Integrating the Fleet Assignment Model with Uncertain Demand

Anwar H. Ahmed

Centre for the analysis of Risk and optimisation modelling applications (CARISMA), School of Information Systems, Computing and Mathematics, Brunel University, London

Abstract

Fleet assignment models (FAM) are used by many airlines to assign aircraft to flights in a schedule to maximize profit. We first state the basic FAM before proposing new FAMs (i.e. DFAM1 and DFAM2) that tackles the common problem associated with aircraft utilization. Through the use of a *two-stage Stochastic Programming (SP) with recourse* technique, the DFAMs are extended to SP-FAMs. 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 outline the SP-FAM solution process and 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.

1 Introduction

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.

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 questionable.

2 Basic FAM Formulation

The fleet assignment problem (Hane et al., 1995) 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_j. The event at time t_j occurs before the event at time, t_{j+1}. The last node in the counting line is denoted by t_n. The index t^+ denotes an event after the current time t_j while t^- denotes an event before the current time t_j ($t_j \in T$).
R: the set of nodes in the network indexed by $\{k,o,t_j\}$;
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 $\{k,o,t_j\}$; and
O(k,o,t) is the set of outbound flight legs from node $\{k,o,t_j\}$.

Decision variables

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

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

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

$$Min \sum_{k\in K}\sum_{i\in L} c_{k,i} f_{k,i} \qquad (1)$$

$$subject \ \ to:$$

$$\sum_{k\in K} f_{k,i} = 1 \quad \forall i \in L \qquad (2)$$

$$Y_{k,o,t^-} + \sum_{i\in I(k,o,t)} f_{k,i} - Y_{k,o,t^+} - \sum_{i\in O(k,o,t)} f_{k,i} = 0 \quad \forall\{k,o,t\} \in R \qquad (3)$$

$$\sum_{o\in O} Y_{k,o,t_n} + \sum_{i\in CL(k)} f_{k,i} \leq N_k \quad \forall k \in K \qquad (4)$$

The objective function (1) is to minimize the cost of the fleet assignment; the *cover* constraint (2) ensures that each flight is covered once and only once by a fleet type. Equation 3 is 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* constraint (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.

3 FAM with Aircraft Utilization

A *tactical FAM* is identical to the *basic FAM* discussed earlier except for the modification of the objective function (equation 5) 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.

$$Min \sum_{k\in K}\sum_{i\in L}(c_{k,i} - X_{k,i} \times q_i) \times f_{k,i} \tag{5}$$

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 buffer. 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 6.

$$N_k \times MN_k \leq \sum_{i \in L} f_{k,i} \times B_{k,i} \leq N_k \times MX_k \quad \forall k \tag{6}$$

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 extend as opposed to getting confined within the threshold boundary. If the MN_k value is made smaller (or zero), the unbounded 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 11) 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.

$$Min \sum_{k \in K} \sum_{i \in L} (c_{k,i} - X_{k,i} \times q_i) \times f_{k,i} + \sum_{k \in K} rp_k \times w_k \tag{7}$$

$$subject \;\; to:$$

$$\sum_{k \in K} f_{k,i} = 1 \;\; \forall i \in L \tag{8}$$

$$Y_{k,o,t^-} + \sum_{i \in I(k,o,t)} f_{k,i} - Y_{k,o,t^+} - \sum_{i \in O(k,o,t)} f_{k,i} = 0 \;\; \forall \{k,o,t\} \in R \tag{9}$$

$$\sum_{o \in O} Y_{k,o,t_n} + \sum_{i \in CL(k)} f_{k,i} \leq N_k \;\; \forall k \in K \tag{10}$$

$$rp_k - rm_k = N_k \times MN_k - \sum_{i \in L} f_{k,i} \times B_{k,i} \;\; \forall k \in K \tag{11}$$

Note that the variable splitting technique depicted in equation 11 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).

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 the MN_k value. Using this representation, constraints 8 - 10 will still remain intact as in *tactical* FAM, while the objective function and the utilization constraint will change. The model will be represented as:

$$Min \sum_{k\in K}\sum_{i\in L}(c_{k,i} - X_{k,i} \times q_i) \times f_{k,i} + \sum_{k\in K} r_k \times w_k \quad (12)$$

$$subject \;\; to:$$

$$\sum_{k\in K} f_{k,i} = 1 \;\; \forall i \in L \quad (13)$$

$$Y_{k,o,t^-} + \sum_{i\in I(k,o,t)} f_{k,i} - Y_{k,o,t^+} - \sum_{i\in O(k,o,t)} f_{k,i} = 0 \;\; \forall \{k,o,t\} \in R \quad (14)$$

$$\sum_{o\in O} Y_{k,o,t_n} + \sum_{i\in CL(k)} f_{k,i} \leq N_k \;\; \forall k \in K \quad (15)$$

$$r_k = N_k \times MX_k - \sum_{i\in L} f_{k,i} \times B_{k,i} \;\; \forall k \in K \quad (16)$$

4 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:

$$Z_{hn} \qquad min\ cx + E_\omega(q(\omega)y(\omega)) \tag{17}$$
$$\begin{aligned} subject\ \ to:& \\ Ax &\geq b \\ B(\omega)x + D(\omega)y(\omega) &\geq h(\omega) \\ x \geq 0, y(\omega) &\geq 0. \end{aligned}$$

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

$$Z_{hn} \qquad min\ cx + E_\xi Q(x, \xi(\omega)) \tag{18}$$
$$\begin{aligned} subject\ \ to:& \\ Ax &\geq b \\ x &\geq 0 \end{aligned}$$

$$where \quad Q(x,\xi(\omega)) = min\ \ q(\omega)y(\omega) \tag{19}$$
$$\begin{aligned} subject\ \ to:& \\ B(\omega)x + D(\omega)y(\omega) &\geq h(\omega) \\ y(\omega) &\geq 0 \end{aligned}$$

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

$$Z_{hn} \qquad min\ cx + \sum_{s=1}^{|S|} p_s q_s y_s \tag{20}$$
$$\begin{aligned} subject\ \ to:& \\ Ax &\geq b \\ B_s x + D_s y_s &\geq h \quad \forall\ s \in S \\ x \in \Re^{n_1}, y_s \in \Re^{n_2}, y_s &\geq 0 \end{aligned}$$

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.

5 The SP-FAM

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. The *two-stage recourse FAM* is an extension of the deterministic equivalent models 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 a concept of standby aircraft that can be assigned in the presence of demand uncertainty.
- Parameters: An additional standby (spare or buffer) aircraft is introduced that attracts a penalty cost $j_{k,i}$. Unlike the other scheduled aircraft, the buffer 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 25 and 31).

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

The SP-FAM1

$$Min\ \ C_{U1} + Z_{AP} + Z_{VP} \tag{21}$$

$Subject\ \ to:$

$$\sum_{k \in K} f_{k,i} = 1 \quad \forall i \tag{22}$$

$$Y_{k,o,t^-} + \sum_{i \in I(k,o,t)} f_{k,i} = Y_{k,o,t^+} + \sum_{i \in O(k,o,t)} f_{k,i} \quad \forall \{k,o,t\} \in R \tag{23}$$

$$\sum_{o \in O} Y_{k,o,t_n} + \sum_{i \in CL(k)} f_{k,i} \leq N_k \quad \forall k \tag{24}$$

$$(f_{k,i} + b_{k,i,s}) \times seats_k \leq D_{k,i,s} \quad \forall k,i,s \tag{25}$$

$$rp_k - rm_k = N_k \times MN_k - \sum_{i \in L} f_{k,i} \times B_{k,i} \quad \forall k \tag{26}$$

The SP-FAM2

$$Min\ \ C_{U2} + Z_{AP} + Z_{VP} \tag{27}$$

$Subject\ \ to:$

$$\sum_{k \in K} f_{k,i} = 1 \quad \forall i \tag{28}$$

$$Y_{k,o,t^-} + \sum_{i \in I(k,o,t)} f_{k,i} = Y_{k,o,t^+} + \sum_{i \in O(k,o,t)} f_{k,i} \quad \forall \{k,o,t\} \in R \tag{29}$$

$$\sum_{o \in O} Y_{k,o,t_n} + \sum_{i \in CL(k)} f_{k,i} \leq N_k \quad \forall k \tag{30}$$

$$(f_{k,i} + b_{k,i,s}) \times seats_k \leq D_{k,i,s} \quad \forall k,i,s \tag{31}$$

$$r_k = N_k \times MX_k - \sum_{i \in L} f_{k,i} \times B_{k,i} \quad \forall k \tag{32}$$

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_j. The event at time t_j occurs before the event at time, t_{j+1}. The last node in the counting line is denoted by t_n. The index t^+ denotes an event after the current time t_j while t^- denotes an event before the current time t_j ($t_j \in T$);
R: the set of nodes in the network indexed by $\{k,o,t_j\}$;
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 $\{k,o,t_j\}$; and
O(k,o,t) is the set of outbound flight legs from node $\{k,o,t_j\}$.

Decision variables

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

$Y_{k,o,t_j} \geq 0$ is the number of aircraft of fleet type k, on the ground at station o, and time t_j; Y_{k,o,t^+} is the number of aircraft of fleet type k, on the ground at station o, just following time t_j ($t_j \in T$); 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_j ($t_j \in T$);
$b_{k,i,s}$ 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 buffer 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 buffer 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. buffer 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 (21 and 27) minimizes the fleet assignment cost (it is a re-expression of maximization of fleet assignment profitability). The *cover* constraints (22 and 28) ensures that each flight is covered once and only once by a fleet type. Equation 23 and 29 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 (24 and 30) ensures that only the number of available aircraft are used. The *demand-spill* constraints (25 and 31) 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 (26 and 32) computes the number of utilized block-time on aircraft of type k.

5.1 The Buffer Aircraft

The distinct feature of the SP formulation is the buffer aircraft that is reassigned in the case of demand fluctuation but at the cost of the original scheduled aircraft. The buffer aircraft concept is used widely by both charter operators (Ronen, 2000) and scheduled operators. Ideally, rather than having a buffer 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 buffer when on ground and not immediately scheduled. One way of accounting for buffer 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 buffer aircraft and an utilization (idle) cost for making the original scheduled aircraft under-utilized.

5.2 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}). 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 buffer aircraft penalty cost ($j_{k,i}$). $c_{k,i}$ is incurred by both scheduled and buffer aircraft while $j_{k,i}$ is exclusive for buffer 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
- $Z_{TP} = Z_{RP} + Z_{VP}$

5.3 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 utilization aircraft cost (C_{U1} or C_{U2}) and the virtual profit (i.e. Z_{VP}) that are tied to the buffer aircraft swaps are the main underlying logics of the SP-FAMs formulation. The major problem with FAM solutions, and especially for a large network, is the generation of infeasible solutions. As in real life the artificial variable does not have much meaning, to generate a feasible optimal solution, $b_{k,i,s}$ is relaxed (i.e. ≥ 0).

5.4 Demand-Spill Constraint

The demand-spill constraint (25 or 31) 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 three 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 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 1
Passenger demand scenarios.

Table 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 2
Demand-spill constraint.

5.5 Aircraft Utilization Constraint

The idea of the cost function C_{U1} and constraint 26 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 buffer 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 buffer 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 32 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 buffer 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 3 is in reference to *SP-FAM1* while Table 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.

5.6 Aircraft Utilization Cost

If we further assume that the utilization variable has a corresponding utilization cost as shown in column **(e)** of Table 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 buffer 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

(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 £
$f_{A1,X-Y5} = 1$	2				
$f_{A1,Y5-X} = 1$	2	5	1	1000	1000
$f_{A1,X-Y2} = 1$	4				
$f_{A1,Y2-X} = 1$	4	7	-1	2000	0
$f_{A3,X-Y3} = 1$	6				
$f_{A3,Y3-X} = 1$	6				
$f_{A3,X-Y4} = 1$	5				
$f_{A3,Y4-X} = 1$	5	10	4	1500	4000
$f_{B4,X-Y1} = 1$	4				
$f_{B4,Y1-X} = 1$	4	8	0	0	0

Table 3
Constraint 26 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 £
$f_{A1,X-Y5} = 1$	2				
$f_{A1,Y5-X} = 1$	2	7	3	1000	3000
$f_{A1,X-Y2} = 1$	4				
$f_{A1,Y2-X} = 1$	4	9	1	2000	2000
$f_{A3,X-Y3} = 1$	6				
$f_{A3,Y3-X} = 1$	6				
$f_{A3,X-Y4} = 1$	5				
$f_{A3,Y4-X} = 1$	5	12	14	1500	21000
$f_{B4,X-Y1} = 1$	4				
$f_{B4,Y1-X} = 1$	4	14	6	0	0

Table 4
Constraint 32 and function C_{U2}.

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 castigated down on hourly (or minutes) rate. Thereafter, the values are adjusted further based on the desired utilization balanced required for each aircraft type. The stabilized values could then be used in solving the SP-FAM problem.

5.7 The Buffer Aircraft Penalty Cost

The aircraft penalty cost is associated with the buffer 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 buffer 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 buffers are either the assigned aircraft in circulation or on standby, 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.

6 The Scenario Generator

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 1, depicts a typical relationship between the two.

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

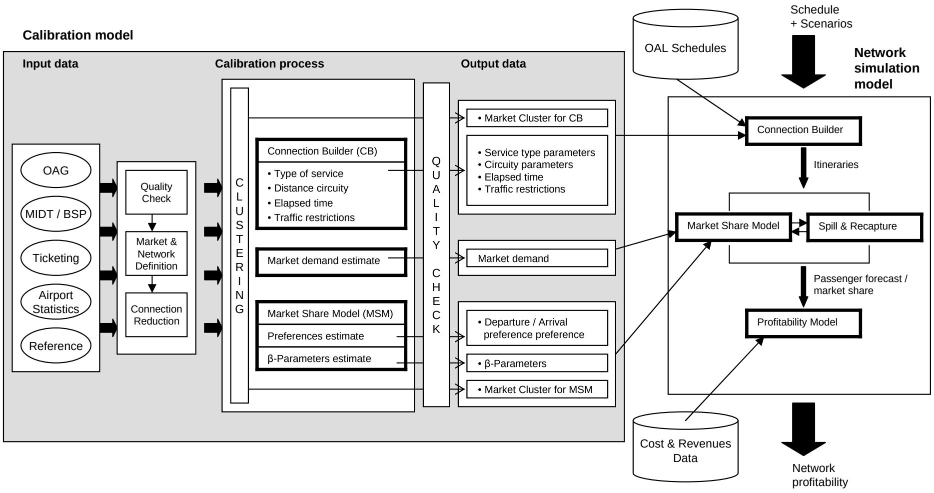

Fig. 1. Calibration process and network-simulation model.

from the former to generate itineraries for the CB and passenger forecast for the MSM.

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.

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 2.

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 Sea-

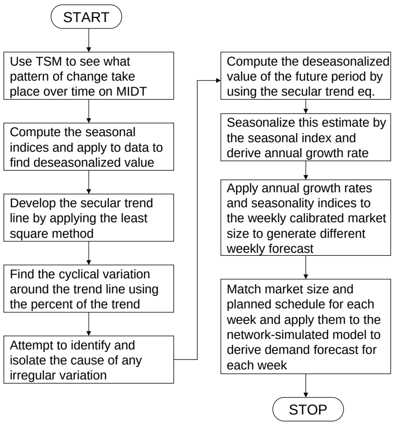

Fig. 2. TSM flow chart.

sonality 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.

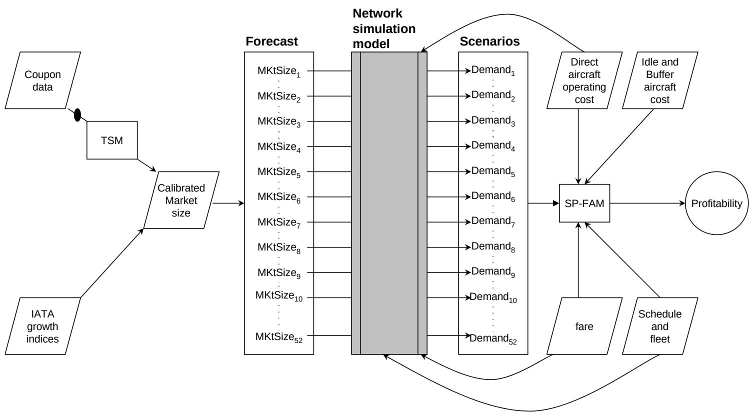

Fig. 3. Overview of the solution approach.

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 3 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;

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

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 *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

7 SP-FAM Solution Process

In solving the SP-FAM problem, we follow a series of steps as described in Figure 4. 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.

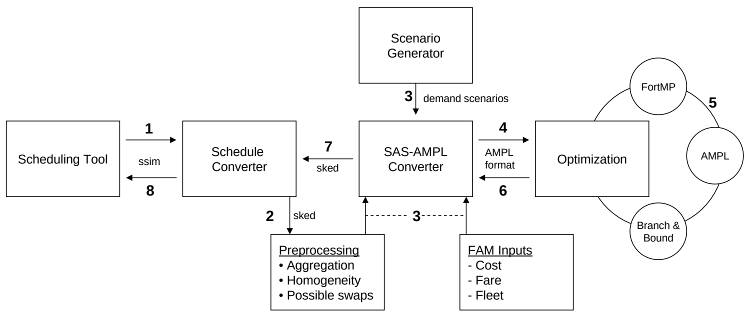

Fig. 4. SP-FAM solution approach.

Schedule Aggregation
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 5. 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.

Homogeneity
Assume the input schedule before optimization is as shown in *part A* of Figure 5. 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 6. Note that the optimized schedule and the corresponding assigned

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					

Fig. 5. Before and after schedule aggregation.

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.

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					

Fig. 6. Before and after defining fleet homogeneity.

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*.

Possible Swaps

Determining the possible swaps, homogeneity and aggregation steps are executed concurrently. *Part D* of Figure 6 could be aggregated as shown in *part F* of Figure 7. 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.

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

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					

Fig. 7. Aggregated swap enumeration.

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 8 shows the interactive shell-level modelling environment of AMPL.

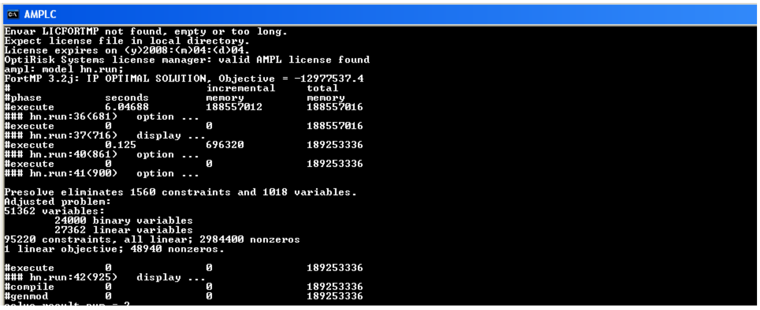

Fig. 8. AMPL Shell.

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 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, we would have first needed to represent the problem in SMPS format that could be acceptable to other solvers. 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.

8 Case Study

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 buffers (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 buffer 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 4.6*.
- 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 (that is 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}. The objective value of the two-stage SP with recourse model is denoted by Z_{hn} (where HN stand for here-and-now). Through these two models we construct the *Expected Value 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 SP-FAM Statistics

Table 5 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 5
Here-and-now model statistics

The results at Table 5 indicate that for a given high computing processing power and memory, solving such a large-scale problem is highly efficient.

8.2 SP-FAM Stochastic Measures

Table 6, 7 and 8 shows the objective values given in UK £ of the various SP models for SP-FAM1 and SP-FAM2, respectively.

Table 6 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. 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.

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 7. For our purpose, we consider the Z_{ev} to be the same as shown in Table 6. Table 7 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. Table 8 portrays similar conclusions as in Table 7 but in addition,

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	18,903,127	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 6
Z_{TP}: Objective values and SP measures

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	∞	129,734	∞
125	1,567,091	795,084	948,529	∞	153,446	∞

Table 7
Z_{RP}: Objective values and SP measures

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.

8.3 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 and solving the optimization problem on each of these scenario sets, the optimal objectives are similar. A scenario-generator is said to manifest out-of-sample stability, if, when generating several scenario sets of the same size and solving the optimization model on each of these scenario sets, the optimal solutions obtained yield similar *true* objective function values (i.e. the solutions obtained are evaluated on the *true* distribution of the random variables involved).

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
Z_{AP}: Objective values and SP measures

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, we use SP-FAM2 to investigate the stability of the scenario-generator .

In-Sample Stability Measure
When testing the in-sample stability, the same scenario set is used for both making a decision and evaluating it. The in-sample stability is measured as the difference of the objective functions obtained by solving the same problem with different scenario trees (of the same size) generated with the same method. In other words, the scenario generation method is compared with itself. 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.

- *Min*: Min represents the minimum objective value of all the simulation runs (in our case 100).
- *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.

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 9
In-sample stability.

- *RMnD*: Represents the *relative mean deviation* and is expressed by the fraction between the Stdev and the Mean (i.e. $RMnD = \frac{Stdev}{Mean}$).

We observe that the value given by the RMnD 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.

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 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:

$$o_1....o_n = F(x_1^*, \bar{\xi})....F(x_n^*, \bar{\xi}) \tag{33}$$

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

$$\bar{d} = \frac{1}{(n^2 - n)/2} \sum_{i=1}^{n} \sum_{j=i+1}^{n} |O_i - O_j| \tag{34}$$

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):

$$s = 1 - \frac{\bar{d}}{\mu} \tag{35}$$

Figure 9 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.

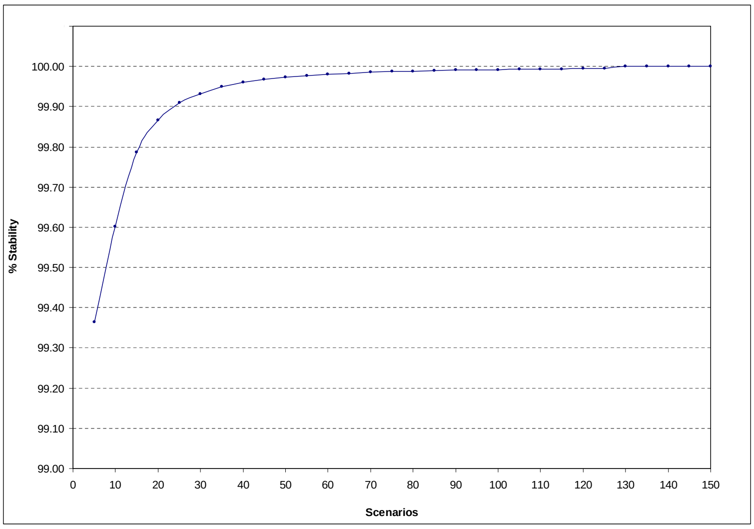

Fig. 9. Out-of-sample stability.

9 Future Work

Robustness within FAM remains a vibrant research area with many challenges on model formulation, accurate passenger forecast and integration of such models.

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.

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.

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; 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

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.

References

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.

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.

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.

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

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

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.

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.

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

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