qquad (4.3.1)
\end{aligned}
$$

Let $(\bar{x}, \bar{y})$ be the optimal solution to P_{ev} and $Z_{ev}(\bar{x}, \bar{y})$ be the corresponding value. In order to evaluate the quality of the first-stage solution, we fix the first-stage variables in Z_{hn} to $\bar{x}$, and define the resulting model as:

$$
\begin{aligned}
P_{eev} \; &= \; min \;\; c\bar{x} + \sum_{s=1}^{|s|} p_s q_s y_s \\
subject \;\; to: & \\
A\bar{x} &\geq b \\
B_s\bar{x} + D_s y_s \geq h_s \qquad &\forall s \qquad (4.3.2)
\end{aligned}
$$

and having the objective value Z_{eev}. The objective value, Z_{eev}, obtained by the model 4.3.2 is called the *expectation of the expected value* solution.

Wait-and-See Problems: In contrast to here-and-now problems, which yield optimal solutions that achieve a given level of confidence, wait-and-see problems involve a category of formulations that shows the effects of uncertainty on optimum design. A wait-and-see problem involves deterministic optimal decision at each scenario or random sample, equivalent to solving several deterministic optimization problems. Consider a family of scenarios dependent (wait-and-see) models,

$$
\begin{aligned}
P_{ws}(s) \; &= \; min \;\; cx + q_s y_s \\
subject \;\; to: & \\
Ax &\geq b \\
B_s x + D_s y_s &\geq h_s \qquad (4.3.3)
\end{aligned}
$$

For a given s, let (x_s^*, y_s^*) be the optimal solution to $P_{ws}(s)$ and $Z_{ws}(x_s^*, y_s^*)$ be the corresponding objective value. Processing all the scenarios and aggregate them, that is

$$Z_{ws} = \sum_{s=1}^{|s|} p_s Z_{ws}(x_s^*, y_s^*)$$

where p_s is the probability for all the scenario s.

4.4 Chance-Constrained Problems

Another important class of SP models are the chance-constrained problems (CCP) accredited to Charnes and Cooper (1959). These problems are characterized with a constraint that holds with a probability. The general formulation (Valente et al., 2005) of a chance-constrained problem is:

$$\begin{aligned} Z_{CCP} &= min \;\; cx \\ subject\ to: &\quad A_0x = b_0 \\ P\{A_ix \geq h_i\} \geq \beta_i &\quad i = 1..I \end{aligned} \tag{4.4.1}$$

where $\beta_i \in$ [0,1] is a reliability level and $\xi_i = (A_i, h_i) \;\; \forall \;\; i = 1..\mathrm{I}$ is a random vector on the probability space $(\Omega, \digamma, Pr)$. If the A_i is a row vector, the i-th constraint is called individual chance constraint. If A_i is a $r \times c$ matrix with $r > 1$, then the i-th constraint is referred to as joint chance constraint.

4.5 The Inter-relationships and Bounds

It can be easily shown that the three solutions, Z_{ws}, Z_{hn}, Z_{eev} are connected by the ordered relationship: $Z_{ws} \leqslant Z_{hn} \leqslant Z_{eev}$. The difference $(Z_{eev} - Z_{hn})$ is known as the value of stochastic solution (VSS). VSS measures how much better the solution of the stochastic optimization problem is in relation to the expected solution of the expected value problem.

The difference $(Z_{hn} - Z_{ws})$ is known as the expected value of perfect information (EVPI). EVPI is interpreted as the expected value or the amount the decision-maker is willing to pay to have perfect information, which is, knowledge about all the future scenarios. A relatively small EVPI indicates that better forecasts will not lead to significant improvement while a relatively large EVPI means that incomplete information about the future may prove costly. The EVPI and VSS can also be shown to be bounded as:

$$\begin{aligned} 0 \leqslant EVPI \leqslant Z_{hn} - Z_{ws} \leqslant Z_{eev} - Z_{ev} \\ 0 \leqslant VSS \leqslant Z_{eev} - Z_{ev} \end{aligned}$$

4.6 Solution Approaches

Given the discrete representation as denoted by equation 4.2.4, we can create very large-scale linear programs whose solution is the same as that of the stochastic program. These Deterministic Equivalents (DE) have a very peculiar algebraic structure which can be exploited by different solution methods. Some of the common known techniques are described in this section.

4.6.1 Universe

In this approach (Valente et al., 2005), the SP problem is expressed as DE and solved by general-purpose linear programming solvers. The discrete distribution can be represented through an *event tree* with nodes associated with the realisations of the stochastic quantities. The introduction of an event tree to describe uncertainty, allows the formulation of the DE problem which can have implicit or explicit non-anticipativity constraints.

In the case of explicit constraints, the event tree (refer to section 6.1.1 for a detail description) is splitted path-wise and the decision process follows the scenario evolution. The decision hierarchy is in this case forced along every scenario consistently with the original tree structure. The procedure leads to S dynamic problems, where S represents the number of scenarios, characterized by the same time structure where each scenario describes a unique path from the root of the tree to the leaf node. Non-anticipativity constraints are thus added explicitly to ensure feasibility of the decisions with respect to the set of information constraints.

In the case of implicit constraints the property of non-anticipativity is automatically fulfilled by introducing a unique vector of decision variables for each node of the tree making sure that the random coefficients of the problem are properly associated.

4.6.2 Decomposition

There are two types of decomposition based approaches depending on whether the scenario tree is split according to time stages (primal decomposition) or scenarios (dual decomposition). In this subsection, we discuss the popular method of each type i.e the *L-Shaped* method (based on primal decomposition) and *Lagrangian* method (based on dual decomposition).

L-Shaped Method: The first solution procedure proposed for two-stage stochastic linear programs with recourse is the L-Shaped method (Slyke and Wets, 1969). The L-Shaped method decomposes the problem by stage; the first-stage problem leads to a *master problem* and the second-stage problem leads to a subproblem. In reality, the method is simply an adaptation of *Benders decomposition* to the structure of the second-stage problem. A general principle behind the L-shaped approach is that, since the recourse function, $Q(x, \xi(\omega))$, involves a solution of all second-stage recourse linear programs, we would like to avoid numerous function evaluations for it.

The algorithm is initiated by solving the master problem first and then checking feasibility with the subproblems. If the subproblem is infeasible, a *feasibility cut* constraint is added to ensure feasibility of the first-stage decision. Next, the algorithm checks for optimality with the subproblem, and if not optimal, adds an *optimality cut* constraint that is a linear approximation of *Q(x)* on its domain of finiteness, and is determined based on the dual of the second-stage problem. The procedure is repeated until all the subproblems have been exhausted.

Depending on the starting point of the chosen subproblem and the number of cuts added, the algorithm could be time consuming before getting a solution. However, substantial amount of time could be saved if we take advantage of a solution for a single subproblem in solving the others (i.e. a *bunching* technique). In many ways, the L-Shaped method is by now a *classic* technique for solving two-stage stochastic programs. Although it is well suited for an introduction to stochastic linear-programming solution methodology, it is no longer computationally effective for large scale problems. One of the first major improvements to this basic methodology involved the introduction of a *regularizing* term (a quadratic proximal term added to the objective, $||x - x^k||^2$, for example).

Lagrangian Method: Lagrangian relaxation is a technique used when difficult problems can be turned into easy problems by eliminating a subset of the constraints. Penalties for violating the constraints are introduced into the objective. These penalties are then adjusted in an attempt to *price* the constraint violation, and the problem is solved again. The objective value of the solution to the pricing problem provides a bound to the original problem.

The basic motivation behind the Lagrangian approaches is that only the non-anticipativity constraints linking the scenarios are relaxed. Consider the dual representation of equation 4.2.4 and let λ_s denote the dual multipliers associated with second-stage constraints for scenario s. Since the non-anticipativity constraints are the *hard* constraints, we can place them in the objective and penalize them. If we define $\tilde{\lambda} = (\lambda_1, ..., \lambda_{|s|})$, then the lagrangian formulation becomes:

$$\begin{aligned} Max_{\tilde{\lambda}}\ LG(\tilde{\lambda}) &= cx + \sum_{s=1}^{|S|} p_s q_s y_s + \sum_{s=1}^{|S|} p_s(\lambda_s(h_s - B_s x + D_s y_s)) \qquad (4.6.1) \\ & subject\ \ to: \\ & Ax \geq b \end{aligned}$$

From the general optimality condition, the optimal value of the lagrangian dual (equation 4.6.1) is equal to the optimal value of the primal problem (equation 4.2.4). The common lagrangian solution algorithms include the Dual Ascent method, Lagrangian Finite Generation Method for Linear-Quadratic Stochastic Programs and Linear Progressive Hedging algorithm (Birge and Louveaux, 1997).

4.6.3 Statistically Based Method

One of the major handicaps of the L-Shaped method is the need to solve a subproblem for each scenario. In large-scale problems, the number of scenarios is much too high for this to be a reasonable approach and the possibility of using statistical estimations (Higle, 2005) of the recourse function becomes computationally attractive. Conceptually, the simplest method for incorporating statistical approximations in the solution procedure is to replace the recourse function, $Q(x, \xi(\omega))$, by a sample mean approximation. That is, if $\{\omega^t\}_{t=1}^{n}$ is a collection of independent and identically distributed observations of $\xi(\omega)$, then one might consider undertaking the solution of the *sample mean problem*:

$$
\begin{aligned}
Z_{hn} \qquad\qquad & Min \;\; cx + \frac{1}{n}\sum_{t=1}^{n} Q(x, \omega^t) \qquad (4.6.2)\\
subject \;\; to: & \\
& Ax \geq b \\
& x \geq 0
\end{aligned}
$$

In this fashion, the sample mean problem is a stochastic program with an alternate distribution. When the sample size is significantly smaller than the number of scenarios in the original distribution, equation 4.6.2 will be much easier to solve than Z_{hn}. On the surface, this approach is quite appealing, however, we note that the solution obtained is dependent on the specific sample that was drawn. Consequently, it is subject to error in the same manner that the mean value solution is subject to error. In the absence of relatively complete recourse, it is possible that the solution obtained is actually infeasible! That is, there may be some scenarios that are not represented in the sample for which the solution to equation 4.6.2 is not feasible. For this reason, relatively complete recourse is a critical property for problems that are to be solved using statistical approximation schemes.

4.6.4 Decomposition Methods with Sampling

A potential disadvantage of statistical based method, is that some effort might be wasted on optimizing when the approximation is not accurate. An approach to avoid this problem is to use sampling within another algorithm without complete optimization. Birge (1997) considered two such approaches, one uses *importance sampling* to reduce variance in deriving each cut based on large samples and another based on a sample that grows as the algorithm progress.

The first approach (Dantzig and Glynn, 1990) is to sample $Q(x,\xi(\omega))$ in the L-shaped method instead of actually computing $Q(x)$. Given an iterate x^k, the result is an estimate, $Q^v(x^k)=\frac{1}{v}\sum_{i=1}^{v}Q(x^k,\xi^i(\omega))$, and an estimate of $\nabla Q(x^k)$ as $\bar{\pi}_k^v = (\frac{1}{v}\sum_{i=1}^{v}\pi_k^i)$ where $\pi_k^i \in \partial Q(x^k,\xi^i(\omega))$. Now, for Q convex in x one obtains

$$Q(x,\xi^i(\omega)) \geq Q(x^k,\xi^i(\omega)) + (\pi_k^i)^T(x - x^k) \qquad \forall x \tag{4.6.3}$$

Assuming that we also have a finite value in $Q(x^k,\xi(\omega))$ for any $\xi(\omega)$ to prevent problems with infeasibility,

$$Q^v(x) = (\frac{1}{v})\sum_{i=1}^{v}Q(x,\xi^i(\omega)) \geq Q^v(x^k) + (\bar{\pi}_k^v)^T(x - x^k) = LB_k^v(x) \tag{4.6.4}$$

where, by the *central limit theorem*, $\sqrt{v}$ times the right-hand side in equation 4.6.4 is asymptotically normally distributed.

An alternate technique that integrates elements of decomposition techniques and statistical approximation techniques is Stochastic Decomposition (SD) (Higle and Sen, 1991). Unlike the sample mean optimization in equation 4.6.2, which operates with a fixed sample size, SD operates with an adaptive sample size; increasing the sample size as iterations progress. Unlike the L-Shaped method, which solves a subproblem for each scenario for each cutting plane constructed, SD uses recursive approximation methods based on previously solved problems in order to bypass the solution of the vast majority of the subproblems that would otherwise be solved. The combination of adaptive sampling and subproblem approximations has proven to be quite powerful, especially when a regularized master program is used.

4.7 Examples of SP Solvers

Most of the solvers would require the problem to be represented in SMPS format before execution. The SMPS format is an extension of MPS (Mathematical Programming System) format that converts the existing deterministic linear programs into stochastic ones through the addition of information about the dynamic and stochastic structure of the model.

BNBS:(bouncing nested Bendes solver) is an implementation of the nested Benders algorithm for solving multi-stage stochastic linear programming problems (Altenstedt, 2009). The name bouncing nested benders solver comes from one of the implemented sequencing protocols. In this protocol, the active stage bounces up and down in the tree.

FortSP: (the SP extensions to FortMP) is an integrated modelling and solver system based on extended language constructs designed to facilitate the formulation of scenario based recourse problems (OptiRisk, 2009). FortSP is embedded within the Stochastic Programming Integrated Environment (SPInE) that produces model instances in SMPS format. The integrated algorithms include Benders decomposition, DE with Implicit and Explicit Non-antipativity (Deteqi and Deteqe).

MSLiP: (multi-stage stochastic linear programming) implements a multi-stage version of the 2-stage L-Shaped method of Van Slyke and Wets using nested Bender's decomposition for the multi-stage stochastic linear programming problem (Gassman, 1990). The solver supports an arbitrary number of time periods and various types of random structures for the input data (in SMPS format).

DDSIP: (dual decomposition in two-stage stochastic mixed-integer programming) implements a number of scenario decomposition algorithms for stochastic linear programs with mixed-integer recourse (Caroe et al., 2009). The implemented algorithm is the Lagrangian relaxation of the non-anticipativity constraints and a branch-and-bound algorithm that re-establishes non-anticipativity.

SLP-IOR: An interactive model management system for two-stage (multi-stage) recourse and chance-constrained models (Kall and Mayer, 1996). The solver handles algebraic structure and scenario generation problems. The implemented algorithms include: Benders, Regularized and Stochastic Decomposition, Discrete Approximation, Interior point methods, Supporting hyperplane, Central cutting plane, etc.

4.8 Conclusion

We have now introduced the recourse-based concept that would be used to extend the DFAM to SP-FAM. Similarly, the SP solutions algorithms discussed can be related to the SP-FAM solution approach. The various SP models and the inter-bound relationship have also been introduced to build a base for the proof of concept.

Chapter 5

The SP-FAM

In this chapter we start by outlining the problem statement and discuss previous attempts in the formulation. We then present two alternate ways of formulating the SP-FAM and describe the models salient features.

5.1 Problem Description

A flight is characterized by a pairing of origin-destination, aircraft type, aircraft operating costs and estimated passenger revenue. While aircraft operating costs are fairly stable and known with certainty, revenue estimates that comprise average fare and demand are highly uncertain. Given a heterogeneous fleet and a scheduling horizon under conditions of demand uncertainty, we are looking for the best assignment of aircraft types to flight.

Our approach is to model the problem as a two-stage SP with recourse. The main distinction of the *recourse FAM* with other FAMs is that, given a number of passenger demand scenarios, it gives a strategic fleet assignment solution that hedges against all possible tactical solutions. In addition, we have a tactical solution for every scenario. The four unique features of the modelling approach are: an introduction of buffer aircraft within the available fleet; introduction of an artificial fleet assignment variable that satisfies the recourse definition; inclusion of aircraft utilization constraint; and inclusion of demand-spill constraint that links the first-stage decision variable and the second-stage decision variable of the recourse formulation.

5.2 Previous Attempts

According to Smith (2004), (Berge and Hopperstad, 1993) pioneered the research area using the *Demand Driven Dispatch* (D^3) concept. D^3 makes swaps after crew scheduling instead of prior to and reassigns capacity near the departure date with crew-compatible aircraft only. As the swapping is limited to only crew-compatible aircraft, the adoption of the concept has not only been at a relatively slow pace but its application has been limited to carriers that have crew-compatible aircraft.

Next, Lister and Dekker (2002) developed the idea of robust FAM through a scenario aggregation approach. The proposed model had the objective of determining the fleet composition that maximized profit in a system that allows capacity swapping close to departure. They used a two-stage stochastic linear programming model; the first stage solved a single scenario and the second stage solved a deterministic FAM model for each set of scenarios. While this is an SP-based approach, the solution is not optimal. (Pilla et al., 2005) quickly noticed that since the initial assignments are based on a single scenario, the result cannot be robust relative to variations in demand.

With the above limitations, (Pilla et al., 2005) came with a more robust FAM that utilizes a two-stage SP alongside the concept of D^3 to assign crew-compatible aircraft in the first-stage, so as to enhance the demand capturing potential of swapping in the second-stage. To overcome the problem of using a single scenario as in Lister's model, they used an average of the scenarios to estimate the expected revenue value of the recourse function for the two-stage SP model. In generating the demand scenarios, they used a statistical model (multivariate adaptive regression splines) fitted into an optimized-based computer experimental design (using Latin hypercube). The above approach has three inherent limitations: the optimization of recourse function was not considered; in a true SP approach, averaging of scenarios (also known as the expected value model), does not give relatively better solutions than when solving an SP model with all scenarios (called here-and-now model); and finally, the reliability of using computer experimental data to generate the uncertain demand is highly questionable.

5.3 Mathematical Formulation

The two-stage recourse FAM

In this section we present a new formulation of the SP-FAM that overcomes the inherent deficiencies of the previous attempts. The *two-stage recourse FAM* is an extension of the deterministic equivalent models described in *Section 3.5* and uses the same mathematical notations with modification in the following areas:

- Scenarios: The introduction of scenarios, and probability p_s for each scenario s
- Decision variable: An additional artificial aircraft related variable is created, *virtual* assigner. The artificial variable corresponds to *virtual profit* that is accrued to the *false* flight legs.
- Parameters: An additional standby (spare or buffer) aircraft is introduced that attracts a penalty cost $j_{k,i}$. Unlike the other scheduled aircraft, the standby aircraft does not attract the utilization (or idleness) cost (i.e. w_k) as described for the case of DFAMs.
- Demand-spill constraint: Addition of the demand-spill constraint that links the first-stage decisions and the second-stage recourse actions (see constraint 5.3.5 and 5.3.11).

The resulting models, SP-FAM1 and SP-FAM2, follow the two deterministic equivalent models discussed earlier, DFAM1 and DFAM2, respectively.

The SP-FAM1

$$Min \;\; C_{U1} + Z_{AP} + Z_{VP} \tag{5.3.1}$$

$$Subject \;\; to:$$

$$\sum_{k \in K} f_{k,i} = 1 \;\; \forall i \tag{5.3.2}$$

$$Y_{k,o,t^-} + \sum_{i \in I(k,o,t)} f_{k,i} = Y_{k,o,t^+} + \sum_{i \in O(k,o,t)} f_{k,i} \;\; \forall \{k,o,t\} \in R \tag{5.3.3}$$

$$\sum_{o \in O} Y_{k,o,t_n} + \sum_{i \in CL(k)} f_{k,i} \leq N_k \;\; \forall k \tag{5.3.4}$$

$$(f_{k,i} + b_{k,i,s}) \times seats_k \leq D_{k,i,s} \;\; \forall k,i,s \tag{5.3.5}$$

$$rp_k - rm_k = N_k \times MN_k - \sum_{i \in L} f_{k,i} \times B_{k,i} \;\; \forall k \tag{5.3.6}$$

The SP-FAM2

$$Min \;\; C_{U2} + Z_{AP} + Z_{VP} \tag{5.3.7}$$

$Subject \;\; to:$

$$\sum_{k \in K} f_{k,i} = 1 \;\; \forall i \tag{5.3.8}$$

$$Y_{k,o,t^-} + \sum_{i \in I(k,o,t)} f_{k,i} = Y_{k,o,t^+} + \sum_{i \in O(k,o,t)} f_{k,i} \;\; \forall \{k,o,t\} \in R \tag{5.3.9}$$

$$\sum_{o \in O} Y_{k,o,t_n} + \sum_{i \in CL(k)} f_{k,i} \leq N_k \;\; \forall k \tag{5.3.10}$$

$$(f_{k,i} + b_{k,i,s}) \times seats_k \leq D_{k,i,s} \;\; \forall k, i, s \tag{5.3.11}$$

$$r_k = N_k \times MX_k - \sum_{i \in L} f_{k,i} \times B_{k,i} \;\; \forall k \tag{5.3.12}$$

where,

$Z_{AP} = C_A$ - R_A actual assignment profit (cost-revenue);

$Z_{VP} = C_V$ - R_V virtual assignment profit (cost-revenue);

$C_{U1} = \sum_{k \in K} rp_k \times w_k$ cost of aircraft under-utilization (or idleness);

$C_{U2} = \sum_{k \in K} r_k \times w_k$ cost of aircraft utilization (or idleness);

$C_A = \sum_{k \in K} \sum_{i \in L} (c_{k,i} + j_{k,i}) \times f_{k,i}$ actual cost of assignment;

$R_A = \sum_{s \in S} p_s \sum_{k \in K} \sum_{i \in L} X_{k,i,s} \times q_i \times f_{k,i}$ expected actual assignment revenues;

$C_V = \sum_{s \in S} p_s \sum_{k \in K} \sum_{i \in L} (c_{k,i} + j_{k,i}) \times b_{k,i,s}$ expected virtual cost of assignment; and

$R_V = \sum_{s \in S} p_s \sum_{k \in K} \sum_{i \in L} X_{k,i,s} \times q_i \times b_{k,i,s}$ expected virtual buffer revenues.

Sets and indices

L: the set of flight legs indexed by i;

K: set of different aircraft types indexed by k;

O: set of stations (airports) indexed by o;

T: the sorted set of all event (arrival or departure) times at all airports, indexed by t. The event at time t occurs before the event at time, $t + 1$. The last node in the counting line is denoted by t_n. The index t^+ denotes an event after the current time t while t^- denotes an event before the current time t ($t \in T$);

R: the set of nodes in the network indexed by $\{k,o,t\}$;

CL(k) denotes the set of flight legs crossing the count time (i.e. t_n) flown by k i.e.

the set of flight legs where an aircraft of type k may be in the air at time t_n;
I(k,o,t) is the set of inbound flight legs to node $\{k,o,t\}$; and
O(k,o,t) is the set of outbound flight legs from node $\{k,o,t\}$.

Decision variables

$$f_{k,i} = \begin{cases} 1 & \text{if fleet type } k \text{ is assigned to flight leg } i \text{ ;} \\ 0 & \text{otherwise.} \end{cases}$$

$Y_{k,o,t} \geq 0$ is the number of aircraft of fleet type k, on the ground at station o, and time t; Y_{k,o,t^+} is the number of aircraft of fleet type k, on the ground at station o, just following time t $(t \in T)$; and Y_{k,o,t^-} is the number of aircraft of fleet type k, on the ground at station o, just prior to time t $(t \in T)$;
$b_{k,i,s}$ is an artificial (virtual assigner) non-integer variable ≥ 0 (aircraft);
r_k is the number of utilized block-time of aircraft of fleet type $k \geq 0$ (minutes or hours);
rp_k is the number of under-utilized block-time of aircraft of fleet type $k \geq 0$ (minutes or hours); and
rm_k is the number of expected-utilized block-time of aircraft of fleet type $k \geq 0$ (minutes or hours).

Parameters

$c_{k,i}$ denotes aircraft *direct operating cost* of aircraft k on flight leg i (£ per aircraft flight leg);
N_k denotes the number of available aircraft of type k (including standby aircraft);
$seats_k$ denotes the number of seats on aircraft k;
$d_{i,s}$ denotes passenger demand on flight leg i under scenario s;
u_k denotes the maximum spill factor on aircraft k;
$D_{k,i,s}$ denotes the maximum of the number of aircraft seats of type k or passenger demand after spill on aircraft k for flight i under scenario s, that is,
$D_{k,i,s} := \max\,(seats_k,\, d_{i,s} \times (1 - u_k))$;
$X_{k,i,s}$ denotes the minimum of the number of aircraft seats of type k or passenger demand after spill on aircraft k for flight i under scenario s), that is,
$X_{k,i,s} := \min\,(seats_k,\, d_{i,s} \times (1 - u_k))$;
$j_{k,i}$ denotes penalty cost of using standby aircraft on flight leg i (£ per aircraft flight

leg);

q_i denotes average ticket price on flight leg i (£ per aircraft);

w_k denotes the utilization cost of the scheduled aircraft k i.e. standby aircraft do not attract utilization cost (£ per aircraft minutes or hours);

p_s denotes the probability of scenario s;

MN_k denotes the average minimum block-time threshold value that aircraft of type k can fly (minutes or hours);

MX_k denotes the average maximum block-time threshold value that aircraft of type k can fly (minutes or hours); and

$B_{k,i}$ denotes the block-time of aircraft of type k on flight leg i (minutes or hours).

Note that $f_{k,i}$ corresponds to the first-stage solution that gives overall fleet composition for all the scenarios while $b_{k,i,s}$ corresponds to the second-stage solution unique for each of the scenario. The *objective* functions (5.3.1 and 5.3.7) minimizes the fleet assignment cost (it is a re-expression of maximization of fleet assignment profitability). The *cover* constraints (5.3.2 and 5.3.8) ensures that each flight is covered once and only once by a fleet type. Equation 5.3.3 and 5.3.9 are the *conservation flows* constraints that ensures aircraft balance, that is, all aircraft going into a station must leave the station at some time. The *count* constraints (5.3.4 and 5.3.10) ensures that only the number of available aircraft are used. The *demand-spill* constraints (5.3.5 and 5.3.11) ensures that passenger demand after spill does not exceed the maximum of the number of aircraft seats of type k or passenger demand after spill on aircraft k for flight i under scenario s. The aircraft *utilization* constraints (5.3.6 and 5.3.12) computes the number of utilized block-time on aircraft of type k.

5.4 Model Features

5.4.1 The Standby Aircraft

The fleet assignment problem is basically a multi-commodity flow problem with aircraft being commodities that need to be assigned (or produced) to satisfy passenger demand. As such, we can envisage the multi-commodity flow problem under conditions of certain passenger demand and uncertain passenger demand. Under certainty conditions, the demand needs to be satisfied with the available number of aircraft. In uncertainty conditions, demand needs to be satisfied with an alternative aircraft type that was not originally scheduled to operate the flight. This can be done by ferrying an aircraft, leasing of an aircraft at a short notice or use of a standby aircraft.

As such the distinct feature of the SP formulation is the standby aircraft that is reassigned in the case of demand fluctuation but at the cost of the original scheduled aircraft. The standby aircraft concept is used widely by both charter operators (Ronen, 2000) and scheduled operators. Ideally, rather than having a standby aircraft waiting to be assigned in the presence of demand uncertainty, a robust schedule is built in such a way that the aircraft in circulation (on the network) become a potential standby when on ground and not immediately scheduled. One way of accounting for standby aircraft types is depicted in the development of the scheduling wave system where the aircraft on ground measure (Bian et al., 2005) becomes an important attribute in the test of schedule robustness. Note that by swapping an originally scheduled aircraft with a standby, the former becomes redundant and under-utilized, a situation not welcomed by network planners. To counter this concept and discourage its practice, we introduce a penalty cost for using the standby aircraft and an utilization (idle) cost for making the original scheduled aircraft under-utilized.

5.4.2 The Objective Function

The objective function computes the *Total Profit* (Z_{TP}) that comprises the *Real Profit* (Z_{RP}) and the *Virtual Profit* (Z_{VP}). While Z_{RP} can easily be measured and tracked in the financial books of accounting, the Z_{VP} is false and cannot be measured. In other words, the airline does not really receive this profit. Both Z_{RP} and Z_{VP} have revenue and cost elements. The revenue component computes the passenger revenue by multiplying the projected passenger number with average fare (i.e. $X_{k,i,s} \times q_i$). The cost element has two components, the aircraft *direct operating cost* ($c_{k,i}$) and the standby aircraft penalty cost ($j_{k,i}$). $c_{k,i}$ is incurred by both scheduled and standby aircraft while $j_{k,i}$ is exclusive for the standby aircraft. The Z_{RP} is associated with the first-stage decision variable ($f_{k,i}$) that comprises the Utilization Cost (C_{U1} or C_{U2}) and Z_{AP}; while the Z_{VP} is associated with the second-stage decision variable ($b_{k,i,s}$).

- $Z_{AP} = C_A$ - R_A
- $Z_{RP} = C_{U1}$ (or C_{U2}) + Z_{AP}

- $Z_{VP} = C_V$ - R_V
- $Z_{TP} = Z_{RP} + Z_{VP}$

5.4.3 The Artificial Variable

The $b_{k,i,s}$ tries to assign superficial aircraft taking into account both revenue and the *aircraft direct operating cost*. In particular, the artificial variable is introduced to satisfy recourse condition but does not violate or contradict the first-stage decision variable. As such, the aircraft utilization cost (C_{U1} or C_{U2}) and the virtual profit (i.e. Z_{VP}) that are tied to the aircraft swaps are the main underlying logics of the formulation. The major problem with FAM solutions, and especially for a large network, is the generation of infeasible solutions. If $b_{k,i,s}$ is treated as a binary variable, the model will yield a feasible optimal solution for a model of smaller size but infeasible for a large size. As in real life the artificial variable does not have much meaning, to generate a feasible optimal solution for a large network, $b_{k,i,s}$ is relaxed (i.e. ≥ 0).

In a typical multi-commodity network flow problem, $b_{k,i,s}$ would have corresponded to a standby aircraft (instead of Z_{VP}). But this assumption is only valid if the original scheduled aircraft are capable of being entirely removed or added from circulation (as opposed to being under-utilized or over-utilized) and where there is a wide variants of standby aircraft. In this case, we visualize a situation where to satisfy the uncertain demand, assign (or produce) $f_{k,i}$ of the scheduled aircraft; and where there is a surplus or shortage in demand, acquire (or hire) $b_{k,i,s}$ of the standby aircraft. However, this is not possible for the fleet assignment problem since the scheduled aircraft are fixed and the standby aircraft are limited.

5.4.4 Demand-Spill Constraint

The demand-spill constraint (5.3.5 or 5.3.11) links the first-stage decision variable, $f_{k,i}$, and the second-stage artificial variable, $b_{k,i,s}$. The constraint tries to ensure that the assigned aircraft does not exceed the maximum of the assigned aircraft capacity or passenger demand after spill. If eliminated, the constraint will lead to unboundedness irrespective of changing any of the cost or revenue parameters. Similarly, an unbounded value is obtained if the equation sign is changed to $\geq$. In understanding the constraint we make an illustration with the following assumptions:

• Two routes denoted by X-Y1-X and X-Y2-X are to be operated. Where X-Y1-X means a complete rotation, with an outbound flight X-Y1 and an inbound flight Y1-X. The outbound flight departs from station X and arrives at station Y1, while the inbound flight departs from station Y1 and arrives back at station X.

• Three aircraft denoted by A1, A2, and B4 (where B4 is a standby) are available to operate any of the two routes. In algebraic notation, this is expressed as $N_{A1} = 1$, $N_{A2} = 1$ and $N_{B4} = 1$.

• The number of seats on each aircraft are given as 100, 120 and 150 seats for A1, A2 and B4, respectively.

• Two passenger demand scenarios (scenario 1 and scenario 2) are given for each flight leg (as shown in *Table 5.1*).

• The probability for scenario 1 is 0.7 while that of scenario 2 is 0.3.

• The maximum spill factor for each aircraft is 0.1.

• For clarity, the artificial second-stage decision variable ($b_{k,i,s}$) is treated as binary (instead of being relaxed).

Flight leg i	*Scenario 1*	*Scenario 2*
X-Y1	100	120
Y1-X	150	80
X-Y2	200	100
Y2-X	130	70

Table 5.1: Passenger demand scenarios.

Table 5.2 depicts typical results of the demand-spill constraint. After the optimization process, assume the fleet assignment is as shown in column 2. The computation shows that each aircraft is capable of being assigned; there is no preference in assigning a bigger aircraft to a smaller one or a smaller one to a bigger one. The main underlying logic of the constraint is that if a standby aircraft (i.e. B4) is assigned to a particular flight leg, in our case $f_{A1,i}$= 1 and $f_{B4,i}$= 1, then the buffer variables associated with the assigned aircrafts are automatically set to zero (i.e. $b_{A1,i,s}$=0 and $b_{B4,i,s}$=0), for both scenario 1 and 2. The other values for $b_{k,i,s}$ will be allocated to the un-assigned aircraft flight legs, taking a value of 0 or 1 in one or both scenarios.

Constraint	$f_{k,i}$	$b_{k,i,s}$	$seats_k$	$(f_{k,i} + b_{k,i,s}) \times seats_k$	$d_{k,i} \times u_k$	$D_{k,i,s}$
A1:Sc1: X-Y1	1	0	100	100	90	100
A1:Sc2: X-Y1	1	0	100	100	108	108
A1:Sc1: Y1-X	1	0	100	100	135	135
A1:Sc2: Y1-X	1	0	100	100	72	100
A1:Sc1: X-Y2	0	{0,1}	100	{0,1} × 100	180	180
A1:Sc2: X-Y2	0	{0,1}	100	{0,1} × 100	90	100
A1:Sc1: Y2-X	0	{0,1}	100	{0,1} × 100	130	130
A1:Sc2: Y2-X	0	{0,1}	100	{0,1} × 100	117	117
A2:Sc1: X-Y1	0	{0,1}	120	{0,1} × 120	90	120
A2:Sc2: X-Y1	0	{0,1}	120	{0,1} × 120	108	120
A2:Sc1: Y1-X	0	{0,1}	120	{0,1} × 120	135	135
A2:Sc2: Y1-X	0	{0,1}	120	{0,1} × 120	72	120
A2:Sc1: X-Y2	0	{0,1}	120	{0,1} × 120	180	180
A2:Sc2: X-Y2	0	{0,1}	120	{0,1} × 120	90	120
A2:Sc1: Y2-X	0	{0,1}	120	{0,1} × 120	130	130
A2:Sc2: Y2-X	0	{0,1}	120	{0,1} × 120	117	120
B4:Sc1: X-Y1	0	{0,1}	150	{0,1} × 150	90	150
B4:Sc2: X-Y1	0	{0,1}	150	{0,1} × 150	108	150
B4:Sc1: Y1-X	0	{0,1}	150	{0,1} × 150	135	150
B4:Sc2: Y1-X	0	{0,1}	150	{0,1} × 150	72	150
B4:Sc1: X-Y2	1	0	150	150	180	180
B4:Sc2: X-Y2	1	0	150	150	90	150
B4:Sc1: Y2-X	1	0	150	150	130	150
B4:Sc2: Y2-X	1	0	150	150	117	150

Table 5.2: Demand-spill constraint.

The allocation of $b_{k,i,s}$ to non-existence flight legs can have a deceptive interpretation where the Z_{VP} value appears to be much higher than the Z_{RP} and Z_{AP} values. Thus, when subjecting Z_{TP} to the stochastic measures, we expect the superiority of the here-and-now model (Z_{hn}) and wait-and-see model (Z_{ws}) to be much higher than the expected-value model (Z_{ev}). Consequently, it would only make sense when the tests are also applied to Z_{RP} and Z_{AP} values. Theoretically, stochastic measures are only applied to the overall objective function (Z_{TP}) since, in our case, the expected value model (Z_{ev}) would have the same value for Z_{TP}, Z_{RP} and even Z_{VP}. However as we are more concerned with EVPI and VSS values, where Z_{ev} does not play a direct role in establishing the inter-bound relationship (i.e. $Z_{ws} \leqslant Z_{hn} \leqslant Z_{eev}$), its reporting only serves as a benchmark and the extension of the stochastic measures to Z_{RP} and Z_{AP} becomes a valid assumption. Since the model logic seeks to trade-off the Z_{RP} with Z_{VP} value, the optimizer will always try and set the value of $b_{k,i,s}$ to 1. Consequently, while the inter-bound relationship will always hold for Z_{RP}, it will do so at the expense of Z_{VP} value (i.e. it will never hold for Z_{VP} value). As the allocation of $b_{k,i,s}$ happens only to virtual flights that have not been assigned to $f_{k,i}$, it does not pose any implementation impediment from the scheduling point of view.

The Z_{VP} paradigm will still not stop airline enthusiast from questioning the main motive behind such an assumption, does the false profit not jeopardize the airline business? In response, we make crystal clear that the Z_{VP} drives the fleeting decisions (our primary concern) and if Z_{RP} is measured distinctly, the monetary benefit will become more apparent. As a simple illustration, assume on solving the SP-FAM, the following values are obtained:

- Z_{AP} = £4 million
- Z_{RP} = £3 million (where C_{U1} or C_{U2}= £1 million)
- Z_{VP} = £17 million
- Z_{TP} = £20 million

The Z_{AP} (£4 million) and Z_{RP} (£3 million) values represent the standard measures that can easily be tracked in the financial books of accounting. However, the Z_{VP} (£17 million) value cannot be trailed and purely serve as a driver to the fleeting decision (i.e. $f_{k,i}$). If the Z_{VP} value did not exist or gets eliminated, constraint 5.3.5 (or 5.3.11) will become redundant, and consequently, the SP-FAM will be reduced to DFAM but with average passenger demand (i.e. the Expected Value model, Z_{ev}). Further, the standby aircraft that was originally in the SP-FAM will still remain intact in the *reduced* DFAM. Ideally the standby aircraft should not be included when solving the DFAM, since in a deterministic environment, we assume that the schedule will be flown as planned. In this case, the original Z_{RP} value (£3 million) will be different from both the *reduced* DFAM and DFAM solved without the standby. The 3 million value will only be identical to the *reduced* DFAM if the resultant fleeting decisions (i.e. $f_{k,i}$) remains intact. But this is rare for a large network, as the Z_{VP} (£17 million) value plays a significant role in determining the resultant fleeting decisions for the SP-FAM. The backbone of our research is to account for the uncertain demand, but the DFAM does not only integrate the uncertain demand but will be proved (in *Chapter 8*) to be relatively inferior to the SP-FAM.

In many optimization problems, we either include high penalty cost to restrict certain variables or introduce dummy variables to satisfy certain conditions. If such a practice is universally accepted within the optimization world, there is little room for doubting the Z_{VP} concept. In the context of schedule robustness and as discussed in section 2.4, a more realistic analogue is given by Love et al. (2002). To tackle the disruptive situation, three strategies were deployed, that is, delaying a flight, cancelling a flight or swap a scheduled aircraft with a standby. In modelling the objective function, just like in the case of our SP-FAM, both the *real* and the *virtual* cost (profit) were minimized.

5.4.5 Aircraft Utilization Constraint

The idea of the cost function C_{U1} and constraint 5.3.6 for the SP-FAM1 is to incur a utilization cost whenever utilization falls below the MN_k value (i.e. incur penalty, w_k, when there is under-utilization, rp_k). However, in our *recourse* argument we wish to incur the utilization cost (w_k) whenever there is a standby aircraft swap, irrespective of under-utilization or expected-utilization (acceptable-utilization). This implies that if the original scheduled aircraft was initially utilized as expected (in which case w_k=0) and later on there was a standby aircraft swap, there is no guarantee that the utilization would fall below the MN_k value to attract the penalty. In other words, the the cost would only be incurred if the swapped aircraft was initially under-utilized (or had a utilization value below MN_k).

Constraint 5.3.12 and the utilization cost (i.e. C_{U2}) that corresponds to SP-FAM2 tries to overcome the inherent limitation. The constraint has MX_k as opposed to MN_k value and the utilization cost will be incurred whenever there is a standby aircraft swap. Although the two approaches (i.e. SP-FAM1 and SP-FAM2) might yield different fleet assignment decisions, it is expected that one will out-perform the other. In deducing the superiority of one over the other, the Z_{AP} value will be an important measuring criteria.

As an illustration, assume we now have a fleet size of six denoted by A1, A2, three A3s and B4; where B4 is a standby. Five aircraft are to be assigned to five routes X-Y1-X, X-Y2-X, X-Y3-X, X-Y4-X and X-Y5-X. Table 5.3 is in reference to

SP-FAM1 while Table 5.4 corresponds to *SP-FAM2*. After the optimization process, and for simplicity, suppose the same fleet assignment decisions took place for both approaches (indicated by column **(a)**). Column **(b)** indicates the flight leg block-time; **(c)** the average minimum threshold value for SP-FAM1 and the average maximum threshold value for SP-FAM2. Column **(d)** gives the constraint computation denoted by variable rp_k and r_k for each approach.

(a) $f_{k,i}$	**(b)** $B_{k,i}$	**(c)** MN_k	**(d)** rp_k	**(e)** w_k, UK £	**(f)** $rp_k \times w_k$, UK £
$f_{A1,X-Y5} = 1$ $f_{A1,Y5-X} = 1$	2 2	5	1	1000	1000
$f_{A1,X-Y2} = 1$ $f_{A1,Y2-X} = 1$	4 4	7	0	2000	0
$f_{A3,X-Y3} = 1$ $f_{A3,Y3-X} = 1$ $f_{A3,X-Y4} = 1$ $f_{A3,Y4-X} = 1$	6 6 5 5	10	8	1500	12000
$f_{B4,X-Y1} = 1$ $f_{B4,Y1-X} = 1$	4 4	8	0	0	0

Table 5.3: Constraint 5.3.6 and function C_{U1}.

(a) $f_{k,i}$	**(b)** $B_{k,i}$	**(c)** MX_k	**(d)** r_k	**(e)** w_k, UK £	**(f)** $r_k \times w_k$, UK £
$f_{A1,X-Y5} = 1$ $f_{A1,Y5-X} = 1$	2 2	7	3	1000	3000
$f_{A1,X-Y2} = 1$ $f_{A1,Y2-X} = 1$	4 4	9	1	2000	2000
$f_{A3,X-Y3} = 1$ $f_{A3,Y3-X} = 1$ $f_{A3,X-Y4} = 1$ $f_{A3,Y4-X} = 1$	6 6 5 5	12	14	1500	21000
$f_{B4,X-Y1} = 1$ $f_{B4,Y1-X} = 1$	4 4	14	6	0	0

Table 5.4: Constraint 5.3.12 and function C_{U2}.

5.4.6 Aircraft Utilization Cost

If we further assume that the utilization variable has a corresponding utilization cost as shown in column **(e)** of Table 5.3, the computation of the utilization cost function C_{U1} and C_{U2} will then be computed as shown in column **(f)**. Note that one of the A3s is idle and incurring a high under-utilization cost, relatively to other aircraft. If there was no standby aircraft swap, the resultant utilization cost would have been much lower.

The aircraft utilization cost, as an input to the SP-FAM models, can be derived in several ways. For instance, in the fleet planning process aircraft can be acquired in different ways; from an outright cash flow, borrowing from a financial institution, raising debt from capital market or on a lease basis. Acquisition through *operating lease* method is the most commonly used. Under this method, the lessor (owner) gives rights to the lessee (user), for a given period of time, to operate the lessor's equipment in exchange for an obligation to pay rent without transferring ownership to the lessee. The monthly lease payment can be one way of determining the utilization cost. Normally, and this applies to any mode of financing, a variable aircraft ownership cost is allocated in the route direct operating costs. The allocation is done by taking the aircraft monthly lease rent and apportioning it over the block hours.

As a strategy in deriving the utilization cost (w_k), it is preferred to solve the deterministic equivalent model first. The initial w_k values could be set, for instance, based on the average lease payment cascaded down on hourly (or minutes) rate. Thereafter, the values are adjusted further based on the desired utilization balance required for each aircraft type. The stabilized values could then be used in solving the SP-FAM problem.

5.4.7 The Standby Aircraft Penalty Cost

The aircraft penalty cost is associated with the standby aircraft that is substituted for an originally scheduled aircraft. Even in normal circumstances of assigning aircraft to flights, the practice is to discourage certain possible swaps that interfere with smooth planning. The swapping affects several processes including; aircraft maintenance activity that needs to be rescheduled; revenue management that needs to account for passenger offload when the swapped aircraft has lower capacity (seats); and even more important, we could have compatibility issues that go with the crew constraints. For instance, it could happen that the standby aircraft is not ideal to operate a particular affected flight or no crew are available to operate the aircraft. As such, deriving the penalty cost when the potential standby aircraft are either the assigned aircraft in circulation or non-assigned, is not difficult. Different airlines have different conventions in calculating the penalty cost. One way would be to penalize whenever there is a swap on a specific aircraft type (i.e. j_k), but the most ideal is to penalize the swaps based on aircraft block-time or aircraft type per flight leg (i.e. $j_{k,i}$). This way, a swap on a longer flight will incur more penalty than on a shorter flight.

5.5 Conclusion

The chapter proposed the SP-FAMs that effectively integrates the uncertain demand, the core of the research. In our approach, the DFAM was extended by introducing an *artificial* second-stage decision variable (b_{kis}), standby aircraft, demand scenarios, aircraft standby cost, scheduled aircraft idle cost and demand-spill constraint. For the recourse condition, the first-stage decision variable ($f_{k,i}$) corresponds to real profit (Z_{RP}) while the second-stage decision variable (b_{kis}) corresponds to the virtual profit (Z_{VP}). For clarity, prototype examples are given to illustrate the models' salient features.

Chapter 6

Scenario Generator

Stochastic demand remains a major challenge in the quest to produce optimal schedules. In the context of integrating the FAM with uncertain demand, there have been few attempts to come up with a realistic model that generates realistic scenarios. Such models tend to be simplistic with the use of descriptive statistics and applying an average of the demand scenarios in solving the FAM. As such, the models have not only been unpopular but enjoyed less industrial application. The problem is attacked through a sophisticated network-simulation model (also called network-planning model). Conceptually, network-simulation models refer to a collection of models that are used to determine how many passengers want to fly, what flight or sequence of flights they choose, and the profitability of transporting passengers on their chosen flights. The network-simulation model gives a snapshot of one week that is representative of any other week of the scheduling season. In this chapter, the previous attempts are reviewed, the description of the network-simulation model process is briefly illustrated and extended by having each week accounted for independently. This is done through an integration of a *time-series* demand generator to the network-simulation model. This approach establishes a forecasting process that can be used as a scenario generator when solving the FAM formulated as a two-stage SP with recourse.

6.1 Integrating Scenario Generator and SP Model

A mathematical programming problem in which some of the data are unknown, that is, they are subject to uncertainty, random influences, or statistical variations is called an SP problem. SP provides a general framework to model path dependence of the stochastic process within an optimization model. Furthermore, it permits innumerable states and actions, together with constraints, time-lags, etc. Unlike dynamic programming, SP separates the model formulation activity from the solution algorithm. One advantage of this separation is that it is not necessary for the SP models to obey the same mathematical assumptions. This leads to a rich class of models for which a variety of algorithms can be developed. SP formulations, however, can lead to very large-scale problems. This requires the development of efficient solution methods in order to process progressively larger models (Poojari et al., 2007). When SP is used for tackling any problem, models of *optimum allocation* and of *randomness* are solved. Whereas the optimum decision model constitutes the core of the problem, the randomness model determines the distribution of the stochastic parameters with underlying probability distribution for different scenarios. Hence, the first component of SP is essential for the realistic representation of the problem and the second component leads to better quality decisions being taken. Thus, a major issue in any application of SP is the representation of the underlying random data process.

6.1.1 Scenario Trees

The major challenge in an SP problem is modelling data to correspond to the algebraic structure of the model; in essence, the two are inter related. The advantage of this relationship is that the modeler is able to analyze the data and solution values using the same structure. Two-stage SP with recourse is nothing more than a large linear program where realization of random parameters are included explicitly. Let ξ represent the vector of all random parameters in a model. The probability of each realization is defined as: p_k=P(ξ_k) for k=1,...,K where $p_k \geq 0$ and $\sum_k p_k$=1.

For multi-stage SP one assumes that the random vector ξ follows a stochastic process ξ_t over the planning horizon. If the process is assumed to be discrete, with probability, P(ξ_t), the uncertainty can be represented through a multilevel event tree which defines the possible sequence of realizations also known as data paths. In general these are called scenarios, over the whole planning horizon (see Figure 6.1).

The random data that represent the uncertainty is called the *scenario tree* and encapsulates the first and the second stage phases. A root node represents the first stage and extends linearly until the end of the stage time period. At the second stage, the tree branches into nodes at level t=k+1 as shown in Figure 6.2. From

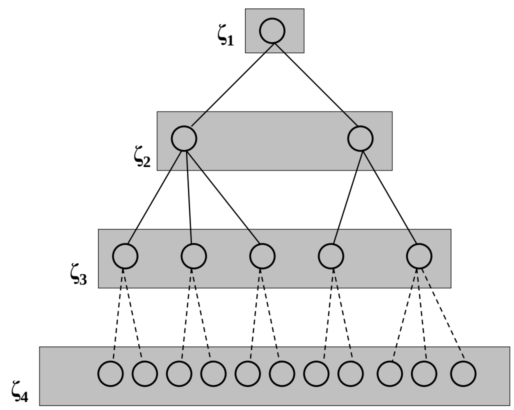


Figure 6.1: Multi-stage tree.

each of these nodes discrete flat scenarios commence with nodes at each time period, an optimum decision has to be taken until level t=T. This means that the scenario tree is effectively a fan of individual scenarios $\omega_s = \omega_{1,s},...., \omega_{T,s}$ which occur with probabilities p_s=P($\omega_{k,s}$) ∀ s.

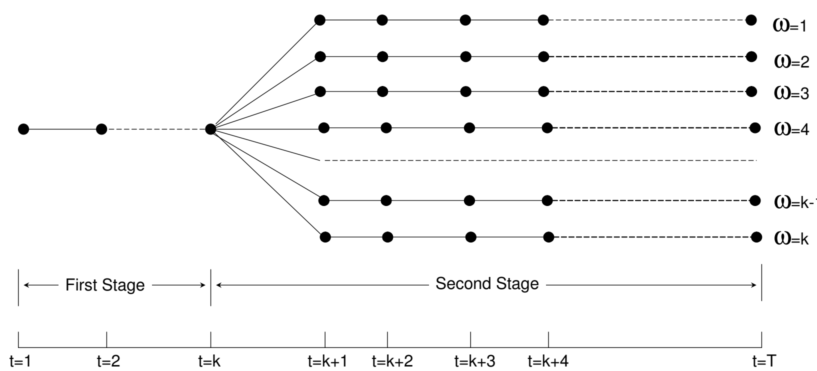


Figure 6.2: Scenario tree for two-stage SP.

6.1.2 Scenario Generation

As outlined above, the data for a stochastic optimization model is provided in the form of *scenario trees* that are created using scenario generation methods which may be very specific to the domain of application. The major challenge is the creation of a close to ideal tree-structure that approximates the underlying distribution of the random parameters. Some defined criteria include a model that explains the behavior of the random parameters and the corresponding estimation of those parameters.

Another criteria includes discretization of the distributions using statistical properties and sampling of the scenarios. Table 6.1 shows the most common techniques that have been used to develop scenario generators (Domenica et al., 2007) using some of the mentioned criteria. For our scenario-generator, we use one of these technique (i.e. time series analysis).

6.1.3 Measure of Quality of a Scenario Tree

Scenario generation methods differ in their ability to model randomness. Kaut and Wallace (2007) discussed how to assess the quality of scenario generation and argue that it should be performance based rather than based on theoretical properties. They outlined two minimal requirements a scenario-generation method must satisfy. Since most of the methods involve some randomness, the first requirement is *stability*: if we generate several trees (with the same input) and solve the optimization problem with these trees, we should get the same optimal value of the objective function.

The other requirement is that the scenario tree should not introduce any *bias*, compared to the true solution. In testing for bias, we wish to determine if the scenario generation method itself introduces any bias producing the optimal solution. This can be theoretically achieved by comparing the optimal solution of the SP model to the optimal solution obtained from the *true* statistical process. Practically this is not applied as solving the theoretically optimal solution would imply the SP model is redundant.

6.2 Scenario Generation for FAM

The uncertainty of demand for a future flight departure can be represented with a probability distribution (density) of expected demand. Historically, a Gaussian (Normal) distribution of demand has been assumed, with a mean and standard deviation that depend on the market being studied. Based on many empirical studies of actual airline data (Belobaba, 2006), the standard deviation of total demand for a flight relative to the mean demand is typically between 0.20 and 0.40.

In the work done by (Lister and Dekker, 2002), selection of demand realization and their mutual combination for the scenario generator was done using a descriptive

Purpose	**Methods**
Generation of data trajectories	Econometric Models and Time Series ∗ Autoregressive Models: AR(p) ∗ Moving Average Models: MA(q) ∗ Autoregressive MA Models: ARMA(p,q) ∗ Generalized Autoregressive Conditional Heteroscedasticity: GARCH(p,q) ∗ Vector Auto Regressive Models: VAR ∗ Bayesian VAR ∗ Reduced Rank Regression Diffusion Process ∗ Wiener Processes (Brownian Motion) ∗ Generalized Wiener Process Other Methods ∗ Neural Networks
Discretization	Statistical Approximation ∗ Property Matching ∗ Moment Matching ∗ Non-parametric methods Sampling ∗ Random sampling ∗ Stratified sampling ∗ Bootstrapping
Tree construction and Conditional Sampling	Optimal discretization ∗ Optimal Discretisation ∗ Barycentric Approximation ∗ Sequential Clustering
Reduction	∗ Scenario Reduction
Internal Sampling	∗ Stochastic Decomposition ∗ Stochastic Quasi-gradient ∗ EVPI-based Importance Sampling

Table 6.1: Techniques used in scenario generation

sampling method. For illustrative purpose, they assumed one payload class is available in each aircraft type. The demand for seats on each flight leg i = 1, 2,....,N is assumed to follow a normal distribution, d_i $N(\mu_i; \sigma_i)$, with probability distribution function, F_i (the demands are assumed to be independent). They specified the number of scenarios (S) to be generated with values, d_i[1], d_i[2],..., d_i[S] sampled from distribution i and equally quartile spaced;

$$d_i[j] = F_i^{-1}(\frac{j - 0.5}{S}),\ j = 1, 2, ..., S;\ i = 1, 2, ..., N \tag{6.2.1}$$

In this way more sample values were generated from a range of higher distribution density and less values from low density regions. Since the inverse of the distribution function F_i is not available analytically, they used accurate numerical approximations generated with the Newton-Raphson method. Subsequently, a random permutation of the values d_i[j]; j = 1, 2,....,S, was generated for each i = 1, 2,....,N. Then each of the scenario representative vectors, $(d_1[j], d_2[j], ..., d_N$[j]), j = 1, 2,.....,S was assigned a probability, $\frac{1}{S}$. Thus the sample variability by a random combination of the S values of each distribution was maintained with each other.

Another scenario generator approach presented was a Design and Analysis of Computer Experiment (DACE) (Pilla et al., 2005). Typically, the computer experiment is a simulation model; however, in this case, it was an optimization model that was constructed based on knowledge of how the system operated. Design of Experiments (DoE) was then used to select a set of sample points as input to the optimization model, which then provided the corresponding responses with a fitting of a statistical model into the data.

6.3 Calibration Process and Network-Simulation Model

The simulated-network tool imitates a real-life operating environment with competitive forces where planners forecast profitability under different scheduling scenarios such as entry into a new market, code-share with a potential partner, change in flight timing, hub restructuring, budgeting for the airline, etc.

A vigorous calibration process, that is done outside the simulated-network model, generates connection-builder (CB) parameters, beta-parameters for market share model (MSM), market size estimates and cluster structures for both CB and MSM which then become part of the input to the simulated-network model. Figure 6.3, depicts a typical relationship between the two.

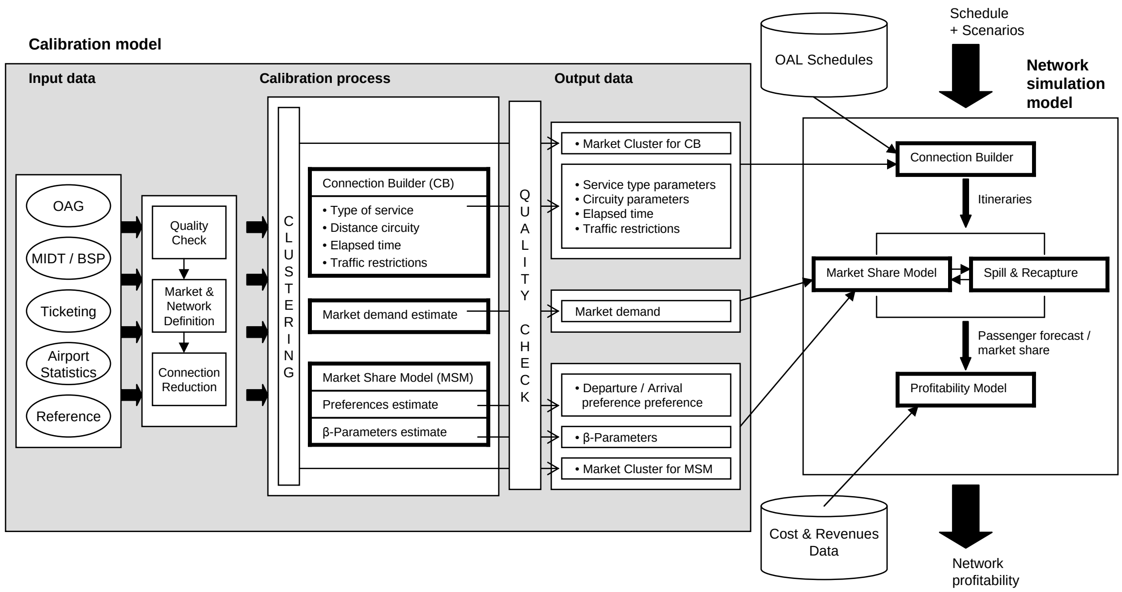


Figure 6.3: Calibration process and network-simulation model.

The calibration process has four major stages, that of market clustering, market demand estimation, generation of beta-parameters for both the CB and MSM. The main distinction between the CB and MSM for the *calibration model* to that of the *network-simulation model* is that the latter takes inputs from the former to generate itineraries for the CB and passenger forecast for the MSM. The major processes can be summarized as follows:

Input data and clustering: The input data that goes into the calibration tool includes OAG, global passenger data, coupon uplift, airport statistics, among other. Once the data has been input, initial data processing takes place followed by quality checks, market definition and connection reduction. A clustering process follows suits where group definition of Origin-Destination is defined with respect to station, city, country, sub-region and region.

Connection Builder: The CB calibration sets parameter values from MIDT by creating relevant itineraries (sequence of flights) from OAG data (schedule). In so doing, realistic connections linking airports are created. Although different airlines use different logic to generate the itineraries, the underlying algorithms are similar. Some of these algorithms include the service type, distance-dependent circuit, elapsed time, and traffic restrictions.

Market demand estimation: The seasonal MIDT data is mapped on a weekly basis using ticketing (coupon uplift) and airport statistics data. Ratios deduced through matching ticketing and airport data to MIDT scales the MIDT with the appropriate correction factors. A weekly average is then computed and a proportion based on either weighted mean by distance or passenger number is applied to the segment scaled MIDT data to derive the O-D levels.

Market share model: The MSM estimates the probability of a traveller selecting a specific itinerary (or path i.e., a sequence of flight segments a passenger can use to make a continuous trip from an origin to a destination) connecting an airport. The earliest market share model employed a demand allocation methodology referred to as *Quality of Service Index* (QSI). The QSI model assigns points (weights) to path (itinerary) based on its attributes. The *Multinomial Logit Model* (MNL) is the refinement of the QSI model and uses multiplicative weighting.

The spill and recapture model: Spill refers to passenger demand in excess of aircraft capacity. The *spill model* has extensively been elaborated (Swan, 1999) and further enhanced (Swan, 2001) into what is known as *the revised spill model*. The relationship between spill and recapture has been elaborated (Barnhart et al., 2002) on the *Passenger Mix Model*. Since demand for certain flights may exceed the available capacity, spill and recapture models are used to reallocate passengers from full flights to flights that have not exceeded capacity.

Profitability model: Finally, revenue and cost allocation models are used to determine the profitability of an entire schedule (or a specific flight).

6.4 The Scenario Generator

Calibration is done for a previous period (the world as it was). Thereafter, schedule and market sizes are adjusted to reflect the current and future period. The schedule adjustment is relatively straight-forward as this entails updating OAG schedule and the airline's own schedule. However, the market size adjustment is subjective since the calibrated week represents an ideal week for a whole planning horizon. One way of refining the market adjustment is to forecast by applying IATA published continental growth factors (IATA, 2007), however, this breeds subjectivity as the factors neither reflect O-Ds growth rates nor capture seasonal fluctuations. Our aim is to determine close to actual demand forecast for each schedule week. The tactical FAM requires a representative demand for a typical week while SP-FAM requires demand for all the weeks in the scheduling season; by forecasting demand for every week and applying equal probability, it represents the underlying concept of our scenario generator.

6.4.1 Holt-Winter's Additive Time Series Model

To achieve this objective, the first step is to derive a weekly Time Series Model (TSM) from historical MIDT, with data spanning a period no less than three years. In particular, develop a trend equation and seasonal index. Using the trend equation, forecast the *market sizes* and determine the annual growth rates. Using the calibrated *market sizes*, apply the deduced growth factors and seasonal index to predict *market sizes* for each particular week. Finally map the planned schedule with the corresponding forecast *market sizes* to the network-simulation model to predict the demand for each week. The TSM was coded in SAS and executed stepwise as shown at Figure 6.4.

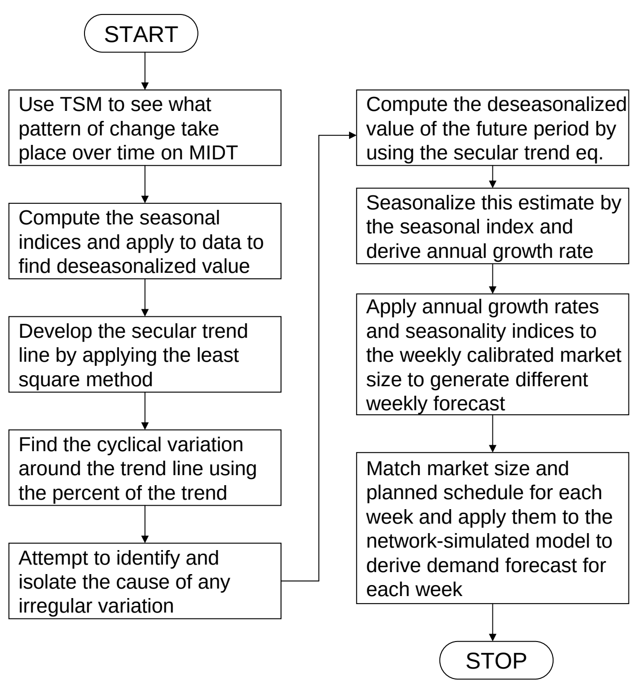


Figure 6.4: TSM flow chart.

As an illustration and for simplicity, suppose the MIDT data for an O-D (given in quarters as opposed to weeks) is as shown on Table 6.2. The objective is to determine the annual growth rate for 2008.

YEAR	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
2003	160	210	90	180	640
2004	150	200	100	180	630
2005	170	240	130	220	760
2006	170	250	110	210	740
2007	180	260	140	250	830

Table 6.2: Marketing Information Data Tape

The first step is to compute the seasonal indices by deseasonalizing the time series on a quarterly basis, as shown in Table 6.3 and then compute the quarterly values, as shown in Table 6.4.

YEAR	Quarter	Code	Pax	4-Quarter MT	MA	Centered MA	% MA
2003	1	1	160	-	-	-	-
2003	2	2	210	-	-	-	-
2003	3	3	90	640	160.0	158.25	56.7
2003	4	4	180	630	157.5	156.25	115.2
2004	1	5	150	620	155.0	156.25	96.0
2004	2	6	200	630	157.5	157.50	127.0
2004	3	7	100	630	157.5	160.00	62.5
2004	4	8	180	650	162.5	167.50	107.5
2005	1	9	170	690	172.5	176.25	96.5
2005	2	10	240	720	180.0	185.00	129.7
2005	3	11	130	760	190.0	190.00	68.4
2005	4	12	220	760	190.0	191.25	115.0
2006	1	13	170	770	192.5	190.00	89.5
2006	2	14	250	750	187.5	186.25	134.2
2006	3	15	110	740	185.0	186.25	59.1
2006	4	16	210	750	187.5	188.75	111.3
2007	1	17	180	760	190.0	193.75	92.9
2007	2	18	260	790	197.5	202.50	128.4
2007	3	19	140	830	207.5	-	-
2007	4	20	250	-	-	-	-

Table 6.3: Computation of seasonal index

Note that in Table 6.3, the quarters are coded for unique identification. We then add up the *four-quarter moving total* (4-Quarter MT) before computing the *moving average* (MA) and the *Centered MA*. The MA ratio (percent) is obtained by dividing Pax with *Centered MA*. The computation of seasonal index average in Table 6.4 is done by eliminating the highest and the lowest values from each quarter and thus reducing the extreme cyclical and irregular variations.

YEAR	Quarter 1	Quarter 2	Quarter 3	Quarter 4
2003	-	-	56.7	115.2
2004	96.0	127.0	62.5	107.5
2005	96.5	129.7	68.4	115.0
2006	89.5	134.2	59.1	111.3
2007	92.9	128.4	-	-
Total	188.9	258.1	121.6	226.3
Mean	94.45	129.05	60.80	113.15
Index	95.1	129.9	61.2	113.9

Table 6.4: Seasonal index

where,

Sum of mean = 94.45 + 129.05 + 60.80 + 113.15 = 397.45

Adjusting factor = 400 / 397.45 = 1.0064

Index = Mean * Adjusting factor

Thereafter, develop a trend line by applying the least squares method to the deseasonalized time series as shown on Table 6.5; where,

$$a = \bar{Y} = (\sum_{i}^{n} Y_i/n) = (3609/20) = 180 \tag{6.4.1}$$

$$b = (\sum_{i}^{n} x_i Y_i / \sum_{i}^{n} x_i^2) = (4205/2660) = 1.6 \tag{6.4.2}$$

$$\hat{Y} = a + bx = 180 + 1.6x \tag{6.4.3}$$

To forecast the OD for 2008 Quarter 1, use the trend-line equation and then multiply the value with the index deduced on Table 6.4.

YEAR	Quarters	Code	Pax	SI	Y_i	TV	x_i	x_iY_i	x_i^2
2003	1	1	160	0.951	168	-9.5	-19	-3192	361
2003	2	2	210	1.299	162	-8.5	-17	-2754	289
2003	3	3	90	0.612	147	-7.5	-15	-2205	225
2003	4	4	180	1.139	158	-6.5	-13	-2054	169
2004	1	5	150	0.951	158	-5.5	-11	-1738	121
2004	2	6	200	1.299	154	-4.5	-9	-1386	81
2004	3	7	100	0.612	163	-3.5	-7	-1141	49
2004	4	8	180	1.139	158	-2.5	-5	-790	25
2005	1	9	170	0.951	179	-1.5	-3	-537	9
2005	2	10	240	1.299	185	-0.5	-1	-185	1
2005	3	11	130	0.612	212	0.5	1	212	1
2005	4	12	220	1.139	193	1.5	3	579	9
2006	1	13	170	0.951	179	2.5	5	895	25
2006	2	14	250	1.299	192	3.5	7	1344	49
2006	3	15	110	0.612	180	4.5	9	1620	81
2006	4	16	210	1.139	184	5.5	11	2024	121
2007	1	17	180	0.951	189	6.5	13	2457	169
2007	2	18	260	1.299	200	7.5	15	3000	225
2007	3	19	140	0.612	229	8.5	17	3893	289
2007	4	20	250	1.139	219	9.5	19	4161	361
					3609			4205	2660

Table 6.5: Identifying the trend component.

$$\begin{aligned}
\hat{Y_{10}} &= (180 + 1.6(21)) \times 0.951 = 203\ passengers \\
\hat{Y_{11}} &= (180 + 1.6(22)) \times 1.299 = 279\ passengers \\
\hat{Y_{12}} &= (180 + 1.6(23)) \times 0.612 = 133\ passengers \\
\hat{Y_{13}} &= (180 + 1.6(24)) \times 1.139 = 249\ passengers
\end{aligned}$$

Total passengers = 203 + 279 + 133 + 249 = 882 passengers

Since the year 2007 had a total of 830 passengers, the annual growth rate is 6.3% (i.e., (882-830)/830).

Later apply the deduced growth rate and seasonal index to the calibrated market size before inputting to the network-simulation model. As an example, suppose the calibrated market size for 2007 is 185 (this represents the reality based on corrected MIDT, YC). The forecast for 2008 will now be:-

$$
\begin{aligned}
\hat{YC}_1 &= 185 \times 1.06 \times 0.951 = 186\, passengers \\
\hat{YC}_2 &= 185 \times 1.06 \times 1.299 = 254\, passengers \\
\hat{YC}_3 &= 185 \times 1.06 \times 0.612 = 120\, passengers \\
\hat{YC}_4 &= 185 \times 1.06 \times 1.139 = 223\, passengers
\end{aligned}
$$

6.4.2 Scenario Generator Assumptions

Since we could not obtain historical MIDT, we resorted to using the airline's own passenger uplift data (also referred as coupon data), to deduce the Seasonality Indices (SI). The superiority of MIDT over coupon data is that the former has competitors information while the latter does not. However, during the calibration process, we have striven to match the bookings based MIDT data to the coupon data, the use of coupon data will have an indicative market seasonality, at least for the markets in which the airline has a strong presence.

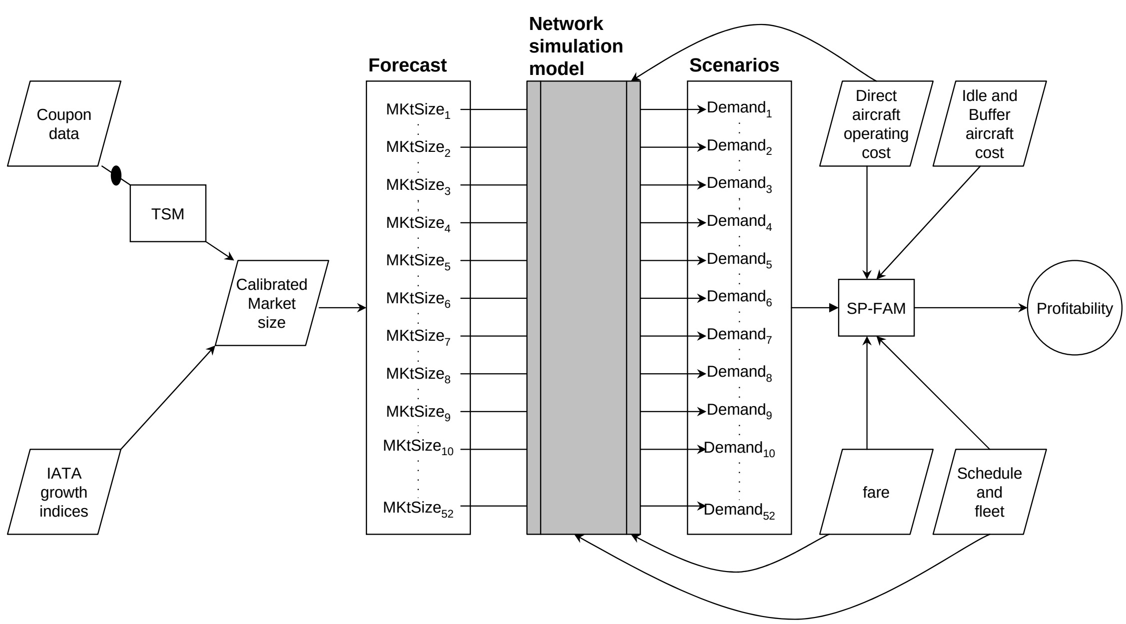


Figure 6.5: Overview of the solution approach.

Prior to the execution of the TSM, we removed inconsistencies within data by eliminating Origin-Destination (O-D) pairs that had many observations missing. Where few observations were missing, we used a mean of the period within the series in which the observation is missing (for other approaches of tackling missing observations refer to Fung (2006)). We constructed a TSM model on three levels, at O-Ds, regional and global level. Further, instead of using the trend equation to forecast the *market sizes* and determine the annual Growth Indices (GI), we resorted to using IATA (2007) published growth rates; also available on three levels i.e. city, sub-region and regional level.

We later applied the deduced SI and GI to the *Calibrated Market* (CM) sizes to generate the *Forecasted Market* (FM) size for each particular week using the *fall-back principle*. When applying the SI on a *fall-back principle*, in the first instance, we used the index at an O-D level, and if a particular O-D was missing on the coupon, we use the regional index, and if still missing, we used the global index. Figure 6.5 depicts the integration between the TSM, network-simulation model and the SP-FAM that explains the solution approach for our robust FAM. The tables on the next page illustrates the *fall-back principle* methodology. *Part A* shows the SI generated for a particular week, on three levels for two regions (that is, Gulf-Africa and Gulf-Europe). To map the the SI (GI) from *Part A* (*Part C*) to the CM sizes (*Part B*) we adopt the following algorithm:

Mapping 1 (Mapping 2): *Part A* (*Part C*) to *Part B*

begin;
if O-D (Country) in *Part B* is in *Part A* (*Part C*) ***then*** SI (GI):= Level 1;
else if O-D (Country) in *Part B* is not in *Part A* (*Part C*) ***then*** SI (GI):= Level 2;
else SI (GI):= Level 3;
end;

The forecasted market size represents a particular week in the subsequent year of the calibration period. Finally, we mapped the planned schedule with the corresponding forecasted *market sizes* to the network-simulation model to predict the unconstrained (not constrained or limited to aircraft capacity) demand for each week. Once the unconstrained demand has been realized, we apply ± 0.5-5% variation to depict twenty additional demand scenarios for that week. That is, if the demand for

Fallback	O_Region	O-D	D_Region	SI
Level 1	Gulf	DXB-NBO	Africa	1.05
	Gulf	DXB-DAR	Africa	0.95
	Gulf	DXB-EBB	Africa	1.20
Level 2	Gulf - Africa			1.07
Level 1	Gulf	DXB-LHR	Europe	1.08
	Gulf	DXB-FRA	Europe	0.95
	Gulf	DOH-CDG	Europe	1.15
Level 2	Gulf – Europe			1.06
Level 3	Gulf – World			1.06

Coupon Data: Part A

Mapping 1

O_Region	O_Country	O-D	D_Country	D_Region	CM Size	SI	GI	FM Size
Gulf	UAE	DXB-NBO	Kenya	Africa	1000	1.05	1.25	1313
Gulf	UAE	DXB-DAR	Tanzania	Africa	750	0.95	1.06	755
Gulf	UAE	DXB-EBB	Uganda	Africa	500	1.20	1.35	810
Gulf	Qatar	DOH-LUN	Zambia	Africa	200	1.07	1.22	261
Gulf	UAE	DXB-LHR	UK	Europe	2500	1.08	1.18	3186
Gulf	UAE	DXB-FRA	Germany	Europe	3000	0.95	1.20	3420
Gulf	Qatar	DOH-CDG	France	Europe	1500	1.15	1.07	1846
Gulf	Qatar	DOH-MEL	Australia	Australia	2000	1.06	1.19	1514

Calibrated Market Size: Part B

Mapping 2

Fallback	O_Country	O-D	D_Country	GI
Level 1	Qatar	DXB-NBO	Kenya	1.25
	Qatar	DXB-DAR	Tanzania	1.06
	Qatar	DXB-EBB	Uganda	1.35
Level 2	Gulf – Africa			1.22
Level 1	Qatar	DXB-LHR	UK	1.18
	Qatar	DXB-FRA	Germany	1.20
	Qatar	DOH-CDG	France	1.07
Level 2	Gulf – Europe			1.15
Level 3	Gulf – World			1.19

IATA Growth Rates: Part C

week 30 is 400 passengers, the following scenarios will be generated:

380 382 384 386 388 390 392 394 396 398 402 404 406 408 410 412 414 416 418 420

6.5 Conclusion

It must be emphasized that an airline's network planning is not only complex but highly volatile and the need for sophisticated network-simulation model is inevitable. In this chapter, we have briefly described the interface between the calibration process and network-simulation model. The heart of airline planning is to come up with a forecast that is close to actual passenger demand. Forecasting becomes even more complex when we have uncertainties surrounding the SP-FAM problem; specifically, tactical demand scenarios are required for the scheduling season under consideration. The previous two SP-FAM scenario generation attempts have not enjoyed application because of the simplistic nature of the underlying methodology. Generating scenarios by use of state-of-the-art network-simulation model gives an alternate way of a scenario-generator for the SP-FAM problem.

The ideal way of generating the scenarios would have been calibrating one week of MIDT instead of averaging several weeks and later deducing weekly demand based on seasonality indices. However, using the current process this is a highly involved and time-consuming process not to mention limitations in computing processing power and memory size that would be required to handle such-large volume of data. The calibration process also relies heavily on personal experience during the fine-tuning phase that cannot be readily substituted by a model; as such, calibration of more than one week will be even more demanding of a calibrator's time. Future research will dwell on substituting personal experience with a robust model, multiple calibration of weekly demand and even more important, calibrating the future as opposed to historical data. In this way, there would be no need to apply growth factors or further interfere with the calibrated data.

Chapter 7

SP-FAM Solution Process

In solving the SP-FAM problem, we follow a series of steps as described in Figure 7.1. We first extract a built up one-week rotated schedule from a scheduling tool in ssim format. The prescribed IATA ssim format (standard schedules information manual) is unreadable to both SAS and AMPL environments. In the second step, we use a special schedule converter (SchedConv) developed by Lufthansa Systems that converts the ssim format into a user-friendly text format, known as the sked format. Next, we perform three preprocessing steps to the resultant schedule before simultaneously inputting the outputs of the preprocessing steps, the scenario generator and others into a SAS-AMPL converter. The SAS-AMPL converter prepares all the data into readable AMPL format. Finally we execute the optimizer using FortMP solver (integrated in AMPL) that invokes branch-and-bound algorithm automatically. After the optimization process, we convert back the resultant schedule into sked format initially, and ssim format subsequently.

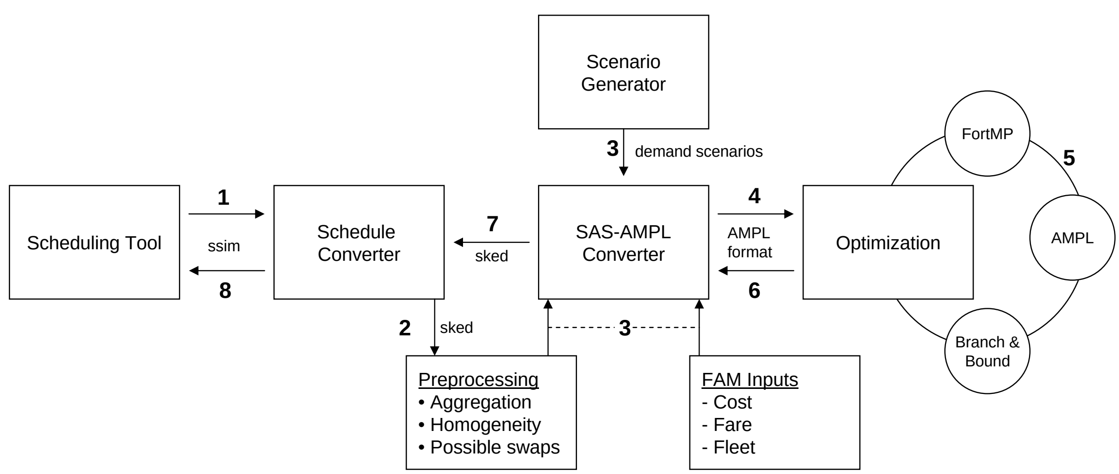


Figure 7.1: SP-FAM solution approach.

Schedule Aggregation

As explained in *section 3.3*; the sheer size of the problem and the required computational time necessitates an aggregation step that makes the representation more compact. The schedule aggregation step can be explained in a simple conventional way as illustrated in Figure 7.2. The diagram shows typical scheduled flights that operate between two stations (DOH and LHR) for a given week. *Part A* shows the representation of the unique flights for each of the operating days before an aggregation step; while *part B* shows the same flights after the aggregation step. Note that all the flights have identical departure and arrival times with similar aircraft type.

Part A

Orig Dest	FNo	Day	Depart	Arrive	Aircraft
DOH LHR	1	1......	750	1075	B772
DOH LHR	1	.2.....	750	1075	B772
DOH LHR	1	...3....	750	1075	B772
DOH LHR	1	...4...	750	1075	B772
DOH LHR	1	....5..	750	1075	B772
DOH LHR	1	.....6.	750	1075	B772
DOH LHR	1	7	750	1075	B772
LHR DOH	2	1......	1275	370	B772
LHR DOH	2	.2.....	1275	370	B772
LHR DOH	2	...3....	1275	370	B772
LHR DOH	2	...4...	1275	370	B772
LHR DOH	2	....5..	1275	370	B772
LHR DOH	2	.....6.	1275	370	B772
LHR DOH	2	7	1275	370	B772

Before schedule aggregation

Part B

Orig Dest	FNo	Day	Depart	Arrive	Aircraft
DOH LHR	1	1234567	750	1075	B772
LHR DOH	2	1234567	1275	370	B772

After schedule aggregation

Figure 7.2: Before and after schedule aggregation.

Homogeneity

Assume the input schedule before optimization is as shown in *part A* of Figure 7.2. Further assume that four aircraft types can operate the given schedule, that are, A332, A333, B772 and B773. The resultant schedule after optimization is shown in *part C* of Figure 7.3. Note that the optimized schedule and the corresponding assigned fleet may not be rational for implementation. For example, we could incur heavy costs associated with crew-related expenses and under-utilization. Consequently, by offering an inconsistent product in the market, the airline's competitive edge is significantly weakened.

To counter such an assignment, we can define a criteria such that only two aircraft should operate the schedule with a balance mix of three and four rotations for each aircraft type. Such a criteria could lead, for example, in an optimized fleet assignment decision as shown in *part D*.

Part C

Orig Dest	FNo	Day	Depart	Arrive	Aircraft
DOH LHR	1	1……	750	1075	A332
DOH LHR	1	.2…..	750	1075	B772
DOH LHR	1	...3….	750	1075	B773
DOH LHR	1	…4…	750	1075	A333
DOH LHR	1	….5..	750	1075	A332
DOH LHR	1	…..6.	750	1075	B772
DOH LHR	1	……7	750	1075	B773
LHR DOH	2	1……	1275	370	A332
LHR DOH	2	.2…..	1275	370	B772
LHR DOH	2	...3….	1275	370	B773
LHR DOH	2	…4…	1275	370	A333
LHR DOH	2	….5..	1275	370	A332
LHR DOH	2	…..6.	1275	370	B772
LHR DOH	2	……7	1275	370	B773

After optimization but before homogeneity

Part D

Orig Dest	FNo	Day	Depart	Arrive	Aircraft
DOH LHR	1	1……	750	1075	A332
DOH LHR	1	.2…..	750	1075	B772
DOH LHR	1	...3….	750	1075	A332
DOH LHR	1	…4…	750	1075	B772
DOH LHR	1	….5..	750	1075	A332
DOH LHR	1	…..6.	750	1075	B772
DOH LHR	1	……7	750	1075	A332
LHR DOH	2	1……	1275	370	A332
LHR DOH	2	.2…..	1275	370	B772
LHR DOH	2	...3….	1275	370	A332
LHR DOH	2	…4…	1275	370	B772
LHR DOH	2	….5..	1275	370	A332
LHR DOH	2	…..6.	1275	370	B772
LHR DOH	2	……7	1275	370	A332

After optimization and homogeneity

Figure 7.3: Before and after defining fleet homogeneity.

Possible Swaps

Determining the possible swaps, homogeneity and aggregation steps are executed concurrently. *Part D* of Figure 7.3 could be aggregated as shown in *part F* of Figure 7.4. However, before the resultant aggregation, one has to enumerate all possible swap options as shown in *part E*. This is particularly necessary during the optimization process where branch-and-bound algorithm is invoked.

Part E

Orig Dest	FNo	Day	Depart	Arrive	Aircraft
DOH LHR	1	1.3.5.7	750	1075	A332
DOH LHR	1	1.3.5.7	750	1075	A333
DOH LHR	1	1.3.5.7	750	1075	B772
DOH LHR	1	1.3.5.7	750	1075	B773
DOH LHR	1	.2.4.6.	750	1075	A332
DOH LHR	1	.2.4.6.	750	1075	A333
DOH LHR	1	.2.4.6.	750	1075	B772
DOH LHR	1	.2.4.6.	750	1075	B773
LHR DOH	2	1.3.5.7	1275	370	A332
LHR DOH	2	1.3.5.7	1275	370	A333
LHR DOH	2	1.3.5.7	1275	370	B772
LHR DOH	2	1.3.5.7	1275	370	B773
LHR DOH	2	.2.4.6.	1275	370	A332
LHR DOH	2	.2.4.6.	1275	370	A333
LHR DOH	2	.2.4.6.	1275	370	B772
LHR DOH	2	.2.4.6.	1275	370	B773

Enumerated swaps

Part F

Orig Dest	FNo	Day	Depart	Arrive	Aircraft
DOH LHR	1	1.3.5.7	750	1075	A332
DOH LHR	1	.2.4.6.	750	1075	B772
LHR DOH	2	1.3.5.7	1275	370	A332
LHR DOH	2	.2.4.6.	1275	370	B772

Resultant optimized aggregated schedule

Figure 7.4: Aggregated swap enumeration.

SAS-AMPL Converter

In solving the SP-FAM problem, we model the problem using the algebraic modelling language AMPL (Fourer et al., 2002). The AMPL choice is appealing for several reasons;

first, it has been used successfully by large airlines (e.g. US Airways) in modelling the FAM problem. AMPL has succinct structural features represented by sets, nodes, arcs, scenarios, etc that offers flexibility in the representation and modelling of many mathematical programming problems. AMPL has not only integrated the SP features that are pertinent to the problem under consideration but also well integrated with many solvers such as FortMP (OptiRisk, 2008*a*) and FortSP solver. While the former is designed to solve a wide range of well known optimization problems (including SP), the latter is only designed for solving SP problems. During the solution process, FortMP automatically selects branch-and-bound as the most appropriate algorithm for the SP-FAM.

The SAS application is expensive but highly flexible and powerful for both modelling and analysis. A necessary step before execution of the optimizer in the *AMPL Shell* is the preparation of the input data. In essence, several conversion programs need to be created that prepare the input files for the optimization process. Further, conversion programs also need to be created that take the output of the optimization process and convert it into an integrative format for the scheduling tool. Figure 7.5 shows the interactive shell-level modelling environment of AMPL.

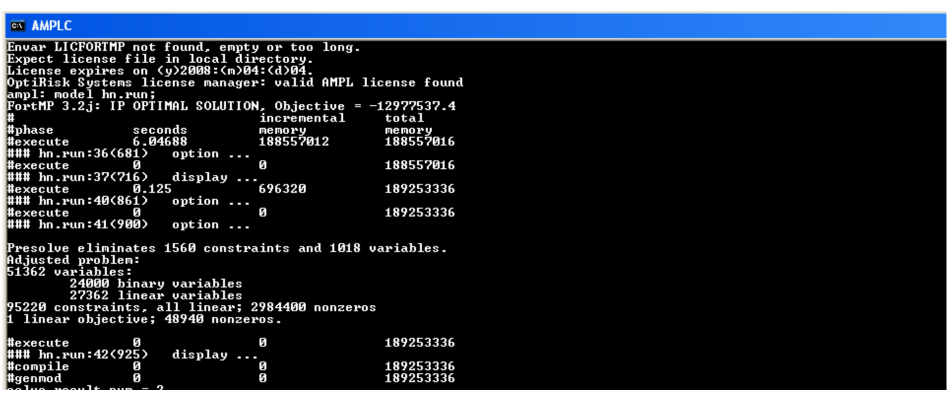


Figure 7.5: AMPL Shell.

Optimization

As noted above, in solving the SP-FAM problem with FortMP, branch-and-bound is automatically selected as the most appropriate solution algorithm. Branch-and-bound is a general search method algorithm for finding solutions of various optimization problems, especially in discrete and combinatorial optimization. It consists of a systematic enumeration of candidate solutions, where large subsets of fruitless candidates are discarded *en masse*, by using upper and lower estimated bounds of the quantity being optimized.

The branch-and-bound algorithm starts by considering the root problem, that is, the original problem with the completely feasible region, and applying the lower-bounding and upper-bounding procedures to the root problem. If the bounds match, then an optimal solution has been found and the procedure terminates. Otherwise, the feasible region is divided into two or more regions that become sub problems of the partitioned feasible region.

The algorithm is applied recursively to the sub problems. If an optimal solution is found to a sub problem, it is a feasible solution to the full problem, but not necessarily globally optimal. If the lower bound for a node exceeds the best known feasible solution, no globally optimal solution can exist in the subspace of the feasible region represented by the node. Therefore, the node can be removed from consideration. The search proceeds until all nodes have been solved or pruned, or until some specified threshold is met between the best solution and the lower bounds for all unsolved sub problems. For a more intuitive understanding, consider the following IP model:

$$Max \;\; z = \sum_j c_j x_j \tag{7.0.1}$$

$$subject \;\; to: \quad \sum_j a_{ij} x_j \leqslant b_i \quad i = 1, ...m \tag{7.0.2}$$

$$x_j \geq 0 \; and \; integer \quad j = 1, ...n \tag{7.0.3}$$

Step 0: *Initialization.* Let the master list initially include only the original linear program, and let the first iteration be denoted by t=1, and z_1= -∞.

Step 1: *Branching.* Stop if the master list is empty. Otherwise select a program from the master list.

Step 2: *Relaxation.* Solve the problem taken from the master list. If the problem has no feasible solution, or if its objective function value z is less than z_t (this branch has fathomed), let $z_{t+1} = z_t$ and go to Step 1. Otherwise go to *Step 3*.

Step 3: If the solution to the solved LP satisfies the integer constraints, then store the solution and let z_{t+1} equal the objective function value for the solution. Since this branch has been fathomed, go to *Step 1*. If the integer condition is not satisfied, go to *Step 4*.

Step 4: *Separation*. Select any variable x_j whose value b_j in the current solution does not satisfy the integer requirement. Add two problems to the master list; these problems are identical to the one just solved except that in one we add:

$x_j \geqslant [b_j]{+}1$

and in the other we add:

$x_j \leqslant [b_j]$.

Let $z_{t+1} = z_t$ and go to *Step 1*.

The major limitation on our SP-FAM testing is the sole use of FortMP that invoked branch-and-bound algorithm. For testing the model with SP algorithms discussed in section 4.6, we would have first needed to represent the problem in SMPS format that could be acceptable to other solvers (see section 4.7). Unfortunately, the SMPS format has not enjoyed universal acceptance because of difficulties with its rigidity and varying constructs. In fact, to date, there is no optimization SP solver that will accept all the wide variation SP problem sets. In our case, we encountered two difficulties; the lack of a typical example of a test example that suits the SP-FAM problem and second, the real, live, problem has thousands of variables and millions of constraints, the accurate representation of which is not trivial. It is probable that SPInE (OptiRisk, 2008*b*) using the FortSP solver would have offered a solution to the above; but during the research, SPInE was still evolving and had not fully integrated into AMPL syntax.

Chapter 8

Case Study

In this chapter we give a proof of concept using real data from a Middle East airline. We quantify the robustness of both the SP-FAM solutions and the scenario generator forecasts through the stochastic measures and stability measures, respectively.

8.1 SP-FAM

We applied the methodology described in the previous sections to a real airline carrier with a weekly schedule containing 79 stations, 1356 legs and five fleet types with a total of 62 aircraft (40 wide-body and 22 narrow-body) including 5 standby (3 wide-body and 2 narrow-body). Due to sheer size of the problem, the solver would allow a maximum of 130 demand scenarios before reaching the iteration limit. The tests are thus conducted using 50, 100 and 125 number of demand scenarios. The following assumptions were made:

- The spill factor was taken as 10% of aircraft capacity.
- Passenger revenue effect and recapture were not considered in the model.
- The penalty cost ($j_{k,i}$) of using the standby aircraft was taken as 2% of the aircraft *direct operating cost.*
- The aircraft utilization cost (w_k) was derived using the strategic methodology discussed in section 5.4.5.
- The MN_k values for wide-body was taken as 17 hours per day while for narrow-body 14 hours per day.
- The MX_k values for wide-body was taken as 21 hours per day while for narrow-body 19 hours per day.

We consider the deterministic demand to be the weighted average (i.e. the expected value) of the stochastic demand. The model so constructed using the deterministic demand is called the Expected Value (EV) model and the corresponding objective is denoted as Z_{ev}. Specifically, we solved the DFAMs with the weighted passenger demand and without using the standby aircraft. In a deterministic environment, we assume that the schedule will be flown with the planned aircraft and the need of a standby aircraft does not arise. Conventionally, a deterministic solution can always be found from the recourse model (i.e. Z_{hn} where HN stand for here-and-now) by setting the number of scenarios to 1 and the corresponding probability of the weighted passenger demand to 1. In this case, the approach of solving the EV model will be identical to solving the wait-and-see (WS) model except that we use an average of all the demand scenarios. But in our case, this approach is not representative as we do not wish to have a standby aircraft with the corresponding penalty cost value (j_{ki}) in the first place. If we set the standby aircraft values to zero, the SP-FAMs will still give high EV values (mainly driven by b_{kis} variable) for the virtual profit. But even then, still this does not represent the deterministic equivalent as the virtual profits are non-existent and cannot be tracked anywhere in the financial books of account.

Through the Z_{ev} and Z_{hn} models we construct the *Expectation of the Expected Value* solution (EEV) model by fixing the first stage (non-recourse decisions) in the HN model to that of the EV model and solving for the remaining variables. The objective value so obtained (denoted by Z_{eev}) indicates the impact of implementing a deterministic solution in a stochastic environment. VSS represents the additional gain obtained on modelling and solving the stochastic model. When the VSS is small then the expected value deterministic model is as good as the stochastic model and we can ignore the uncertainties. We also process each of the scenario models individually; such models are known as *wait-and-see* and the probability weighted objective for all the scenarios is denoted by Z_{ws}.

We modelled the problem using the algebraic modelling language AMPL (Fourer et al., 2002), generated all the input data using SAS and solved with FortMP (OptiRisk, 2008*a*);

where branch-and-bound is invoked automatically. All experiments are performed on Windows 2000 having a dual processor of 3.4 GHz and 3.24 GB of RAM.

8.1.1 SP-FAM Statistics

Table 8.1 shows the computer runtime and statistics of the *here-and-now* model under varying scenarios for both the SP-FAM1 and SP-FAM2 (using the same data source).

Scen.	Runtime (Sec.)		Binary	Linear Variables		Constraints
	SP-FAM1	SP-FAM2	Variables	SP-FAM1	SP-FAM2	
50	166	160	1200	62237	62228	62370
100	655	568	1200	122335	122326	122370
125	948	881	1200	152385	152376	152370

Table 8.1: Here-and-now model statistics

The results at Table 8.1 indicate that for a given high computing processing power and memory, solving such a large-scale problem is highly efficient.

8.1.2 SP-FAM Stochastic Measures

Table 8.2, 8.3 and 8.4 shows the objective values given in UK £ of the various SP models for SP-FAM1 and SP-FAM2, respectively.

Scenarios	Z_{ev}	Z_{ws}	Z_{hn}	Z_{eev}	EVPI	VSS
$SP-FAM1$						
50	2,102,768	19,572,735	19,569,195	∞	3,540	∞
100	2,213,569	19,792,435	19,786,168	∞	6,266	∞
125	2,308,793	19,960,050	19,951,351	∞	8,699	∞
$SP-FAM2$						
50	1,361,502	18,803,127	18,798,075	∞	5,051	∞
100	1,471,918	19,023,220	19,015,363	∞	7,857	∞
125	1,567,091	19,190,818	19,197,834	∞	7,016	∞

Table 8.2: Z_{TP}: Objective values and SP measures

Table 8.2 indicate that, for both SP-FAM1 and SP-FAM2, solving the *here-and-now* model is much superior to the *expected-value* and *expected of the expected-value* models.

Scenarios	Z_{ev}	Z_{ws}	Z_{hn}	Z_{eev}	EVPI	VSS
$SP-FAM1$						
50	2,102,768	1,403,510	1,494,987	∞	91,477	∞
100	2,213,569	1,478,657	1,544,635	∞	65,979	∞
125	2,308,793	1,552,784	1,657,952	∞	105,169	∞
$SP-FAM2$						
50	1,361,502	645,897	744,796	∞	98,899	∞
100	1,471,918	721,529	811,346	∞	89,817	∞
125	1,567,091	795,084	948,529	∞	153,446	∞

Table 8.3: Z_{RP}: Objective values and SP measures

Scenarios	Z_{ev}	Z_{ws}	Z_{hn}	Z_{eev}	EVPI	VSS
$SP-FAM1$						
50	2,640,867	1,931,366	2,028,001	∞	96,635	∞
100	2,751,668	2,000,851	2,076,137	∞	75,285	∞
125	2,846,893	2,067,012	2,187,949	∞	120,938	∞
$SP-FAM2$						
50	2,667,417	1,945,741	2,050,230	∞	104,488	∞
100	2,777,833	2,015,565	2,115,268	∞	99,703	∞
125	2,873,007	2,080,979	2,250,946	∞	169,967	∞

Table 8.4: Z_{AP}: Objective values and SP measures

Note that Z_{ws} and Z_{hn} values are much higher than Z_{ev}, this is simply because of the deceptive behaviour inherent by $b_{k,i,s}$ assignment (refer to section 5.4.4).

If we extend the measurement and apply to the *Real Profit* (Z_{RP}) i.e. a component of the the *Total Profit* (Z_{TP}), we obtain Table 8.3. For our purpose, we consider the Z_{ev} to be the same as shown in Table 8.2. Table 8.3 indicate that, for both SP-FAM1 and SP-FAM2, solving the *here-and-now* model outperforms the *wait-and-see* and *expected of the expected-value* models. Further, we have the inter-bound relationship holding (i.e. $Z_{ws} \leqslant Z_{hn} \leqslant Z_{eev}$).

If we further extend the measurement and apply to the *Actual Profit* (Z_{AP}), we obtain Table 8.4. Table 8.4 portrays similar conclusions as in Table 8.3 but in addition, we have SP-FAM2 outperforming SP-FAM1. The Z_{AP} values depict the standard agreeable measure that can easily be tracked in the financial books of accounting. A striking feature about the overall measurement is the EVPI values that appears significantly higher for Z_{RP} and Z_{AP} compared to Z_{TP}. While a higher EVPI means that more gain will be obtained in knowing the future, in our context, it also emphasizes the fact that we cannot ignore the Z_{RP} and Z_{AP} values.

Scenarios	Z_{ev}	Z_{ws}	Z_{hn}	Z_{eev}	EVPI	VSS
$SP-FAM1$						
50	2,102,768	18,169,225	18,074,207	∞	95,018	∞
100	2,213,569	18,313,778	18,241,533	∞	72,245	∞
125	2,308,793	18,407,267	18,293,399	∞	113,868	∞
$SP-FAM2$						
50	1,361,502	18,157,230	18,053,279	∞	103,951	∞
100	1,471,918	18,301,692	18,204,017	∞	97,675	∞
125	1,567,091	18,395,735	18,249,305	∞	146,430	∞

Table 8.5: Z_{VP}: Objective values and SP measures

Finally, if we now subject the measurement to the *Virtual Profit* (Z_{VP}) and assume that the Z_{ev} to be the same as shown in Table 8.2, we obtain Table 8.5. The results in Table 8.5 portrays contradictory conclusions as that of Z_{RP} and Z_{AP} relationship, that is, the inter-bound relationship ($Z_{ws} \leqslant Z_{hn} \leqslant Z_{eev}$) does not hold. This behaviour depicts the trade-off logic of making the Z_{RP} value obeying the inter-bound relationship at the expense of the Z_{VP} value.

8.2 Stability Measure

The stability requirement has two desirable aspects which need consideration, that is, the *in-sample* and the *out-of-sample* measures. A scenario-generator is said to manifest in-sample stability if, when generating several scenario sets of the same size (from the same sample) and solving the optimization problem on each of these scenario sets, the optimal objectives are similar. If the optimal solution does not change significantly for inputs taken from outside the sample, we have out-of-sample stability.

The out-of-sample stability is the important one, since this says that the real performance of the solution is stable, i.e. it does not depend on which scenario set we solved the optimization problem. In-sample stability does not imply the out-of-sample one or vice versa. It is possible to have in-sample instability (of the objectives) but stability of the solutions in this case, it is likely to have out-of-sample stability, since, in this case, all the solutions are tested on the same scenario set. Given that SP-FAM2 outperforms SP-FAM1 in Table 8.4, we use SP-FAM2 to investigate the stability of the scenario-generator .

8.2.1 In-Sample Stability Measure

In our case study, we decided to use a tree with twenty scenarios and subsequently carried out fifty simulation runs. The distribution of the objectives obtained by the simulation is used to extract the following measures of stability.

Stability measured by	Value
Min, in UK £	18,011,783
Max, in UK £	19,178,926
Range, in UK £	1,167,143
Mean, in UK £	18,563,886
Stdev, in UK £	306,659
RMnD	1.65%

Table 8.6: In-sample stability.

- *Min*: Min represents the minimum objective value of all the simulation runs.
- *Max*: Max represents the maximum objective value of all the simulation runs.
- *Range*: The value *Range* is the difference between Max and Min and represents the maximum spread between all the runs.
- *Mean*: Represents the average of all the objective values.
- *Stdev*: Stdev is the standard deviation of all the objective values.
- *RMnD*: Represents the *relative mean deviation* and is expressed by the fraction between the Stdev and the Mean (i.e. $RMnD = \frac{Stdev}{Mean}$).

When plotting the distribution of the objectives functions (as shown in Figure 8.1), a flat curve is obtained. We also observe that the value given by the RMnD in Table 8.6 is less than 2%, therefore can assume that the scenario generation method used in this study is stable. Use of a tree with a larger number of demand scenarios in the simulation runs, will provide more reliable values of stability measures with a smaller error interval.

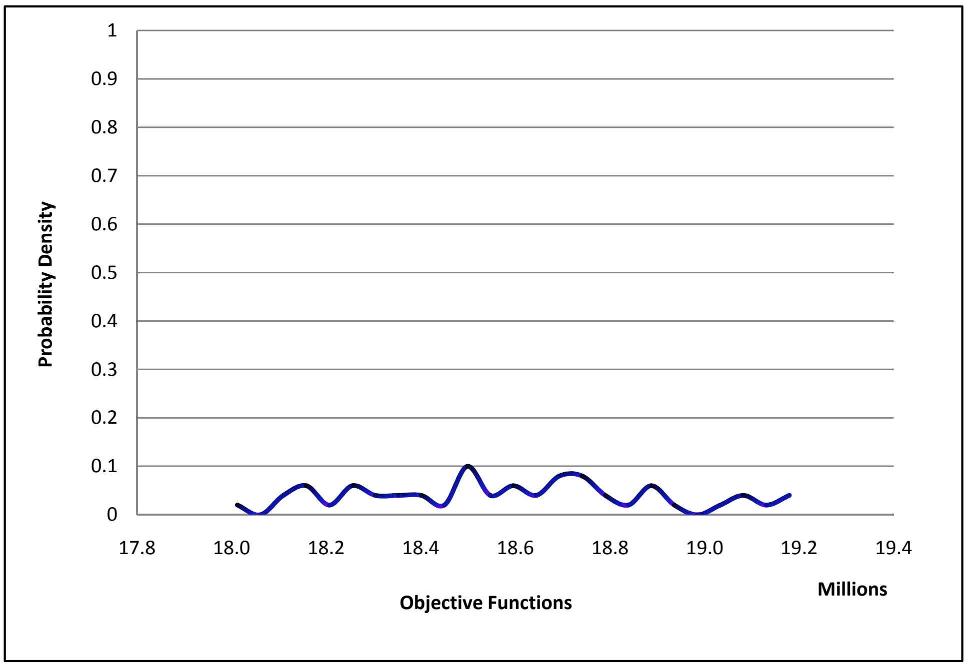


Figure 8.1: In-sample stability.

8.2.2 Out-of-Sample Stability Measure

The *out-of-sample* stability is concerned with the robustness of the optimal decisions obtained by solving an SP problem with a given scenario generation method. In this case, a very large scenario tree, generated with a different method, is used as the *real* stochastic process, and the performance of the optimal solutions is computed in relation to this tree. In other words, the scenario generation method is compared in *absolute* terms with what is supposed to be the real stochastic process (this in reality can never be done, so the large tree is used as a substitution of the real process). More formally:

Let $\xi_1 ... \xi_n$ be n sets of scenario trees and let the optimum decisions obtained by solving

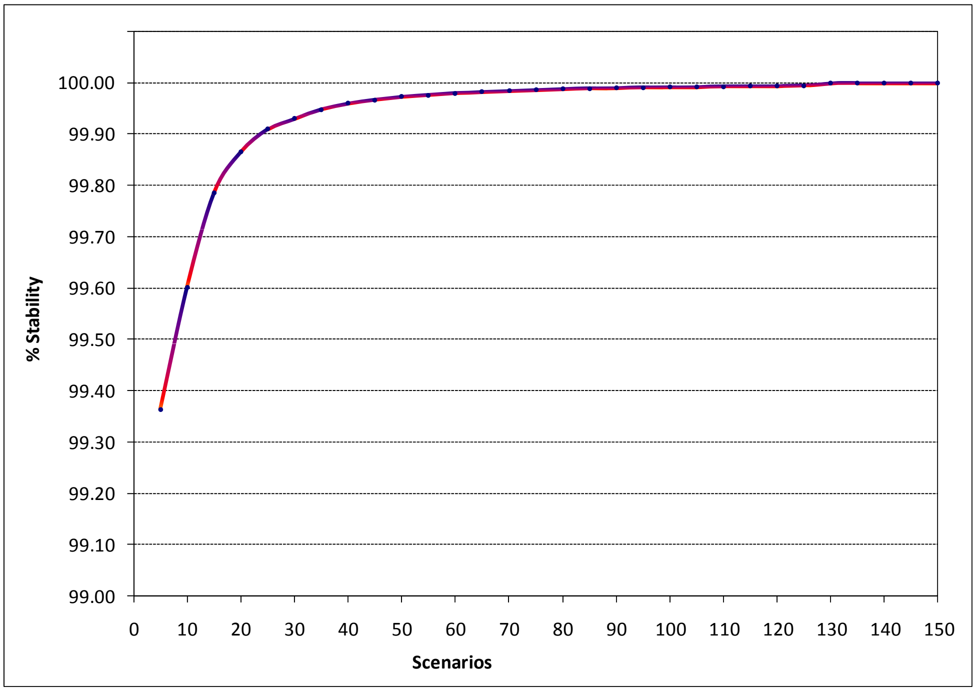


Figure 8.2: Out-of-sample stability.

the problem be represented by $x_1^*...x_n^*$. Let also $\bar{\xi}$ be a large scenario tree which is assumed to be the best available approximation of the *real* stochastic process. The objective function values obtained by evaluating $x_1^*...x_n^*$ using the tree $\bar{\xi}$ are represented by:

$$o_1....o_n \;=\; F(x_1^*, \bar{\xi})....F(x_n^*, \bar{\xi}) \tag{8.2.1}$$

We can then compute the average distance ($\bar{d}$) between all pairs of values:

$$\bar{d} = \frac{1}{(n^2 - n)/2} \sum_{i=1}^{n} \sum_{j=i+1}^{n-1} |O_i - O_j| \tag{8.2.2}$$

As a measure of the stability (s) of the scenario generation method we use the ratio between the mean distance ($\bar{d}$) and the mean objective (for all scenarios) (μ):

$$s = 1 - \frac{\bar{d}}{\mu} \tag{8.2.3}$$

Figure 8.2 shows how the *out-of-sample* stability for our scenario generator is increasing when the number of scenarios generated by the method increases. In this case, the method seems perfectly stable with trees of more than 130 scenarios.

8.3 Conclusion

In our solution approach we have shown that given high computing processing power and memory, solving such a large-scale problem is highly efficient. Through real data from an airline, we have proved that the SP-FAM satisfies the recourse conditions with inter-bound relationship holding (i.e. $Z_{ws} \leqslant Z_{hn} \leqslant Z_{eev}$) for both Z_{TP}, Z_{RP} and Z_{AP}. In addition, we have also proved that SP-FAM2 outperforms SP-FAM1 and that the inter-bound relationship does not hold for Z_{VP}.

In testing the reliability of the scenario generator, both in-sample and out-of-sample tests were done. For in-sample stability, when plotting the distribution of the objectives functions, a flat curved is obtained. We also observed that the value given by the RMnD is less than 2%, therefore assumed that the scenario generation method used in this research is stable. Similarly, the out-of-sample test revealed that the stability for the scenario generator is increasing when the number of scenarios generated by the method increases. The method seems perfectly stable with trees of more than 130 scenarios.

Chapter 9

Conclusion and Future Work

The complexity and the unprecedented growth facing the airline industry pose a major challenge for planning under uncertainty conditions. Currently, the trend is towards the use of deterministic and non-integrated *Decision Support Systems* (DSS); however, with increasing competition, managers require more robust DSS that integrate different models, databases and solution approaches that will cope up with the increasing demand and pressure from airlines. Although several attempts have been made to narrow the gaps; there is still scope for enhancing efficiency and reliability. On the one hand, there is an algorithmic challenge that, although SP provides a natural mechanism for tackling the uncertainty conditions, little application has been enjoyed. Similarly, system integration, and in particular, the use of relational databases within the optimization solvers, has not been fully exploited.

9.1 Conclusion

Robustness within FAM remains a vibrant research area with many challenges on model formulation, accurate passenger forecast and integration of such models. In Chapter 1 we gave an overview of the optimisation application within the airline industry. In Chapter 2, the robust models and solution algorithms that tackle schedule robustness were discussed. Chapter 3 described the FAM in detail while Chapter 4 explained the concept of SP in general.

The main research contribution is summarized as follows:-

- We have presented new FAMs (the DFAMs) that tackles the major problem of aircraft utilization associated with basic FAM. The DFAMs saves an optimization expert's time in re-optimizing aircrafts while seeking trade-off in attaining desired utilization balance.

- Using two-stage SP with recourse, we have provided two ways of formulating the SP-FAM (i.e. SP-FAM1 and SP-FAM2). The distinguishing feature of the model formulation is the introduction of the aircraft utilization variable, standby aircraft and aircraft assignment artificial variable.

- Through an empirical testing we prove that SP-FAM2 outperforms SP-FAM1 in more representation of the recourse model.

- We have come up with a new way of generating passenger demand scenarios through a network-simulation model. In our scenario-generator, we have embedded an *additive TSM* that forecasts the market sizes as an input to the network-simulation model. While previous research tackled the integration of the demand through averaging of scenarios (*expected-value*), our formulation integrates all the scenarios concurrently (*here-and-now*). Through SP-FAM2, we prove reliance of the scenario-generator through the *in-sample* and *out-of-sample* stability measures.

- We then provide a framework showing how the model can be integrated with uncertain demand scenarios through interfacing network-simulation model, scheduling tool and AMPL-SAS converter.

- The hurdle associated with solving FAM in general and SP-FAM in particular, is the sheer size of the problem. While schedule aggregation is one such preprocessing step, we combine other preprocessing steps (such as homogeneity and possible swaps) that, traditionally, have not been combined concurrently.

- Through a real case study, we have shown that solving the *here-and-now* model is superior to both the *expected-value* and the *expectation of the expected-value* models. We thus vindicate the long-standing notion that it is impossible to account for all demand scenarios in the same model.

9.2 Future Work

9.2.1 Solution Algorithm

The major limitation of SP-FAM testing is the sole use of FortMP that invokes branch-and-bound algorithm. Unfortunately, for testing SP problems with other algorithms, the model has to be expressed in an SMPS format that converts the existing deterministic linear programs into stochastic ones by addition of information about the dynamic and stochastic structure of the model. While conversion into SMPS with the issue of thousands of variables is non-trivial, there is no solver that will solve all instances that can be expressed in the SMPS format; and this applies to the SP-FAM problem.

The standard SMPS format representation entails generation of three files known as the *core*, *time* and *stoch*. The *core* file is equivalent to the MPS format and can be generated from the deterministic equivalent of the SP problem. In the *core* file, the problem dimensions and deterministic coefficients as well as the locations of all the stochastic coefficients are determined. The *time* file contains the information needed to specify the dynamic structure of the problem. The *time* file indicates the position (in terms of rows and columns) of the elements of the decision vector identified for each period. The *stoch* file specifies the distribution of the random variables.

To prepare the file formats for the SP-FAM problem, one needs to account for three main challenges. Selection of a solver that has been customized to accept the specialized features of the problem beforehand; generating the required format through automation that will realign data on the three files (*core*, *time* and *stoch*); and finally, be able to interpret the data once the problem has been solved.

9.2.2 Revenue Effects

In the SP-FAM formulation discussed in this research, only the stochastic nature of demand was considered. Actually, in all the SP-FAM attempts made so far, RM effects have never been taken into account. In principal, all the models have used the average fare for one passenger type, which is far from reality. Flight fares are highly uncertain with wide variation between fare-classes being offered. Furthermore, we have the network effects (spill and recapture) that are compounded by passenger type which, if modelled into the problem, will add complexity and pose a major challenge in both formulation and solution algorithms.

9.2.3 Improving Demand Forecast

During the calibration of the network-simulation model we endeavoured to deduce historical weighted load-factor (LF) by flight level. Ideally, as an input to FAM, this could have a severe impact on the realized fleeting decision. In fact, for running SFA, hardly do we take the input of the demand forecast from network-simulation model without fine-tuning the LF at flight level and by day of week; and as such, this is one of the deficiencies of the reliance on the network-simulation model. Although different approaches have been used in modelling passenger time-of-day preference, such as the inclusion of time-of-day dummy variables for each hour of the day; and the weighting parameters for a series of sin and cos curves. The models so applied only focus on time-of-day and not day of week, and this is of concern to network planners who strive to get a realistic representation

The other difficulty is matching the airline's projected market share with the actual share for the chosen representative week in the network-simulation model. As an input to the calibration, the MIDT data is customized on a leg or segment basis (as opposed to O-D level) with different trip building rules. Figure 9.1 shows an O-D pair denoted by JNB-DOH. This O-D pair has two legs (that is, JNB-CPT and CPT-DOH) and three segments (that is, JNB-CPT, CPT-DOH and JNB-DOH).

The network-simulation tool, takes the output of the calibrated data (in segment format) and projects the markets at O-D level that, when compared with MIDT market share, differs significantly. Furthermore, this problem is compounded when one compares the airline's actual passenger numbers with the network-simulation forecast. These two shortcomings are inter twined, in that, if we can deduce market share from MIDT and consider the airline's actual passenger uplift, then the market size can be estimated. Unfortunately, this can only

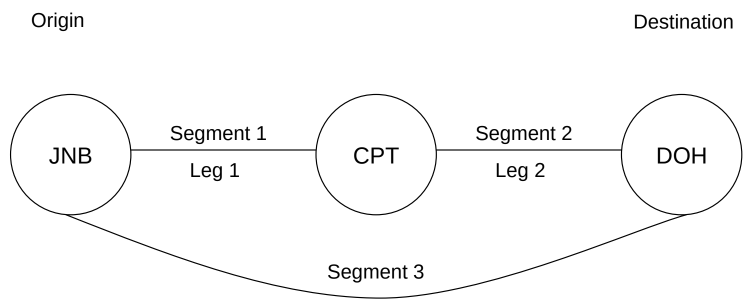


Figure 9.1: O-D pair.

be harmonized, to some extent, during the fine-tuning phase of the calibration process. As such, there is scope for further improving the demand forecast of the network-simulation tool by automating the fine-tuning phase into rough calibration phase.

9.2.4 System Integration

Integrating the FAM with uncertain demand is only one of the challenges; there are still many integration aspects to the resultant SP-FAM. For instance, integrating SP-FAM with RM, integrating SP-FAM with crew planning, integrating SP-FAM with Operations Control Centre, are but a few of the challenges. The airline fleet assignment process affects many other processes and normally comes as a module within a planning system. In their system investment criteria, airlines adopt an enterprize acquisition strategy. They specifically look at a system that has integrated (or is integrative) with many other core processes that support the entire planning cycle. Unfortunately, this has not been easy to achieve and will remain a challenge for many years to come.

Bibliography

Abara J (1989). 'Applying integer linear programming to the fleet assignment problem'. *Interfaces* **19**(4). 20 - 28.

Ahmed A and Taskila T (2007). *Prioritizing launching of new destinations*. AGIFORS Strategic and Scheduling Study Group. Tokyo, Japan.

Ahuja R (2006). *Network Application 4*. University of Florida. http://www.ise.ufl.edu/ANO/. Advanced Network Optimization Class Notes ESI 6912 Section (6129) - Fall 2008.

Ahuja R, Goodstein G, Orlin J, Sharma D, Liu J and Murkherjee A (2003). *Solving multi-criteria through-fleet assignment models*. In: T Ciriani and G Fasano and S Gliozzi and R Tadei(editor). Operations Research in space and air. Springer. chapter 13, pp. 233–256.

Ahuja R, Goodstein J, Mukherjee A, Orlin J and Sharma D (2007). 'A very large-scale neighborhood search algorithm for the combined through-fleet assignment model'. *Informs Journal on Computing* **19**(3). 416 - 428.

Altenstedt F (2009). *Bouncing nested Bendes solver*. http://www.math.chalmers.se. as accessed in July.

Alvarez J (2002). *Iberia Serio*. AGIFORS Symposium Proceeding.

Ambrosini R, Lemaire S, Odu O and Gueguen C (2003). *Solving the aircraft rotations problem in Air France*. AGIFORS Scheduling and Strategic Planning Study Group. Toulouse, France.

Ball M, Barnhart C, Nemhauser G and Odoni A (2007). *Air Transportation: Irregular Operations and Control*. Handbook in operations research and management science: transportation. Elsevier, North Holland. C Barnhart and G Laporte. Volume 14, Page 1-61.

Barnhart C (2003). *Airline Schedule Planning: 1.206 Airline Operations: Lecture No.1*. MIT open courseware. http://ocw.mit.edu/ocwWeb/Civil-and-Environmental-Engineering/index. Airline scheduling.

Barnhart C, Boland N, Clarke L, Johnson E, Nemhauser G and Shenoi R (1998). 'Flight string models for aircraft fleeting and routing'. *Transportation Science* **32**(3). 208 - 202.

Barnhart C, Johnson E, Nemhauser G and Vance P (1999). *Crew scheduling*. In: R Hall (editor). Handbook of Transportation Science. Springer. chapter 14, pp. 493–523.

Barnhart C, Kniker T and Lohatepanont M (2002). 'Itinerary-based airline fleet assignment'. *Transportation Science* **36**(2). 199 - 217.

Barnhart C, Lu F and Shenoi R (1998). *Integrated airline schedule planning.* In: G Yu (editor). OR in the airline industry. Kluwer Academic Publishers. Boston, US. chapter 13, pp. 384–401.

Belobaba P (2006). *16.75J/1.234J Airline Management.* MIT open courseware. http://ocw.mit.edu/ocwWeb/Aeronautics-and-Astronautics/index. Airline fleet planning models.

Berge M and Hopperstad C (1993). 'Demand driven dispatch: A method of dynamic aircraft capacity assignments, models and algorithms'. *Operations research* **41**(1). 153 - 168.

Bian F, Burke E, Jain S, Kendall G, Koole G, Silva J, Mulder J, Paelinck M, Reeves C, Rusdi I and Suleman M (2005). *Measuring the robustness of airline fleet schedule.* In: G Kendall and E Burke and S Petrovic and M Gendreau (editor). 1st Multidisciplinary scheduling: Theory and Applications. Springer Academic Publishers. MISTA. pp. 381–392.

Birge J (1988). *Large-scale linear programming techniques in stochastic programming.* In: Numerical Techniques for Stochastic Optimisation. Springer, Berlin. Y Ermoliev and R Wets.

Birge J (1997). 'State-of-the-art survey: Stochastic programming computations and applications'. *INFORMS Journal on Computing* **9**(2). 111 - 133.

Birge J and Louveaux F (1997). *Introduction to Stochastic Programming.* Springer Verlag. New York, US. Chapter 6. 215 - 224.

Boeing (2003). *Schedule analysis and fleet planning.* Airline planning seminar. Seattle, USA. Session 10.

Boyd E (2006). 'Revenue management in the airline industry: From gumball dispensers to rocket science'. *Journal of Revenue and Pricing Management* **5**(2). 157 - 166.

Bratu S and Barnhart C (2005). 'An analysis of passenger delays'. *Air Traffic Control Quarterly* **13**(1).

Bratu S and Barnhart C (2006). 'Flight operations recovery: New approaches considering passenger recovery'. *Journal of Scheduling* **9**(3). 279-298.

Busutulli L (1999). *Revenue management.* AGIFORS Revenue Management Study Group. London, UK.

Caroe C, Gollmer R, Markert A and Schult R (2009). *Dual decomposition in two-stage stochastic mixed-integer programming*. http://neos.mcs.anl.gov/. as accessed in July.

Charnes A and Cooper W (1959). 'Chance-constrained programming'. *Management Science* **6**(1), 73 – 79.

Clark P (2001). *Buying the big jets: fleet planning for airlines*. Ashgate Publishing Ltd. Aldershot, UK. Page 41.

Clarke J (2004). *The case for robust schedule*. AGIFORS Scheduling and Strategic Planning Study Group. Macau, China.

Clarke M (1997). Irregular airline operations: A review of the state-of-the-art practice in airline operations control centers. Technical report. Massachusetts Institute of Technology. http://web.mit.edu/heart/Public/www/wp1.pdf. Page 6-7.

Coates J (2005). *Revenue management*. In: Vision 2020: Ahmedabad. Richard Ivey School of Business. P Bell. http://www.ivey.uwo.ca. Slide 9.

Cook T (2002). 'Looking at no.1'. *ORMS Today* **29**(1). 28 - 33.

Dantzig G and Glynn P (1990). 'Parallel processors for planning under uncertainty'. *Annals of Operations Research* **22**. 1 - 21.

Dempster M (1980). *Stochastic Programming*. Academic Press. London.

Desaulniers G, Desrosiers J, Dumas Y, Solomon M and Soumis M (1997). 'Daily aircraft routing and scheduling'. *Management Science* **43**(6). 841-855.

Domenica N, Mitra G, Valente P and Birbilis G (2007). 'Scenario generation within a simulation framework: An information system perspective'. *Decision Support Systems* **42**(4). Page 2197-2218.

Dunleavy H and Westermann D (2005). 'Future of revenue management: Future of airline revenue management'. *Journal of Revenue and Pricing Management* **3**(4). 380 - 383.

Erdmann A, Nolte A, Noltemeier A and Schrader R (2001). 'Modeling and solving an airline schedule generation problem'. *Annals of operations research* **107**(1-4). 117 - 142.

Ferguson A and Dantzig G (1955). 'The aircraft routing problem'. *Aeronautical Engineering Review* **14**(4).

Fourer R, Gay D and Kernighan B (2002). *AMPL: A Modeling Language for Mathematical Programming*. Duxbury Press. Belmont, US.

Fung D (2006). Methods for the estimation of missing values in time series. Master's thesis. Edith Cowan University.

Gassman H (1990). 'Mslip: A computer code for the multistage stochastic linear programming problem'. *Mathematical Programming* **47**. 407 - 423.

Gershkoff I (1998). *An approach for JIT airline scheduling*. In: G Yu (editor). OR in the airline industry. Kluwer Academic Publishers. Boston, US. chapter 6, pp. 158–187.

Gotz S, Grothklags S, Kliewer G and Tschoke S (1999). *Solving the weekly fleet assignment problem for large airlines*. Proceedings of the III Metaheuristic International Conference.

Green T (2002). *Evaluating airline schedule for improved dependability*. AGIFORS Airline Operations Study Group. Rome, Italy.

Hane C, Barnhart C, Johnson E, Marsten R, Nemhauser G and Sigismondi G (1995). 'The fleet assignment problem: Solving a large-scale integer program'. *Mathematical Programming* **70**(2). 211 - 232.

Higle J (2005). *Stochastic Programming: Optimization When Uncertainty Matters*. In: Tutorials in Operations Research: Emerging Theory, Methods, and Applications. INFORMS Annual Meetings. H Greenberg and J Smith. Chapter 2. 30 - 53.

Higle J and Sen S (1991). 'Stochastic decomposition: An algorithm for two-stage linear programs with recourse'. *Mathematics of Operations Research* **16**. 650 - 669.

Higle J and Sen S (2001). *A stochastic programming model for network resource utilization in the presence of multi-class demand uncertainty*. In: S Wallace and W Ziembe (editors). Application of Stochastic Programming. MPS-SIAM Series on Optimization.

Huag S (1998). Airline delay pertubation under minor disturbance. Technical report. University of Califonia at Berkeley.

IATA (2007). *Passenger Forecast*. IATA. www.iata.org. Annual publication.

Jeschke O (2006). *Optimizing hub systems with operational constraints*. AGIFORS Strategic and Scheduling Study Group. Cranfield, UK.

Kall P and Mayer P (1996). 'Slp-ior: an interactive model management system for stochastic linear programs'. *Mathematical Programming: Series A and B archive* **75**(2), 221 – 240.

Kang L (2004). Degradable airline schedule. PhD thesis. Massachusetts Institute of Technology.

Kaut M and Wallace S (2007). 'Evaluation of scenario generation methods for stochastic programming'. *Pacific Journal of Optimization* **3**(2). Page 257-271.

Klabjan D (2005). *Large-scale models in the airline industry.* In: Desaulniers G, Desrosiers J and Solomon M (editors). Column Generation. Springer Verlag. New York, US.

Klophaus R and Polt S (2007). 'Airline overbooking with dynamic spoilage costs'. *Journal of Revenue and Pricing Management* **6**(1). 9 - 18.

Lan S, Barnhart C and Clarke J (2003). *New approaches to add robustness into airline schedule*. MIT open courseware. http://ocw.mit.edu/ocwWeb/Civil-and-Environmental-Engineering/index. Airline scheduling.

Lee K and Moore J (2003). *On-time performance*. AGIFORS Airline Operations Study Group. Auckland, New Zealand.

Lee L, Barnhart C and Tan Y (2003). *Improving robustness of flight schedule*. AGIFORS Scheduling and Strategic Planning Study Group. Toulouse, France.

Lister O and Dekker R (2002). 'A scenario aggregation based approach for determining a robust airline fleet composition'. *Econometric Institute Report* (E1). 17.

Lohatepanout M and Jacobs T (2003). *Quantifying resource impact on schedule profitability*. AGIFORS Scheduling and Strategic Planning Study Group. Toulouse, France.

Love M, Larsen A, Larsen J, Sorensen K and Clausen J (2002). *Disruptive management for an airline - rescheduling of aircraft*. In: R Hall (editor). Applications of evolutionary computing. Springer Berlin. pp. 289–300.

Lufthansa (2006). *NetLine-Plan and Netline-Sched Systems*. User's Manual.

Luo S and Yu G (1998). *Airline schedule pertubation problem*. In: G Yu (editor). OR in the airline industry. Kluwer Academic Publishers. Boston, US. chapter 14-15, pp. 404–459.

Mathaisel D (1981). Fleet planning for the air transport industry: network aggregation and cell theory models. PhD thesis. Massachsetts Institute of Technology.

Mathaisel D (2008). Computer models for airline planning and airline operations. Technical report. Massachsetts Institute of Technology. http://web.mit.edu. as browsed in January.

McGill J and Ryzin G (1999). 'Revenue management: Research overview and prospects'. *Transportation Science* **33**(2). 233 - 256.

Mederer M and Frank M (2002). *Increasing robustness of flight schedules through stochastic modeling of planning parameters*. AGIFORS Airline Operations Study Group. Rome, Italy.

Mitra G (2001). Optimum decision making under uncertainty: background and overview. Technical report. Brunel University: Centre for the analysis and optimisation modelling applications (CARISMA). PhD course notes.

OptiRisk (2008*a*). *FortMP*. http://www.optirisk-systems.com. as accessed in April.

OptiRisk (2008*b*). *SPInE*. http://www.optirisk-systems.com. as accessed in April.

OptiRisk (2009). *FortSp*. http://www.optirisk-systems.com. as accessed in July.

Pbuero D (2001). *Dealing with operational irregularities in airline schedules*. AGIFORS Symposium Proceeding. Sydney, Australia.

Pilla V, Rosenberger J, Chen V and Smith B (2005). A statistical computer experiments approach to airline fleet assignment. Technical report. The University of Texas at Arlington. http://ieweb.uta.edu/TechReports/Cosmos05-03pdf. Page 6-7.

Poojari C, Lucas C and Mitra G (2007). 'Robust solutions and risk measures for a supply chain planning problem under uncertainty'. *Journal of Operations Research Society* **62**(10). Page 50-70.

Rexin B, Barnhart C, Knicker C, Jarrah T and Krishnamurthy N (2000). 'Airline fleet assignment with time windows'. *Transportation Science* **34**(1). 1-20.

Ronen D (2000). 'Scheduling charter aircraft'. *Journal of Operations Research Society* **51**(3). 258-262.

Rosenhead J (2002). 'Forum: About robustness analysis'. *EWG-MCDA Newsletter* **Series 3**(10). http://inescc.pt/ewgmcda/Articles.html.

Rushmeier R and Kontogiorgis S (1997). 'Advances in the optimization of airline fleet assignment'. *Transportation Science* **31**(2). 159 - 169.

Seth C, Clarke M and Mathaisel D (1998). *The processes of airline system operations control*. In: G Yu (editor). OR in the airline industry. Kluwer Academic Publishers. Boston, US. chapter 14-15, pp. 312–369.

Sherali A, Bish E and Zhu X (2006). 'Airline fleet assignment concepts, models, and algorithms'. *European Journal of Operational Research* **172**(1). 1 - 30.

Slyke R and Wets R (1969). 'L-shaped linear programs with application to optimal control and stochastic programming'. *SIAM Journal on Applied Mathematics* **58**. 638 - 663.

Smith B (2004). Robust airline fleet assignment. PhD thesis. Georgia Institute of Technology.

Sriram C and Haghani A (2003). 'An optimization model for aircraft maintenance scheduling and re-assignment'. *Transportation Research: Policy and Practice* **37**(1). 29 - 48.

Stone R (1998). *Topics in airline planning*. http://ima.umn.edu/industrial.

Swan M (1999). 'Spill modeling for airlines'. 8th World Conference on Tranport Research Proceedings. Page 225 - 237.

Swan M (2001). Spill modeling for airlines. Technical report. Seabury Airline Planning Group. http://www.seaburyapg.com.

Sylla A (2000). Operations research in the airline industry. Technical report. University of Oulu. http://www.oulu.fi/seydou/aboud.pdf. Page 49-50.

Thengvall B, Bard J and Yu G (2000). 'Balancing user preferences for aircraft schedule recovery during irregular operations'. *IIE Transactions* (32). pp. 181-193.

Valente P, Mitra G and Poojari C (2005). *A Stochastic Programming Integrated Environment*. In: Application of Stochastic Programming. MPS-SIAM Series on Optimization. S Wallace and W Ziemba. Chapter 8. 115 - 136.

Vinke P (2003). 'Forum: About robustness analysis'. *EWG-MCDA Newsletter* **Series 3**(8). http://inescc.pt/ewgmcda/Articles.html.

Wu R and Reid C (2003). *Dream and reality: modeling of airline schedule operations and measuring ATM performance*. AGIFORS Airline Operations Study Group. Auckland, New Zealand.

Yu G (1998). *OR in the airline industry*. Kluwer Academic Publishers. Boston, US. Page XV-XVII.

Yu G and Yang J (1998). *Optimisation application in the airline industry*. Handbook of Combinatorial Optimization. Kluwer Academic Publishers. Du and P.M. Pardalos. Volume 2, Page 3.

Zhao Q (2003*a*). *A multi-criteria optimization approach for building robust aircraft rotations*. AGIFORS Scheduling and Strategic Planning Study Group. Toulouse, France.

Zhao Q (2003*b*). *A robust optimization approach for aircraft schedule dependability*. AGIFORS Airline Operations Study Group. Auckland, New Zealand.