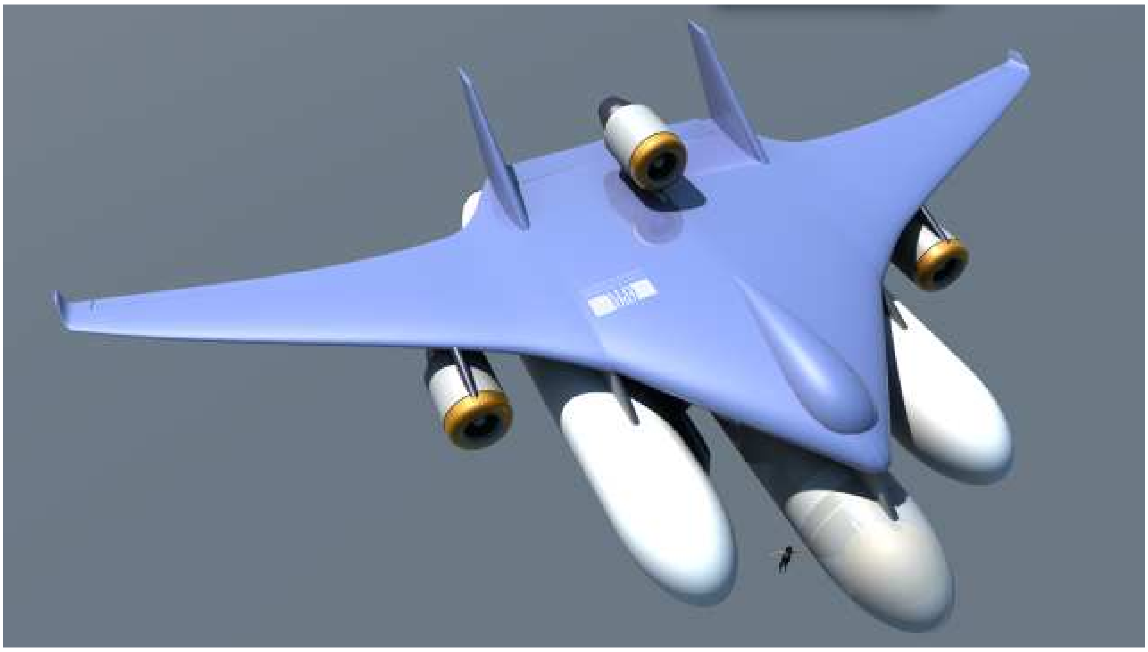

An integrated fleet assignment and itinerary choice model for a new flexible aircraft

Bilge Atasoy

Matteo Salani

Michel Bierlaire

STRC 2012 **May 2012**

TRANSP-OR

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

IDSIA

STRC 2012

An integrated fleet assignment and itinerary choice model for a new flexible aircraft

Bilge Atasoy
TRANSP-OR
ENAC
EPFL
Lausanne
phone: +41 21 693 93 29
fax: +41 21 693 80 60
bilge.kucuk@epfl.ch

Matteo Salani
IDSIA
USI/SUPSI
Manno-Lugano
phone: +41 58 666 66 71
fax: +41 58 666 66 61
matteo.salani@idsia.ch

Michel Bierlaire
TRANSP-OR
ENAC
EPFL
Lausanne
phone: +41 21 693 25 37
fax: +41 21 693 80 60
michel.bierlaire@epfl.ch

May 2012

Abstract

In this paper we present methodologies for improving the demand-responsiveness of air transportation systems. The main ingredients are the flexibility in transportation capacity provided by an innovative aircraft and an integrated model where supply-demand interactions are explicitly formulated. The integrated model benefits from the simultaneous schedule planning and revenue management decisions. The schedule planning consists of schedule design and fleet assignment models. Revenue management decisions are integrated with an itinerary choice model which gives the market shares of the available itineraries in the market according to their price, travel time, number of stops and departure time of the day. The integrated model also includes spill and recapture effects based on the demand model. Furthermore, the demand model is developed for economy and business classes and the seat allocation for these classes is determined by the integrated model. The resulting model is a mixed integer nonlinear problem and we propose a heuristic to tackle with the complexity of the problem.

Keywords

Fleet assignment, supply-demand interactions, integrated schedule planning, discrete choice modeling, itinerary choice, revenue management, spill and recapture, mixed integer nonlinear problem

1 Introduction and Motivation

The increase in air travel demand in the last decades results with frequent delays and cancellations of flights. In such an environment it is difficult to be demand responsive for airlines. We believe that to tackle with the shortcomings of the current air transportation system actions need to be taken from both supply and demand sides. In this study we address improvements in both dimensions. We study the supply side by developing appropriate models for an innovative flexible aircraft. When it comes to modeling demand, we integrate an itinerary choice model into the scheduling model in order to define supply-demand interactions. The objective of this study is to identify the challenges in integrating demand and supply models and develop appropriate methodologies.

A new flexible air transportation concept, called Clip-Air, is developed at EPFL. Clip-Air's flexibility is mainly provided by the detachable load units, *capsules*. The capsules can be detached from the carrying unit, *wing*. This decoupling brings in many advantages in terms of airline and airport operations. In terms of modeling, Clip-Air necessitates two level of fleet assignments for the decoupled units. Therefore we adapt our fleet assignment model to appropriately represent the flexibility of Clip-Air. In order to quantify the potential advantages of Clip-Air we build models for both standard planes and Clip-Air capsules and wings. We refer to Atasoy *et al.* (2011) for a preliminary analysis on the potential performance of Clip-Air in comparison to the existing aircraft.

The focus of this paper is the integrated schedule planning and revenue management model. The schedule planning model is an integrated schedule design and fleet assignment model. Schedule design decision is included with the existence of an optional set of flights that can be canceled. The revenue management decision includes the decisions on the pricing, spill and recapture as well as the seat allocation for economy and business classes. Revenue management is based on an itinerary choice model. The itinerary choice is modeled as a logit formulation using a joint revealed preferences (RP) and stated preferences(SP) data. RP data is a booking data provided by a major European airline. RP data has low variability due to the absence of non-chosen alternatives. Therefore we use the SP data to benefit from its elasticity that is ensured by the design. At the end we use the model for RP data in the optimization. The logit model includes the variables of price and time interacted with the number of stops; and the departure time of the day.

The added-value of the integrated model is illustrated with a set of experiments. However, the integrated model is a mixed integer nonlinear problem where the convexity is not guaranteed. Therefore, we are able to solve medium sized instances with available solvers in reasonable time. In order to overcome these limitations, we propose an heuristic which works on a simplified model and explores the feasible region with price sampling and variable neighborhood search techniques. We provide results on the performance of the heuristic and discuss potential improvements based on Lagrangian relaxation and subgradient optimization.

2 Related literature and the contributions of the paper

In this section we focus on the closely related literature in terms of the demand modeling, integrated schedule planning, revenue management and solution methodologies.

Itinerary choice models have been studied in the literature, with an increasing interest in the last decade, as a more appropriate tool for demand forecasting compared to the classical models. We refer to Garrow (2010) where the motivation for the usage of discrete choice methodology in air travel demand is presented together with several case studies. Various specifications are provided such as logit and nested logit models.

The schedule planning model we consider in this study is inspired by the work of Lohatepanont and Barnhart (2004). They present an integrated schedule design and fleet assignment model where they include spill and recapture effects based on the Quality Service Index (QSI). They take the price and demand values as inputs to the model. We refer to this model as *price-inelastic schedule planning model*. The integrated model presented in this paper considers explicit supply-demand interaction due to the integration of the demand model. Therefore the integrated model is elastic to the price and other attributes of the itineraries in the market. In section 5 we compare the integrated model to the price-inelastic schedule planning model, in order to show the impacts of the integration of the demand model. Since we do not have access to the parameters of the recapture ratios that Lohatepanont and Barnhart (2004) use, we utilize our demand model to estimate the recapture ratios between itineraries. Sherali *et al.* (2010) also present an integrated schedule design and fleet assignment model where they work with itinerary-based demands for multiple fare classes. They optimize the allocation of seats for each fare class as we do in our integrated model. However they do not include supply-demand interactions in the model.

In terms of the integration of discrete choice models in revenue management, we refer to the work of Talluri and van Ryzin (2004a) who introduce a revenue management model based on a discrete choice methodology. They decide on the subset of fare products to offer at each point in time according to the discrete choice model. They consider single-leg, multiple-fare-class products. Schön (2008) presents an integrated schedule design, fleet assignment and pricing model which is similar to our idea. She provides different specifications of the demand model as logit and nested logit where the only explanatory variable is the price. However, she does not consider spill and recapture effects and she provides results based on a synthetic data.

In classical revenue management models the capacity is considered as a fixed input which is assumed to be obtained by the schedule plan (Talluri and van Ryzin, 2004b). We refer to this common practice as *sequential approach*. Our integrated model decides on the capacity and the demand sides simultaneously. In order to see the impact of this simultaneous optimization, we compare our model with the sequential approach in section 5

The presented model in this paper is a mixed integer nonlinear problem (MINLP) where we can not guarantee the convexity. For the difficulties in MINLP and the review of available methodologies we refer to D'Ambrosio and Lodi (2011).

3 Demand model

We develop an itinerary choice model in order to explicitly integrate supply-demand interactions in the schedule planning model. *Itinerary* is referred as each available product, which may include more than one flight leg, for a market segment. The market segments , $s \in S^h$, are defined by the origin and destination (OD) pairs where h represents the cabin class: economy and business. The choice situation is defined for each segment s with a choice set of all the alternative itineraries in the segment represented by I_s. The index i for each alternative itinerary in segment I_s carries the information of the cabin class of the itinerary due to the definition of the segments. In order to better represent the reality, we include *no-revenue options* ($I_s^{'} \subset I_s$), which represent the itineraries offered by competitive airlines.

The utility of each alternative itinerary i, including the no-revenue options, is represented by V_i and the specification is provided in Table 1. The alternative specific constants, ASC_i, are included for each itinerary in each segment except one of them which is normalized to 0 for identification purposes. Other parameters are represented by β for each of the explanatory variables. We have different models for economy and business classes. The superscript E indicates the model for economy itineraries and the parameters with B represent the model for business itineraries. The superscripts NS and S are used to indicate whether the itinerary is a non-stop or a one-stop itinerary. The explanatory variables are described as follows:

- p_i is the price of itinerary i in €, which is normalized by 100 for scaling purposes,
- $time_i$ is the elapsed time for itinerary i in hours,
- $non\text{-}stop_i$ is a dummy variable which is 1 if itinerary i is a non-stop itinerary, 0 otherwise,
- $stop_i$ is a dummy variable which is 1 if itinerary i is a one-stop itinerary, 0 otherwise,
- $economy_i$ is a dummy variable which is 1 if itinerary i is an economy itinerary, 0 otherwise,
- $business_i$ is a dummy variable which is 1 if itinerary i is a business itinerary, 0 otherwise,
- $morning_i$ is a dummy variable which is 1 if itinerary i is a morning itinerary departing between 07:00-11:00, 0 otherwise. The time slot is inspired by the studies in literature

that show that the individuals have higher utility for the departures in this slot(Garrow, 2010).

Table 1: Specification table of the utilities

	Parameters	Explanatory variables
constants	ASC_i^E	$1 \times \text{economy}_i$
	ASC_i^B	$1 \times \text{business}_i$
price	$\beta_{\text{p}}^{E,NS}$	$\ln(\text{p}_i/100) \times \text{non-stop}_i \times \text{economy}_i$
	$\beta_{\text{p}}^{B,NS}$	$\ln(\text{p}_i/100) \times \text{non-stop}_i \times \text{business}_i$
	$\beta_{\text{p}}^{E,S}$	$\ln(\text{p}_i/100) \times \text{stop}_i \times \text{economy}_i$
	$\beta_{\text{p}}^{B,S}$	$\ln(\text{p}_i/100) \times \text{stop}_i \times \text{business}_i$
time	$\beta_{\text{time}}^{E,NS}$	$\text{time}_i \times \text{non-stop}_i \times \text{economy}_i$
	$\beta_{\text{time}}^{B,NS}$	$\text{time}_i \times \text{non-stop}_i \times \text{business}_i$
	$\beta_{\text{time}}^{E,S}$	$\text{time}_i \times \text{stop}_i \times \text{economy}_i$
	$\beta_{\text{time}}^{B,S}$	$\text{time}_i \times \text{stop}_i \times \text{business}_i$
time-of-day	$\beta_{\text{morning}}^{E}$	$\text{morning}_i \times \text{economy}_i$
	$\beta_{\text{morning}}^{B}$	$\text{morning}_i \times \text{business}_i$

As seen in Table 1, the time and price variables are interacted with the number of stops, i.e. the dummies of *non-stop* and *stop*, The motivation behind this interaction is that there are strong correlations between the number of stops and both the time and price of the itinerary. The one-stop itineraries have longer travel time and usually more expensive compared to non-stop itineraries. We specify the price variable as a log formulation since it improves the model significantly. The idea behind is that, the effect of the increase in price is not linear for a low price itinerary and a high price itinerary.

The explanatory variables include the price, p_i, as a policy variable which can be controlled by the integrated model in order to increase the profit. The other explanatory variables are context variables which we denote by the vector z_i. These context variables provide information for the demand and improves the estimation of the market shares but can not be modified by the integrated model. In order to explicitly represent these variables we refer to the utilities V_i as $V_i(p_i, z_i; \beta)$.

The resulting logit model gives the choice probability for each itinerary i in segment s and when multiplied with the total expected demand of the segment, D_s, it provides the estimated demand of each itinerary as represented by equation 1.

$$\tilde{d}_i = D_s \frac{\exp\left(V_i(p_i, z_i; \beta)\right)}{\sum_{j \in I_s} \exp\left(V_j(p_j, z_j; \beta)\right)} \qquad \forall h \in H, s \in S^h, i \in I_s \tag{1}$$

The logit model is also used to model the spill and recapture effects. Passengers, who can not fly on their desired itineraries, may accept to fly on other available itineraries in the same market segment in case of such shortages. Airlines can take advantage of this knowledge when planning for the schedule and the design of fleet capacity. They can keep their capacity at profitable levels by taking into account the possibility of redirecting passengers to the alternative itineraries. We assume that the spilled passengers are recaptured by the other itineraries with a recapture ratio based on the logit formulation. Therefore the recapture ratio is represented by equation (2).

$$b_{i,j} = \frac{\exp\left(V_j(p_j, z_j; \beta)\right)}{\sum\limits_{k \in I_s \setminus \{i\}} \exp\left(V_k(p_k, z_k; \beta)\right)} \quad \forall h \in H, s \in S^h, i \in (I_s \setminus I_s^{'}), j \in I_s. \tag{2}$$

The recapture ratios $b_{i,j}$ represent the proportion of recaptured passengers by itinerary j among $t_{i,j}$ number of spilled passengers from itinerary i. The recapture ratio is calculated for the itineraries that are in the same market segment where the desired itinerary i is excluded from the choice set. Therefore lost passengers may be recaptured by the remaining alternatives of the company or by the no-revenue options.

For the estimation of the demand model we use an RP data provided in the context of ROADEF Challenge 2009[1]. This is a booking data from a major European airline which provides the set of airports, flights, aircraft and itineraries. The information provided for the itineraries includes the corresponding flight legs therefore we can deduce the information on the departure and arrival time of itinerary, the trip length and the number of stops. Additionally, we have information on the demand and average price (€) for each cabin class. Since the RP data does not include non-chosen alternative we have lack of variability in some attributes. This results with statistically insignificant estimation of key parameters of the choice models. Therefore, in this study we combine the RP data with an SP data, where the variability is obtained by design. This SP data is based on an Internet choice survey collected in 2004 in the US. The Internet survey was organized to understand the sensitivity of air passengers to the attributes of an airline itinerary such as fare, travel time, number of stops, legroom, and aircraft. By design, the data has enough variability in terms of price and other variables. For the estimation, the parameters of the logit model corresponding to the RP data are constrained to be the same as those of the SP data. The estimation of the two logit models for the two data sets is carried out simultaneously. For the details on the SP model and the simultaneous estimation we refer to Atasoy and Bierlaire (2012).

The estimation of the parameters is done with a maximum likelihood estimation using the software BIOGEME (Bierlaire and Fetiarison, 2009). The resulting parameters can be seen

[1] `http://challenge.roadef.org/2009/en`

Table 2: Estimated parameters for the model with joint RP and SP data

	β_{p}		β_{time}		
	non-stop	stop	non-stop	stop	β_{morning}
economy	-2.23	-2.17	-0.102	-0.0762	0.0283
business	-1.97	-1.96	-0.104	-0.0821	0.0790

in Table 2. The cost and time parameters have negative signs as expected since the increase in the price or the time of an itinerary decreases its utility. They also indicate that, economy demand is more sensitive to price and less sensitive to time compared to business demand as expected (Belobaba *et al.*, 2009). Departure time of the day parameter, $\beta_{morning}$, is higher for business demand compared to the economy demand, which means that business passengers have a higher tendency to chose morning itineraries.

The details on the demand model and results on the demand indicators such as the price and time elasticities as well as the willingness to pay are provided in Atasoy and Bierlaire (2012).

In order to illustrate the application of the demand model together with the spill and recapture effects we choose an arbitrary OD pair A-B. There are two alternatives of economy itineraries which are both nonstop itineraries with the same travel time. We include the no-revenue itinerary A-B$'$. The values of attributes can be seen in Table 3. According to the attributes the resulting choice probability, which is referred as the *market share*, is presented in the last column. The itinerary 2 has the lowest price and is a morning itinerary. Therefore it attracts the biggest number of passengers.

With the same example we illustrate the spill and recapture effects. The resulting ratios according to the given attributes are presented in Table 4. For example, in case of capacity shortage for itinerary 1, at most 55% of spilled passengers will be recaptured by itinerary 2 and 45% will be lost to the itineraries offered by competitive airlines. Since the price of itinerary 2 is lower than the price of competitors, the probability to be recaptured by itinerary 2 is higher.

Table 3: Attributes of the itineraries and the resulting market shares

OD	price	morning	market share
A-B$_1$	225	0	0.26
A-B$_2$	203	1	0.44
A-B$'$	220	0	0.30

Table 4: Resulting recapture ratios

	A-B$_1$	A-B$_2$	A-B$'$
A-B$_1$	0	0.552	0.448
A-B$_2$	0.487	0	0.513

4 Integrated schedule planning and revenue management model

We present an integrated schedule planning and revenue management model for a single airline. The schedule is based on a time-space network. The parameters of the model is provided in Table 5 and the decision variables of the model are presented in Table 6. We indicate the decision variables as schedule planning and revenue management variables for the ease of explanation. The mathematical formulation of the integrated model is given in Figure 1.

Table 5: Parameters of the integrated model

Set	Definition
F	the set of flight legs indexed by f
F_M	the set of mandatory flight legs
F_O	the set of optional flight legs
CT	the set of flights flying at count time
A	the set of airports indexed by a
K	the set of fleet types indexed by k
T	the set of time of the events in the network indexed by t
$N(k,a,t)$	the set of the nodes in the time-line network for fleet type k, airport a and time t
In(k,a,t)	set of inbound flight legs for node (k,a,t)
Out(k,a,t)	set of outbound flight legs for node (k,a,t)
H	set of cabin classes indexed by h
S^h	the set of market segments indexed by s, for cabin class h
I_s	the set of itineraries in segment s, indexed by i
$I_s^{'}$	the set of no-revenue itineraries, $I_s^{'} \in I_s$
Parameter	**Definition**
$C_{k,f}$	operating cost for flight f when operated by fleet type k
R_k	available number of planes for type k
Q_k	the capacity of fleet type k in number of seats
$min\mathrm{E}_a^-$	the time just before the first event at airport a
$max\mathrm{E}_a^+$	the time just after the last event at airport a
$\delta_{i,f}$	1 if itinerary i uses flight leg f, 0 otherwise
UB$_i$	the upper bound on the price of the itinerary i
V_i	the utility of itinerary i
z_i	the vector of explanatory variables for itinerary i
β	the vector of parameters of the logit model

Objective function(3) maximizes the profit calculated by revenue minus operating costs. Firstly, we have the fleet assignment constraints. Constraints (4) ensure the coverage of mandatory flights which must be served according to the schedule development. Constraints (5) are for the optional flights that have the possibility to be canceled. Constraints (6) maintain the flow conservation of fleet. Constraints (7) ensure that the usage of each plane type is consistent with the number of available planes. It is assumed that the network configuration at the beginning

Table 6: Decision variables of the integrated model

Variable	Definition
	Schedule planning
$x_{k,f}$	1 if fleet type k is assigned to flight f, 0 otherwise
y_{k,a,t^-}	the number of type k planes at airport a just before time t
y_{k,a,t^+}	the number of type k planes at airport a just after time t
	Revenue management
$\tilde{d}_i$	demand of itinerary i based on the logit model
d_i	realized demand of itinerary i
p_i	price of itineary i
$t_{i,j}$	redirected passengers from itinerary i to itinerary j
$b_{i,j}$	recapture ratio for the passengers spilled from itinerary i and redirected to itinerary j
$\pi^h_{k,f}$	assigned seats for flight f in a type k plane for cabin class h

and at the end of the period, which is one day, is the same in terms of the number of planes at each airport (8).

The relation between the supply capacity and the actual demand should be maintained. Therefore we have the constraints (9) which maintain that the assigned capacity for a flight should satisfy the demand for the corresponding itineraries. The actual demand is composed of the original demand of the itinerary minus the spilled passengers plus the recaptured passengers from other itineraries. The same constraints ensure that the itineraries do not realize any demand if any of the corresponding flight legs is canceled. We let the allocation of business and economy seats to be decided by the model as a revenue management decision. Therefore, we need to make sure that the total allocated seats does not exceed the capacity (10).

Demand related constraints include the constraints (11) which maintain that the total redirected passengers from itinerary i to all other itineraries including the no-revenue options do not exceed its realized demand. Finally, we have the nonnegativity constraints and upper bounds (14)-(20) for the decision variables.

5 Results on the added value of the integrated model

The mixed integer nonlinear problem is formulated in AMPL and BONMIN[2] is used to obtain feasible solutions. Since we cannot guarantee the convexity of the problem, BONMIN serves an an approximation method. In order to see the added value of the integration of the demand model we need to support our observations with a set of experiments. For that purpose we identified 18 data instances with different characteristics that are listed in Table 7. For the

[2] `https://projects.coin-or.org/Bonmin`

$$
\begin{aligned}
\max \sum_{h\in H}\sum_{s\in S^h}\sum_{i\in(I_s\setminus I'_s)} & (d_i - \sum_{j\in I_s} t_{i,j} + \sum_{j\in(I_s\setminus I'_s)} t_{j,i}b_{j,i})p_i \\
& - \sum_{\substack{k\in K\\ f\in F}} C_{k,f}x_{k,f} && && (3)\\
\text{s.t.} \sum_{k\in K} x_{k,f} &= 1 && \forall f\in F^M && (4)\\
\sum_{k\in K} x_{k,f} &\leq 1 && \forall f\in F^O && (5)\\
y_{k,a,t^-} + \sum_{f\in \text{In}(k,a,t)} x_{k,f} &= y_{k,a,t^+} + \sum_{f\in \text{Out}(k,a,t)} x_{k,f} && \forall [k,a,t]\in N && (6)\\
\sum_{a\in A} y_{k,a,\text{minE}_a^-} + \sum_{f\in CT} x_{k,f} &\leq R_k && \forall k\in K && (7)\\
y_{k,a,\text{minE}_a^-} &= y_{k,a,\text{maxE}_a^+} && \forall k\in K, a\in A && (8)\\
\sum_{s\in S^h}\sum_{i\in(I_s\setminus I'_s)} \delta_{i,f}d_i - \sum_{j\in I_s}\delta_{i,f}t_{i,j} &+ \sum_{j\in(I_s\setminus I'_s)} \delta_{i,f}t_{j,i}b_{j,i} \\
&\leq \sum_{k\in K}\pi^h_{k,f} && \forall h\in H, f\in F && (9)\\
\sum_{h\in H}\pi^h_{k,f} &\leq Q_k x_{k,f} && \forall f\in F, k\in K && (10)\\
\sum_{j\in I_s} t_{i,j} &\leq d_i && \forall h\in H, s\in S^h, i\in I_s && (11)\\
\tilde{d}_i &= D_s \frac{\exp\left(V_i(p_i,z_i;\beta)\right)}{\sum_{j\in I_s}\exp\left(V_j(p_j,z_j;\beta)\right)} && \forall h\in H, s\in S^h, i\in I_s && (12)\\
b_{i,j} &= \frac{\exp\left(V_j(p_j,z_j;\beta)\right)}{\sum_{k\in I_s\setminus\{i\}}\exp\left(V_k(p_k,z_k;\beta)\right)} && \forall h\in H, s\in S^h, i\in(I_s\setminus I'_s), j\in I_s && (13)\\
x_{k,f} &\in \{0,1\} && \forall k\in K, f\in F && (14)\\
y_{k,a,t} &\geq 0 && \forall [k,a,t]\in N && (15)\\
\pi^h_{k,f} &\geq 0 && \forall h\in H, k\in K, f\in F && (16)\\
0 \leq d_i &\leq \tilde{d}_i && \forall h\in H, s\in S^h, i\in I_s && (17)\\
0 \leq p_i &\leq UB_i && \forall h\in H, s\in S^h, i\in I_s && (18)\\
t_{i,j} &\geq 0 && \forall h\in H, s\in S^h, i\in(I_s\setminus I'_s), j\in I_s && (19)\\
b_{i,j} &\geq 0 && \forall h\in H, s\in S^h, i\in(I_s\setminus I'_s), j\in I_s && (20)
\end{aligned}
$$

Figure 1: Integrated schedule planning and revenue management model

experiments, we present the number of airports and the number of flights in the network. Moreover, the flight density stands for the average number of flights per route. The average demand gives the average number of passengers per flight according to demand forecast. The fleet composition provides information on the number of different plane types in the fleet together with the seat capacity for each type.

Table 7: The experiments

No	Airports	Flights	Flight density	Average demand		Fleet composition
1	3	10	1.67	51.9	2	50-37 seats
2	3	11	2.75	83.1	2	117-50 seats
3	3	12	2	113.8	2	164-100 seats
4	3	26	4.33	56.1	3	100-50-37 seats
5	3	19	3.17	96.7	3	164-117-72 seats
6	3	12	3	193.4	3	293-195-164 seats
7	3	33	8.25	71.9	3	117-70-37 seats
8	3	32	5.33	100.5	3	164-117-85 seats
9	2	11	5.5	173.7	3	293-164-127 seats
10	4	39	4.88	64.5	4	117-85-50-37 seats
11	4	23	3.83	86.1	4	117-85-70-50 seats
12	4	19	3.17	101.4	4	134-117-100-85 seats
13	4	15	1.88	58.1	5	117-85-70-50-37 seats
14	4	14	2.33	87.6	5	134-117-85-70-50 seats
15	4	13	2.6	100.1	5	164-134-117-100-85 seats
16	8	77	2.08	67.84	4	117-85-50-37 seats
17	7	56	2.33	87.84	4	164-117-85-50 seats
18	8	97	3.46	90.84	5	164-117-100-85-50 seats

For the considered data instances, we compared our integrated model with the price-inelastic schedule planning model and the sequential approach. The comparative results are presented in Table 8. In the table, price-inelastic schedule planning model is represented by *PISP*; sequential approach is represented by *SA* and the integrated model is represented by *IM*. Let us note that for the first 15 experiments BONMIN reports 0% duality gap for the integrated model although we cannot guarantee optimality. For the last three experiments the solution has a duality gap which results with lower profit compared to the sequential approach.

It is observed that the price-inelastic schedule planning model is outperformed by the two other models for all the experiments. The flexibility obtained by the control on the demand and price results with superior decisions. The analysis of the comparison between the sequential approach and out integrated model is more interesting because they both have the flexibility on the demand side however our integrated model decides on the schedule planning simultaneously with the revenue management. This simultaneous optimization provides superior decisions on the schedule planning. In Table 9, we report the improvement of the integrated model over the sequential approach, for the experiments with an improvement. It is observed that for 7 of these 15 instances there is an improvement with the integrated model in terms of the profit and transported number of passengers. These are the cases where the simultaneous optimization of the schedule planning and revenue management lead to different scheduling decisions such as the operated number of flights or the number of allocated capacity.

Table 9: The advantage of the integrated model over the sequential approach

Experiments	Profit	Transported pax.
2	5.55%	33.50%
4	1.43%	14.18%
6	0.30%	-
9	0.43%	5.83%
10	0.83%	4.94%
11	3.36%	1.40%
14	1.45%	16.69%

When we analyze the instances where there is an improvement, we observe that the improvement is higher when the demand levels for the flights has high variation but there is a few number of plane types. In those cases, the integrated model is able to adjust the capacity according to the demand and has significant improvement over the sequential approach. Experiment 2 is a good example for this phenomenon. There are 2 different fleet types with 50 and 117 seats. The sequential approach does not use the larger aircraft which is costlier to fly. On the other hand the integrated model uses this large aircraft thanks to its flexibility in controlling the demand by pricing decisions. As a result, there is a 5.55% increase in profit and 33.5% more passengers are transported. Similarly, for the experiments 4, 6, 9, 10, and 11 the integrated model decides to use more capacity with the knowledge on the demand behavior. In addition to the decision on the allocated capacity, the integrated model may decide to operate more flights by changing the attractiveness of the corresponding itineraries. For example, for experiment 14, the integrated model operates 2 more flights with the same overall capacity compared to the sequential approach. We observe a similar increase in the number of flights in experiment 4.

6 Heuristic approach

We are limited by the complexity of the mixed integer nonlinear problem. When we go beyond the presented instances in section 5 we are not able to obtain feasible solutions in reasonable computational time with BONMIN. Therefore we propose a heuristic in order to be able to test the integrated model for larger instances which represent the reality better.

The heuristic method is based on two simplified versions of the model that is presented in Figure 1. The first model, which is referred as FAM^{LS} enables us to explore new fleet assignment solutions based on a *local search* mechanism. The local search is developed by combining a *price sampling* and a *variable neighborhood* procedure. The price sampling is done such that a random price is drawn for each itinerary and according to this price the demand values and recapture ratios are fixed based on the equations 12 and 13. Variable neighborhood procedure is designed by fixing a subset of fleet assignments (Hansen and Mladenović, 2001). The number of fixed assignment is represented by n_{fixed} and varied according to the quality of the solution. When the solution is improved an *intensification* is applied by increasing n_{fixed}. On the other hand when there is not an improvement for a number iterations a *diversification* is utilized by fixing less assignments. The local search mechanism therefore enables us to visit better fleet assignment solutions. The set of fixed assignments is represented by L. Each fixed assignment l indicates a fleet type k_l^{fixed} and a flight f_l^{fixed}. We add this constraint to the model as given by equation 21. Therefore the FAM$^{\text{LS}}$ has the objective function (3) subject to the constraints (4)-(11), (14)-(17), (19), and the new defined constraint (21). Let us note that the variables $\tilde{d}$, p, b are parameters for the model due to the price sampling.

$$x_{k_l^{fixed}, f_l^{fixed}} = 1 \qquad \forall l \in L \tag{21}$$

The second model is referred as REV$^{\text{LS}}$ which optimizes the revenue for the fleet assignment solutions explored by solving the FAM$^{\text{LS}}$ model at each iteration of the local search. Therefore this model has the fleet assignment model variables of x and y as parameters. The objective can be reformulated as in equation 22 and maximized subject to the constraints (9)-(13) and (16)-(20).

$$\max \sum_{h \in H} \sum_{s \in S^h} \sum_{i \in (I_s \setminus I_s')} (d_i - \sum_{j \in I_s} t_{i,j} + \sum_{j \in (I_s \setminus I_s')} t_{j,i} b_{j,i}) p_i \tag{22}$$

The heuristic procedure consists of iteration each of which solves FAM$^{\text{LS}}$ and REV$^{\text{LS}}$ models subsequently until the maximum number of iterations, k_{max}, is reached. When the solution of BONMIN is available we terminate the iterations if the deviation from this solution, referred as z_{opt} is smaller than ϵ. This procedure is presented by Algorithm 1 where n_{min} and n_{max} are defined as the minimum and maximum number of fixed assignments according to the data instance.

6.1 Performance of the heuristic

For testing the performance of the heuristic we use the same set of instances provided in Table 7. The results of the heuristic compared to BONMIN is presented in Table 10. The time limit set for BONMIN is 12 hours, on the other hand maximum computational time allowed for the

Algorithm 1 Heuristic procedure

Require: $\bar{x}_0$, $\bar{y}_0$, $\bar{d}_0$, $\bar{p}_0$, $\bar{t}_0$, $\bar{b}_0$, $\bar{\pi}_0$, z^*, z_{opt}, k_{max}, ϵ, n_{min}, n_{max}
  $k := 0$, $n_{fixed} := n_{min}$
  repeat
    $\bar{p}_k$:= Price sampling
    $\{\bar{d}_k, \bar{b}_k\}$:= Demand model($\bar{p}_k$)
    $\{\bar{x}_k, \bar{y}_k, \bar{\pi}_k, \bar{t}_k\}$:= solve $z_{\text{FAM}^{\text{LS}}}(\bar{d}_k, \bar{b}_k, n_{fixed})$
    $\{\bar{p}_k, \bar{d}_k, \bar{b}_k, \bar{\pi}_k, \bar{t}_k\}$:= solve $z_{\text{REV}^{\text{LS}}}(\bar{x}_k, \bar{y}_k)$
    if improvement($z_{\text{REV}^{\text{LS}}}$) **then**
      Update z^*
      Intensification: $n_{fixed} := n_{fixed} + 1$ when $n_{fixed} < n_{max}$
    else
      Diversification: $n_{fixed} := n_{fixed} - 1$ when $n_{fixed} > n_{min}$
    end if
    $k := k + 1$
  until $||z_{opt} - z^*||^2 \leq \epsilon$ **or** $k \geq k_{max}$

heuristic is 1 hour. For both of them we report the time when the best solution is found. For the experiments 1-3 and 12-15 the heuristic is able to find the best solution of BONMIN in a few seconds. For other experiments we have 10 replications and we report the minimum, average and the maximum deviation from the best solution. Similarly, we report the minimum, average and maximum computational time needed.

In the majority of the instances the heuristic has a considerable reduction in computational time. When we analyze the quality of the solutions, the deviation from the best solution is on the average 2.3 % for the first 15 experiments. Let us note that the last three experiments were the ones where BONMIN reported a duality gap. These are instances with higher complexity due to increased number of flights. It is seen that the heuristic is outperforming BONMIN for experiment 16 with a higher profit and using significantly less computational time. For this particular experiment the maximum profit attained is 204,906 which is still inferior to the result of the sequential approach (see Table 8).

Table 10: Performance of the heuristic versus BONMIN

	Best solution reported by BONMIN		**Heuristic**					
			% deviation			Time(sec)		
Experiments	Profit	Time (sec)	min	avg.	max	min	avg.	max
1	15,091	11	-	0.00%	-	-	1	-
2	37,335	27	-	0.00%	-	-	2	-
3	50,149	56	-	0.00%	-	-	33	-
4	70,904	2,479	1.32%	1.77%	2.06%	288	1,510	3,129
5	82,311	1,493	0.00%	0.13%	0.22%	18	900	3,092
6	906,791	12,964	7.37%	7.37%	7.37%	25	279	1,434
7	135,656	23,662	13.88%	16.36%	18.84%	74	1,714	3,534
8	115,983	209	0.00%	0.01%	0.12%	643	1,955	3,432
9	858,544	7,343	3.42%	4.79%	6.92%	1	762	3,322
10	138,575	37,177	2.76%	3.94%	4.98%	929	1,775	2,891
11	96,486	17,142	0.00%	0.16%	0.90%	236	1,625	3,574
12	49,448	32	-	0.00%	-	-	1	-
13	27,076	36	-	0.00%	-	-	5	-
14	53,128	141	-	0.00%	-	-	2	-
15	26,486	14	-	0.00%	-	-	4	-
16	194,598	42,360	**-5.89%**	**-4.04%**	**-2.41%**	293	1,652	2,990
17	191,091	39,447	0.48%	2.13%	4.46%	32	1,646	3,305
18	351,655	17,424	4.91%	7.94%	11.22%	840	2099	3331

6.2 Future work on the heuristic method

We believe that the performance of the heuristic can be improved when considered in a Lagrangian relaxation framework. If we relax the constraint (10) of the integrated model presented in Figure 1 and introduce Lagrangian multipliers, $\lambda_{k,f}$, for each flight f and fleet type k we can decompose the problem into two subproblems. The first problem is a revenue maximization model which optimizes the pricing and seat allocation decisions. The objective function of this subproblem can then be formulated as in equation 23 where we have the Lagrangian multipliers. The related constraints for the revenue subproblem are (9)-(13) and (16)-(20).

$$\max \sum_{h\in H}\sum_{s\in S^h}\sum_{i\in(I_s\setminus I'_s)}(d_i - \sum_{j\in I_s} t_{i,j} + \sum_{j\in(I_s\setminus I'_s)} t_{j,i}b_{j,i})p_i + \sum_{k\in K f\in F}\lambda_{k,f}\sum_{h\in H}\pi^h_{k,f} \tag{23}$$

The second subproblem on the other hand is a fleet assignment problem where Lagrangian multipliers serves as a penalty on the allocation of the capacity. The objective function can be formulated as in equation 24. The related constraints are (4)-(8) and (14)-(15).

$$\min \sum_{k \in K f \in F} (C_{k,f} x_{k,f} - \lambda_{k,f} Q_k x_{k,f}) \tag{24}$$

These two subproblems can be integrated in a subgradient optimization framework which will provide an upperbound to the problem. This is important for the large instances where we do not have solutions from the BONMIN solver. This work on the Lagrangian relaxation is a work in progress.

7 Conclusions and Future Research

In this paper an integrated schedule planning and revenue management model is presented. The added value of the integration is evaluated in comparison to the models which mimic the state-of-the-art models. It is observed that the explicit representation of supply-demand interactions lead to superior schedule planning decisions.

As a solution method for the MINLP a simple heuristic method is proposed based on a local search procedure. The results on the heuristic are promising in terms of the reduction in the computational time and the quality of the solutions. The future work regarding the heuristic is the utilization of a Lagrangian relaxation based methodology. The heuristic then needs to be tested for larger instances to see the limit of our methodology. For the simplification of the model a piecewise linear approximation of the logit model can be considered.

The demand model included in the integrated model has only the price variable as a policy variable. The other attributes of the itineraries cannot be controlled by the integrated model. Therefore a future direction is the extension of the model where the flights can be rescheduled based on the demand model.

Acknowledgments

The authors would like to thank EPFL Middle East for funding this research. Special thanks go to Claudio Leonardi for the information he provided regarding the design of Clip-Air.

References

Atasoy, B. and M. Bierlaire (2012) An air itinerary choice model based on a mixed RP/SP dataset, *Technical Report*, **TRANSP-OR 120426**, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

Atasoy, B., M. Salani, C. Leonardi and M. Bierlaire (2011) Clip-air, a flexible air transportation system, *Technical Report*, **TRANSP-OR 110929**, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

Belobaba, P., A. Odoni and C. Barnhart (2009) *The Global Airline Industry*, first edn., John Wiley & Sons, Ltd, Chichester, UK.

Bierlaire, M. and M. Fetiarison (2009) Estimation of discrete choice models: extending BIOGEME, paper presented at *Swiss Transport Research Conference (STRC)*, September 2009.

D'Ambrosio, C. and A. Lodi (2011) Mixed integer nonlinear programming tools: a practical overview, *4OR: A Quarterly Journal of Operations Research*, **9**, 329–349.

Garrow, L. A. (2010) *Discrete Choice Modelling and Air Travel Demand: Theory and Applications*, Ashgate Publishing: Aldershot, United Kingdom.

Hansen, P. and N. Mladenović (2001) variable neighborhood search: Principles and applications, *European Journal of Operational Research*, **130**, 449–467.

Lohatepanont, M. and C. Barnhart (2004) Airline schedule planning: Integrated models and algorithms for the schedule design and fleet assignment, *Transportation Science*, **38**, 19–32.

Schön, C. (2008) Integrated airline schedule design, fleet assignment and strategic pricing, paper presented at *Multikonferenz Wirtscaftsinformatik (MKWI)*, February 2008.

Sherali, H. D., K.-H. Bae and M. Haouari (2010) Integrated airline schedule design and fleet assignment: Polyhedral analysis and benders' decomposition approach, *INFORMS Journal on Computing*, **22** (4) 500–513.

Talluri, K. T. and G. J. van Ryzin (2004a) Revenue management under a general discrete choice model of customer behavior, *Management Science*, **50** (1) 15–33.

Talluri, K. T. and G. J. van Ryzin (2004b) *The Theory and Practice of Revenue Management*, first edn., Kluwer Academic Publishers, Boston.

Table 8: The comparative results of the experiments

Experiments	Models	Profit	Transported pax.	Flights	Allocated seats
1	PISP	11,559	281	8	124
	SA	15,091	284	8	124
	IM	15,091	284	8	124
2	PISP	27,872	400	8	150
	SA	**35,372**	**400**	8	**150**
	IM	**37,335**	**534**	8	**217**
3	PISP	41,997	884	10	300
	SA	50,149	859	10	300
	IM	50,149	859	10	300
4	PISP	53,604	943	22	274
	SA	**69,901**	**931**	**22**	**274**
	IM	**70,904**	**1,063**	**24**	**324**
5	PISP	66,129	1,186	16	333
	SA	82,311	1,145	16	333
	IM	82,311	1,145	16	333
6	PISP	763,321	1,466	10	1,148
	SA	**904,054**	1,448	10	**1,148**
	IM	**906,791**	1,448	10	**1,312**
7	PISP	102,756	1,800	32	498
	SA	135,656	1,814	32	498
	IM	135,656	1,814	32	498
8	PISP	82,253	2,207	26	691
	SA	115,983	2,236	26	691
	IM	115,983	2,236	26	691
9	PISP	687,314	1,270	10	1,016
	SA	**854,902**	**1,270**	10	**1,016**
	IM	**858,544**	**1,344**	10	**1,090**
10	PISP	110,055	1,474	34	391
	SA	**137,428**	**1,517**	34	**391**
	IM	**138,575**	**1,592**	34	**476**
11	PISP	78,527	1,143	20	387
	SA	**93,347**	**1,144**	20	**387**
	IM	**96,486**	**1,160**	20	**457**
12	PISP	38,104	982	12	370
	SA	49,448	1,050	12	370
	IM	49,448	1,050	12	370
13	PISP	22,356	446	10	207
	SA	27,076	448	10	207
	IM	27,076	448	10	207
14	PISP	44,499	605	10	267
	SA	**52,369**	**599**	**10**	267
	IM	**53,128**	**699**	**12**	267
15	PISP	19,625	479	6	185
	SA	26,486	504	6	185
	IM	26,486	504	6	185
16	*PISP*	*173,513*	*2,676*	*62*	*958*
	SA	*208,561*	*2,678*	*62*	*958*
	IM	*194,598*	*2,664*	*59*	*873*
17	*PISP*	*162,601*	*2,717*	*46*	*1,044*
	SA	*196,434*	*2,742*	*46*	*1,044*
	IM	*191,091*	*2,929*	*48*	*1,161*
18	*PISP*	*292,956*	*5,362*	*75*	*1,784*
	SA	*365,753*	*5,388*	*75*	*1,784*
	IM	*351,655*	*5,295*	*73*	*1,667*