10th International Conference of the Practice and Theory of Automated Timetabling
PATAT 2014, 26-29 August 2014, York, United Kingdom

A Simulation Scenario Based Mixed Integer Programming Approach to Airline Reserve Crew Scheduling Under Uncertainty

Christopher Bayliss · Geert De Maere ·
Jason Atkin · Marc Paelinck

Abstract Airlines operate in an uncertain environment for many reasons, for example due to the effects of weather, traffic or crew unavailability (due to delay or sickness). This work focuses on airline reserve crew scheduling under crew absence and journey time uncertainty for an airline operating a single hub and spoke network. Reserve crew can be used to cover absent crew or delayed connecting crew. A fixed number of reserve crew are available for scheduling and each requires a daily standby duty start time. Given an airline's crew schedule and aircraft routings we propose a Mixed Integer Programming approach to scheduling the airline's reserve crew. A simulation of the airline's operations with stochastic journey time and crew absence inputs and without reserve crew is used to generate disruption scenarios for the $MIPSSM$ formulation (Mixed Integer Programming Simulation Scenario Model). Each disruption scenario corresponds to a record of all of the disruptions in a simulation for which reserve crew use would have been beneficial. For each disruption in a disruption scenario there is a record of all reserve crew that could have been used to solve or reduce the disruption. This information forms the input to the $MIPSSM$ formulation, which has the objective of finding the reserve schedule that minimises the overall level of disruption over a set of scenarios. Additionally, modifications of the $MIPSSM$ are explored, and a heuristic solution approach and a reserve use policy derived from the $MIPSSM$ are introduced. A heuristic based on the proposed Mixed Integer Programming Simulation Scenario Model or $MIPSSM$ outperforms a range of alternative

---

Christopher Bayliss · Geert De Maere · Jason Atkin
ASAP, University of Nottingham, UK
E-mail: cwb,gdm,jaa@cs.nott.ac.uk

Marc Paelinck
KLM Decision Support, Information services department
KLM Royal Dutch Airlines
KLM Headquarters, The Netherlands
E-mail: Marc.paelinck@klm.com

approaches. The heuristic solution approach suggests that including the right disruption scenarios is as important as ensuring that enough disruption scenarios are added to the $MIPSSM$.

Keywords Airline Reserve Crew Scheduling · Simulation · Mixed Integer Programming

1 Introduction

An airlines primary product is its schedule, due to operating costs airlines maximise the utilisation of resources (crew and aircraft) resulting in schedules with little slack. This makes each resource a critical component of an airlines network and if a component is missing all flights related to that component may be disrupted. Crew can be absent or delayed on connecting flights, in such circumstances airlines may call on reserve crew. This work focusses on reserve crew scheduling, using simulation generated disruption scenarios added to a Mixed Integer Programming model to schedule reserve crew.

A Mixed Integer Programming Simulation Scenario Model ($MIPSSM$) has been developed which will use information from repeat simulations of an airline network where reserve crew are not available. Then reserve crew are to be scheduled in such a way that the level of delay and cancellation that would have occurred in the original simulations (disruption scenarios) is minimised.

Simulation (Section 3.3) is used to generate the set of input disruption scenarios for the $MIPSSM$. A disruption scenario corresponds to the set of disrupted flights in a single run of the airline simulation, where a single run corresponds to executing the airlines schedule in the considered time horizon from start to finish once. For each disruption in a disruption scenario there is a record of all of the reserve crew start times (discretised according to scheduled departure times) which, if scheduled, would allow the corresponding reserve crew to be used to solve completely, or reduce, the given disruption. In the $MIPSSM$ there are 2 types of variables, X the reserve crew schedule and y the reserve use decisions within each disruption scenario that are feasible with respect to X. Reserves can only be used if they are scheduled. Solving the $MIPSSM$ in an appropriate solver finds the reserve crew schedule X and reserve use decisions y that minimises delay and cancellations in the set of disruption scenarios used to form the constraints and objective of the $MIPSSM$.

The remainder of the paper is structured as follows. Section 2 outlines closely related work. Section 3 introduces the simulation used to generate disruption scenarios, how disruption scenarios are derived from the simulation and presents the formulation of the $MIPSSM$. Section 4 covers modifications and variants of the basic $MIPSSM$ formulation. Section 5 gives experimental results. Section 6 concludes the paper with a summary of the main findings. Section 7 discusses future work.

2 Related work

The $MIPSSM$ has similarities to Recoverable Robustness [4]. In [4] Liebchen provides a framework for timetabling problems with the objective that the schedule must be feasible in each of a limited set of disruption scenarios given limited availability of recovery from disruptions. The approach reduces to strict robustness (feasible in all outcomes without recovery actions) if the feature of limited available recovery is removed. The similarity between Recoverable robustness and the $MIPSSM$ lies in the idea of solving a scheduling problem over a limited number of realistic disruption scenarios. The $MIPSSM$ is influenced by stochastic programming, which optimises over a set of explicit independent possible outcomes as opposed to optimising over the expected outcome, which may not even correspond to a possible outcome.

In [7], Bailey et al. present an airline reserve crew scheduling model that takes training days and bidline conflicts into account. Such conflicts arise when crew bid for rosters which overlap with recurring training and this leads to open time (flights without scheduled crew) which have to be covered with reserve crew. In [6], Shebalov tackles the robust airline crew pairing problem using the concept of move-up crews. Move-up crews refers to crews who can swap pairings in the event of delay (the available crew can adopt the delayed crew's pairing). Their objective is to maximise move-up crews. Shebalov measures the robustness of schedules/quality of the scheduled move-up crews in computational experiments in terms of the number of deadheads (crew transported as passengers to the origin of their next flight leg), reserve crew used, number of uncovered flight legs and the cost of crew schedule. For the interested reader other work carried out previously on the problem of airline reserve crew scheduling includes [2,3,5].

3 Deriving and formulating the $MIPSSM$

This section starts by introducing the notation, it then introduces the delay cancellation measure function (Section 3.2), which converts delays into a quantity with units of cancellations. This approach means the $MIPSSM$ remains a single objective problem. The cancellation measure function is used in the disruption scenario generating simulation (Section 3.5) to find the cancellation measures associated with all possible reserve crew start times that if scheduled could be used to solve or reduce a given disruption in the given disruption scenario. Section 3.3 gives details of the single hub airline simulation used for disruption scenario generation and (in Section 5) experimental validation of reserve crew schedules derived from the $MIPSSM$ as well as other methods. Section 3.5 defines what is meant by a disruption scenario and how the information it stores is collected from simulation. Section 3.6 defines the notation used in the $MIPSSM$ formulation. Section 3.7 presents and explains the $MIPSSM$ in terms of its objective and constraints.

3.1 Schedule notation

D_h : Scheduled departure time of flight h
C_h : Crew team number scheduled to flight h
A_h : Aircraft number scheduled to flight h
cd_h : Crew related delay at departure h that occurs in disruption scenario generating simulation
$rd_{h,l}$: Delay when reserve crew with start time index l used to cover disrupted crew of flight h
td_h : Total delay at departure h
$crewSize_h$: number of crew in crew team scheduled to flight h
$ceta_h$: Estimated time of arrival of crew scheduled to flight h
$aeta_h$: Estimated time of arrival of aircraft scheduled to flight h
CT : Cancellation threshold over which delayed flights are cancelled
MS : Minimum sit or minimum rest time required by crew between consecutive flights within a duty shift
TT : Minimum turn/ground time required by aircraft between consecutive flights
$|P_n|$: Length of crew pairing n in terms of hub departures
$P_{n,m}$: Departure number of the m^{th} hub departure of crew pairing n

3.2 Cancellation measure of a delay

To retain the simplicity of a single objective problem Equation 1 converts delay into a measure of cancellation. The simulation cancels flights with a delay over the cancellation threshold so the maximum cancellation measure of a delay is 1. cm_h is the cancellation measure of flight h, td_h (Equation 2) is the total delay of flight h, cd_h (Equation 3) is the delay of flight h due to crew over and above delay due to the aircraft, i.e. the delay which could be absorbed by using reserve crew. Equation 4 gives the delay due to waiting for reserve crew with start time index l (start time=D_l as reserve start times are discretised according to scheduled departure times) to begin their duty shift, counting only delay over and above delay due to the aircraft assigned to the same flight.

$$cm_h = \left(\frac{td_h - cd_h}{CT}\right)^n \tag{1}$$

$$td_h = \max\left(0,\ \ \max\left(aeta_h + TT,\ ceta_h + MS\right) - D_h\right) \tag{2}$$

$$cd_h = \max\left(0,\ ceta_h + MS - \max\left(D_h,\ aeta_h + TT\right)\right) \tag{3}$$

$$rd_{h,l} = \max\left(0,\ D_l - \max\left(D_h,\ aeta_h + TT\right)\right) \tag{4}$$

A decision maker choice is required for the delay exponent n of the cancellation measure function. Choosing higher values for $n > 1$ corresponds to giving lower weight to delays below the cancellation threshold. Using the delay cancellation measure function means that the objective measures of using reserve crew teams to cover delayed connecting crew and using reserve crew to cover absent crew are both in the same units, that of cancellations. In the following $n = 2$ is used.

3.3 Simulation

The simulation of a single hub airline is used without reserve crew to generate disruption scenarios which contain information on the possible benefit of using reserve crew scheduled at specified times in response to the given disruption. These disruption scenarios form the input for the $MIPSSM$ formulation (Section 3.7).

Simulation takes as input the airline's scheduled flights, the crew and aircraft scheduled to each of those flights. The simulation's stochastic inputs are journey times and crew absence, each of which have corresponding statistical distributions derived from real data. Crew and aircraft were scheduled using first in first out scheduling. In the crew schedule 30% of crew connections at the hub involve a change of aircraft. The scheduled journey times correspond to a 0.6 probability of early arrival.

A single run of the simulation proceeds by considering each scheduled departure in departure time order. If a departure corresponds to the start of a crew duty then the number of crew absent is instantiated from the cumulative statistical distribution of possible numbers of absent crew. If reserve crew are not available then the flight has to be cancelled. At this point in the simulation, information on the possible benefit of scheduling reserves at different start times is collected (Section 3.5). If reserve crew are available (as is the case in the validation simulation used in Section 5 to validate reserve crew schedules created using the methods proposed herein) they are considered for use in earliest start time order. If a departure is delayed by more than the delay threshold (15 minutes) all combinations of single crew and aircraft swaps are considered. Swaps are only considered feasible if the swap can take place without invoking additional delay on the flights affected by the swap, the crew must be able to complete each other's duties without violating maximum working hours and it must be possible to undo the swap in the overnight break (same overnight station).

In the disruption scenario generation simulation, after the consideration of swap recovery actions, if the delay is still above the delay threshold, information is collected for the given disruption on the possible benefit of scheduling reserve crew at different possible start times (Section 3.5). In the validation simulation, after the consideration of swap recovery actions, possible combinations of reserve crew are considered for replacing delayed connecting crew. If after delay recovery the delay is above the cancellation threshold (180 minutes) the flight is cancelled.

3.4 Disruption scenario notation

W : Number of disruption scenarios

W_i : Number of disruptions in scenario i

$N_{i,j}$: The number of reserve crew required to cover disruption j in scenario i

$CM_{i,j}$: Cancellation measure contribution when no reserves are used to cover disruption j in scenario i

$N_{i,j}$: Number of crew required to cover disruption j in scenario i

$F_{i,j}$: Set of feasible reserve instances for disruption j in scenario i

$F_{i,j,k}$: k^{th} instance of a reserve feasible to cover disruption j in scenario i

$F^{V}_{i,j,k}$: k^{th} reserve use variable index feasible for disruption j in scenario i

$F^{U}_{i,j,k}$: k^{th} index of reserve use variable first used at disruption j in scenario i which can subsequently be used to absorb crew related delay propagated to a following flight

$F^{CM}_{i,j,k}$: Cancellation measure that occurs as a result of using the k^{th} feasible reserve for disruption j in scenario i

$F^{RD}_{i,j,k}$: reserve delay corresponding to feasible reserve use instance k feasible for disruption j in scenario i

$G_{i,j}$: Set of feasible reserve instances corresponding to reserve crew first used to absorb delay on a preceding flight that also have the knock-on effect of preventing or reducing delay disruption j in scenario i

$G_{i,j,k}$: k^{th} instance of a reserve feasible corresponding to a reserve first used to absorb crew delay on a preceding flight that also has the knock-on effect of reducing delay disruption j in scenario i

$G^{V}_{i,j,k}$: k^{th} reserve use variable index corresponding to a reserve first used to absorb delay on a preceding flight that has the knock-on effect of reducing delay disruption j in scenario i

$G^{CM}_{i,j,k}$: Cancellation measure corresponding to the k^{th} feasible reserve use instance first used to absorb delay on a preceding flight that also has the knock-on effect of reducing delay disruption j in scenario i

$R_{i,k}$: Set of feasible reserve use variable instances corresponding to reserve k in scenario i

$R_{i,k,l}$: Reserve use variable index corresponding to the l^{th} reserve use variable corresponding to reserve k in scenario i

3.5 Simulation derived scenarios

Simulation is used to derive disruption scenarios that are used as input for the $MIPSSM$. This section explains how simulation is used to derive the information for disruption scenarios. A given disruption scenario i corresponds to a single run of the simulation.

In disruption scenario i, a disruption j is a flight which has a delay over the delay threshold after the consideration of swap recovery or has to be cancelled due to crew absence. Such disrupted flights have a positive cancellation measure, where $CM_{i,j}$ denotes the cancellation measure of disruption j in disruption scenario i.

In a given run of the simulation, when a disruption occurs with a positive cancellation measure, data is collected regarding all of the possible feasible reserve start times that, could be used to reduce the disruption. For each such beneficial reserve start time, feasible reserve use instances are generated. A feasible reserve use instance corresponds to a possible scheduled reserve crew duty start time and subsequent use to cover a given crew disrupted flight in a given scenario. The number generated is equal to the number of reserve crew required to cover the given disruption, which is the number of crew absent in the event of a crew absence disruption or the size of the crew team assigned to flight h ($crewSize_h$) in the event of a delay. For each feasible reserve use instance (b) there is a corresponding cancellation measure (b^{CM}) that replaces the cancellation measure ($CM_{i,j}$) of the disruption if the reserve is used, a unique reserve use variable index (b^V), a unique knock on effect reserve use variable index (b^U) (if applicable) and a reserve delay (b^{RD}). Let $F_{i,j}$ denote the set of feasible reserve use instances corresponding to possible reserve start times that could be used to solve or reduce disruption j of disruption scenario i.

For the specific case of delay disruptions it is also possible that if there was crew delay on the preceding flight then the delay on the current flight might possibly be prevented or reduced by reserve crew used to absorb the initial delay. For this purpose the set $G_{i,j}$ is introduced and denotes the set of feasible reserve use instances corresponding to reserves used to cover crew related delay propagated from a previous previous flight. These feasible reserve use instances only apply if the corresponding feasible reserve use instances are used to cover the root crew related delay. G accounts for reserve crew that can have the effect of absorbing knock on crew delays. Algorithms 1 and 2 outline the procedure of collecting information for the disruption scenarios from the single hub airline simulation.

Algorithm 1 Pseudocode for deriving disruption scenario information for a crew absence disruption occurring at simulation run i departure k

```
During simulation run $i$ the $k^{th}$ scheduled flight is disrupted resulting in the $j^{th}$ disruption for which reserve crew use could potentially be beneficial
if Crew absence disruption then
    Create new disruption ($j$) for scenario ($i$), store the size of the disruption if not absorbed by utilising reserve crew, i.e. The number of flights cancelled=size of pairing ($|P_{C_k}|$) and store the number of crew absent ($N_{i,j}$)
    for each hub departure ($m$) in the crew absence disrupted pairing do
        for each reserve duty start time ($l$) feasible to cover the absence disrupted pairing at the $m_{th}$ hub departure of the disrupted pairing do
            Add $N_{i,j}$ new feasible reserve use instances to $F_{i,j}$, each with a unique variable number index ($V$), compute the associated cancellation measure ($CM$), Add the generated feasible reserve use instances to $R_{i,l}$ for the constraints regarding reserves only being used once per scenario.
        end for
    end for
    $j = j + 1$
end if
```

Algorithm 1 is used in the simulation when a crew absence occurs. The number of reserves required to cover this disruption is the number of absent crew (line 3). The cancellation measure of the absence disruption ($CM_{i,j}$) is the number of hub departures in the disrupted crew pairing that would have to be cancelled if reserves are unavailable to cover the absent crew (line 3), with no delay cancellation measure contribution. The algorithm then considers each possible reserve start time (line 5) which can be used to cover absent crew at each hub departure in the disrupted crew pairing (line 4). If reserve start time l is feasible, $N_{i,j}$ new instances of feasible reserve use instances are created with unique reserve use variable indices and cancellation measures equal to the number of flights that have to be cancelled before crew absence is covered at the m^{th} hub departure in the disrupted crew pairing plus a delay cancellation measure contribution from any delay caused by the reserve start time (lines 6). $cm = m - 1 + \left(\frac{rd_{f,l}}{CT}\right)^n$ is the equation for the cancellation measure associated with reserve crew with start time index l being used to cover crew absence disrupted pairing at the m_{th} flight in the crew pairing, where $f = P_{C_k,m}$ is the departure number of the flight the reserve crew are used to cover the absence disrupted crew pairing. The newly generated instances of feasible reserve use are also stored from a reserve perspective (R) (line 6), which is useful later on when creating constraints for feasible reserve use in the $MIPSSM$ formulation.

Algorithm 2 Pseudocode for deriving disruption scenario information for a crew delay disruption occurring at simulation run i departure k

```
1:  During simulation run $i$ the $k^{th}$ scheduled flight is disrupted resulting in the $j^{th}$ disruption for which reserve crew use could potentially be beneficial
2:  if crew delay disruption then
3:      Store the number of delayed connecting crew that need to be replaced ($N_{i,j} = crewSize_k$) and the cancellation measure of the delay $CM_{i,j} = \left(\frac{td_k}{CT}\right)^n$
4:      for Each reserve start time index $l$ that if scheduled could feasibly reduce the crew related delay of departure $k$ do
5:          Generate $N_{i,j}$ reserve use instances with unique reserve use variable indices ($V$), store the corresponding cancellation measure that applies if the reserves with start time $D_l$ are used, add the reserve use instances to $R_{i,l}$
6:      end for
7:      if Current crew delay is crew delay propagated from the crew's previous flight $q$, disruption $o$ then
8:          for $l = 1$ to $|F_{i,o}|$ do
9:              Create new reserve use instance, with unique reserve use variable index ($V$) and store in $G_{i,j}$, store the corresponding cancellation measure that applies if the root crew delay is absorbed using the reserve associated with the reserve use instance $F_{i,o,l}$
10:             $F^U_{i,o,l} = G^V_{i,j,a}$, where $a = |G_{i,j}|$
11:         end for
12:     end if
13:     $j = j + 1$
14: end if
```

Algorithm 2 is used in the simulation when a crew related delay occurs, the number of reserves required to cover this disruption is the number of crew in the delayed crew team (line 3). The cancellation measure of the delay disruption if reserve crew are not available to cover the delayed crew is also computed (line 3). The algorithm then considers each feasible reserve start time (line 4) used to cover the delay, and for each generates $N_{i,j}$ new feasible reserve use instances with unique reserve use variable indices and cancellation measures calculated using Equation 1 with Equation 4 added to the numerator. Lines 7 to 12 of Algorithm 2 apply if the given delay originated from a crew delay in the scheduled crew's previous flight. In this case it's possible that reserve use instances generated for that previous flight may have the effect of preventing delay propagating to the given delayed flight. For such feasible reserve use instances (line 8) U denotes new unique reserve use variable index for the reserve first used earlier, used to reduce the knock-on delay. For the current disruption j the set G stores the same newly generated reserve use variable index and a cancellation measure (line 9) that depends on the amount of delay that would have propagated if the reserve use instance feasible to cover the root crew delay is utilised. The $MIPSSM$ has constraints ensuring that the beneficial knock on effects can only apply if the reserve is actually used to absorb the root crew delay that propagated in the simulation. After the disruption scenarios have been created they can be used to create the constraints and objective of the $MIPSSM$.

3.6 $MIPSSM$ notation

X : Reserve crew schedule
x_k : Number of reserves with start time index k
Y : Set of reserve use variables
y_m : Reserve use instance variable m
$\delta_{i,j}$: Binary variable describing whether or not disruption j in scenario i is left uncovered (1) or covered (0) by reserve crew
$\gamma_{i,j}$: Real valued variable which takes on the cancellation measure of disruption j in scenario i given the reserve recovery decision made by the model
Z : Variable that takes on a value equal to the cancellation measure total of the scenario with the maximum cancellation measure
TR : Total reserve crew available for scheduling
ND : Total flights in reserve crew scheduling time horizon
Rt_q : Reserve use policy, the minimum threshold number of reserve crew remaining for using a team of reserve crew to cover a delayed connecting crew to be considered acceptable
obs_q : Number of times reserve teams are used to cover delayed connecting crew at flight q in reserve use policy derivation
$simRpts$: Number of repeat simulations used to derive a reserve use policy for a given reserve crew schedule

The reserve schedule X specifies the number of reserves which are scheduled to begin duties at a given time index k. $\gamma_{i,j}$ is a real valued variable, which equals the cancellation measure of disruption j in scenario i given the reserve use decisions (y). Each y variable corresponds to an individual reserve with a given start time index being used to cover a given disruption.

3.7 Mixed Integer programming formulation

Minimise:

$$\sum_{i=1}^{W}\sum_{j=1}^{W_i}\gamma_{i,j} \tag{5}$$

s.t.

$$\sum_{k=1}^{|F_{i,j}|} y_{F^V_{i,j,k}} + \sum_{k=1}^{|G_{i,j}|} y_{G^V_{i,j,k}} + \delta_{i,j}N_{i,j} = N_{i,j},\, \forall i \in 1..W,\, \forall j \in 1..W_i \tag{6}$$

$$\sum_{i=1}^{ND} x_i = TR \tag{7}$$

$$\sum_{l=1}^{|R_{i,k}|} y_{R^V_{i,k,l}} \leq x_k,\, \forall k \in 1..ND,\, \forall i \in 1..W \tag{8}$$

$$y_{R^U_{i,k,l}} \leq y_{R^V_{i,k,l}},\, \forall l \in R_{i,k} | \exists y_{R^U_{i,k,l}},\, \forall i \in 1..W,\, \forall k \in 1..ND \tag{9}$$

$$\delta_{i,j}CM_{i,j} \leq \gamma_{i,j}, \forall i \in 1..W,\, \forall j \in 1..W_i \tag{10}$$

$$y_{F^V_{i,j,k}} F^{CM}_{i,j,k} \leq \gamma_{i,j}, \forall i \in 1..W,\, \forall j \in 1..W_i,\, \forall k \in F_{i,j} \tag{11}$$

$$y_{G^V_{i,j,k}} G^{CM}_{i,j,k} \leq \gamma_{i,j}, \forall i \in 1..W,\, \forall j \in 1..W_i,\, \forall k \in G_{i,j} \tag{12}$$

$$y_m \in \{0,1\},\, \forall m \in Y \tag{13}$$

$$\delta_{i,j} \in \{0,1\},\, \forall i \in 1..W,\, \forall j \in 1..W_i \tag{14}$$

$$X_k \in \{0, 1...maxCA_i - 1, maxCA_i\},\, \forall k \in 1..ND \tag{15}$$

Objective 5 minimises the sum of all cancellation measures over all disruptions in all the scenarios included in the model. Constraint 6 ensures that disruptions are only considered covered if the required number of reserves are used for the given disruption. Constraint 6 forces $\delta_{i,j}$ to 1 when no reserve recovery can be applied to disruption j in scenario i and to 0 otherwise. Constraint 6 means that it is acceptable to cover a crew delayed departure with a combination of reserves used now and reserves used to cover a preceding crew delay that propagated, which may be useful if some of the reserves used to cover the root delay are not feasible to cover the following flight. Constraint 7 ensures that no more than the total number of reserves available (TR) are scheduled. Constraint 8 ensures that in each disruption scenario the number of reserves used with the same start time index does not exceed the number of reserves which are scheduled to that start time index. Constraint 9 ensures that disruptions can only be absorbed by reserves which were first used to absorb delay on the preceding flight if the reserve is used to cover that preceding flight. Constraints (10 to 12) ensure that the cancellation measure associated with a given disruption is the maximum of that associated with the recovery actions used for the given disruption. If no reserves are used for a given disruption that disruption gets the cancellation measure CM_{ij}, the same as occurred in the simulation in which the disruption occurred. If reserves are used the cancellation measure corresponds to the reserve used for that disruption that invokes the largest cancellation measure (as the flight can't take off before all the crew are present). Constraints 13 to 15 are integrality constraints.

4 Variants and Modifications

This section firstly considers 2 alternative formulations of the basic $MIPSSM$ formulation given in equations 5 to 15. Then a scenario selection heuristic designed to address the question of whether the types of scenarios or the number of scenarios included in the formulation has the greatest effect on solution quality. The final part in this section introduces an approach for deriving an optimal reserve use policy for a given reserve schedule, by repeated solving of the $MIPSSM$ for a single disruption scenario and fixed reserve schedule and learning the circumstances in which reserve use is beneficial in the long run.

4.1 Alternative objectives for the $MIPSSM$

Several alternative objectives are suggested in this section.

$MiniMax1$

The objective of minimising the sum of cancellations measures over all disruption scenarios included in the model (Objective 5) could be replaced with the alternative objective $MiniMax1$ of minimising the sum of cancellation measures of the disruption scenario with the largest sum of cancellation measures. This is a minimax objective function and can be implemented by replacing

Objective 5 with Objective 16 and adding Constraint 17. Information on implementing minimax objectives in linear programs can be found in [8]. This approach will have the effect of finding a reserve crew schedule that minimises the extent of the worst case scenario as opposed to minimising the average cancellation measure. In Table 1 the *probability of delay over 30 minutes* performance measure is most relevant to the $MiniMax1$ formulation.

$$\text{min: } Z \tag{16}$$

$$\sum_{j=1}^{W_i} \gamma_{i,j} \leq Z,\, \forall i \in 1..W \tag{17}$$

$MiniMax2$

Instead of minimising the total cancellation measure of the disruption scenario with the largest cancellation measure, the same principle can be applied to individual scenarios with the alternative Objective $MiniMax2$. I.e. find the reserve crew schedule that minimises the single largest disruption. To implement this approach replace Constraint set 17 with Constraint set 18. In Table 1 there is no performance measure which is directly relevant to the $MiniMax2$ formulation because in the reserve crew schedule validation simulation the worst single disruption is a cancellation and these will inevitably occur in each method. However in the $MiniMax2$ formulation the worst single disruption is leaving an absence disruption uncovered which would result in all flights on the absent crew's line of flight being cancelled.

$$\gamma_{i,j} \leq Z,\, \forall i \in 1..W,\, \forall j \in 1..W_i \tag{18}$$

4.2 Alternative solution approach

Scenario Selection Heuristic

The basic $MIPSSM$ and the two alternative formulations $MiniMax1$ and $MiniMax2$ are solved over a given set of disruption scenarios in a linear programming solver (CPLEX in this case). Although CPLEX yields optimal solutions, the solutions are only optimal for the set of disruption scenarios considered in the model. This section introduces a scenario selection heuristic (SSH) to address the issue of the choice of scenarios included in the $MIPSSM$. The solution time increases sharply as the number of disruption scenarios increases, which provides another motivation for considering a scenario selection heuristic solution approach, which includes the right scenarios rather than ensuring that enough disruption scenarios are included in the model. The following heuristic for solving the model defined by Equations 5 to 15 is based on adding one disruption scenario to the model at a time and stopping only when a new disruption scenario cannot be found for which the incumbent solution (overall reserve schedule) performs worse than in the worst disruption scenario already added to the model. The heuristic is analogous to column generation in which the master problem and pricing problem are solved iteratively. Repetition ceases when a specified iteration limit ($itLim$) is reached (in which case the whole algorithm terminates and returns a final solution), or, alternatively, when no new scenario with a sub problem ($subObj$) objective value can be found that is larger than the scenario already in the master problem with the largest objective contribution ($\max_j(masterObj_j)$), after $rptLim$ attempts. In summary this scenario selection approach focusses on finding a reserve schedule that can cope with a wide variety of difficult scenarios as opposed to a random set of scenarios representing the average outcome. An outline of the scenario selection heuristic is given in Algorithm 3.

Algorithm 3 Psuedocode for the scenario selection heuristic

```
$newScenarioFound = true$
$its = 0$
while $newScenarioFound \wedge its \leq itLim$ do
    $newScenarioFound = false$
    $rpts = 0$
    while $\neg newScenarioFound \wedge rpts < rptLim$ do
        Run simulation to generate disruption scenario $newScenario$
        Solve new scenario subproblem
        if $subObj > \max_j(masterObj_j)$ then
            $newScenarioFound$ =true
            add new scenario to the master problem
        else
            $rpts = rpts + 1$
        end if
        if $newScenarioFound$ then
            resolve master problem
        end if
    end while
    $its = its + 1$
end while
return solution
```

4.3 Optimal reserve use policy derivation

The simulation (Section 3.3) when used to test reserve schedules has a default policy of using reserve crew whenever this is immediately beneficial. The default policy also uses reserve crew in earliest start time order, so as to leave the largest amount of unused reserve crew capacity available for subsequent disruptions. The $MIPSSM$ approach uses reserve crew in each disruption scenario in an optimal way based on full knowledge of future disruptions. Knowledge of future disruptions is not available in the simulation, if a scenario which was included in the $MIPSSM$ formulation is repeated in the validation simulation, reserves might not necessarily be used in the same optimal way.

In this section an algorithm for deriving an optimal reserve use policy is described. The policy is based on reserve use decisions in response to delayed crew, where a team of reserve crew could be constructed and used to absorb the delay. The policy consists of threshold minimum numbers of reserve crew remaining for each departure for which using reserve teams to absorb crew related delay is deemed globally beneficial. The threshold numbers of reserves remaining are derived from the simulation by solving for the reserve use variables of the $MIPSSM$ model for the given reserve schedule over the disruptions that occur in the given run of the simulation and then averaging the number of reserves remaining at times when the $MIPSSM$ model uses reserve crew to cover for delayed crew.

The default policy is used for reserve crew use in response to crew absence since the penalty for not replacing absent crew with reserves is far too high (cancellation) to consider a crew absence reserve holding policy, and the cost of using teams of reserves to cover delayed crew is too high if this leaves too few reserve crew to cover subsequent absences. In general using teams of reserve crew to cover delayed connecting crew is expensive as it solves a smaller disruption (a delay compared to a cancellation) using more reserves than are usually required to cover absent crew.

5 Experimental results

The reserve crew scheduling approaches $MIPSSM$ of Section 3.7, $MiniMax1$ and $MiniMax2$ of Section 4.1 and SSH of Section 4.2 are tested and compared to one another. IBM CPLEX Optimization Studio version 12.5 with Concert technology is used as the MIP solver, on a desktop computer with a 2.79GHz Core i7 processor. These methods are also compared to a range of alternative methods for reserve crew scheduling (described below).

5.1 methods

Probabilistic reserve crew scheduling under uncertainty

The probabilistic approach ($Prob$) to reserve crew scheduling is an application of the work by the same authors in [1]. Given knowledge of the probabilities of crew absence for each flight in an airline's schedule, the probabilistic model evaluates the effect a given reserve crew schedule has on the probabilities of cancellations due to crew absence. The solution space of reserve crew schedules is then searched to find the reserve crew schedule that minimises the probabilities of cancellations due to crew absence. It was found that constructive heuristics provide near optimal solutions when solving the model put forward in [1]. The work in [1] has been extended to account for different numbers of crew being absent from each crew pairing in an airline schedule. Moreover the constraint that reserve crew are only feasible for disruptions if their duty start time is no later than the scheduled departure time of the disrupted flight has been relaxed so that some reserve delay is permitted, just as in the $MIPSSM$. Reserve delays in the probabilistic approach are accounted for using the delay cancellation measure function (Equation 1).

Area Under the Graph

The Area Under the graph ($Area$) method is based on running a number of simulations and recording the cumulative demand for reserve crew with respect to time in the form of a bar chart (in terms of the cancellation measure that could be avoided if reserve crew were available). Reserve crew are then scheduled at equal area intervals under the reserve demand graph over the whole scheduling time horizon. The $Area$ approach is based on a simulation without reserve crew to find reserve demand independent of the effects of a reserve crew schedule.

Uniform start rate

The Uniform Start Rate method (USR) schedules reserve crew at equal time intervals.

Zeros

The Zeros method ($Zeros$) schedules all reserve crew to begin standby duties at the first departure of the first day.

5.2 Experiment design

The methods stated in Section 5.1 are each solved for a synthetic airline schedule, synthetic in that the schedule is designed to increase the chance of delays due to delayed connecting crew. Other than this the schedule is fully detailed in terms of crew pairings and aircraft routings. Journey time uncertainty is modelled by statistical distributions based on real data, crew absence uncertainty is modelled as each individual scheduled crew member having a 1% chance of being absent and missing their entire crew pairing. All teams of crew consist of 4 individuals with identical rank (primarily aimed at cabin crew, but extending also to technical crew). The schedule is based on a 3 day single hub airline schedule with 243 flight legs a day with half being from the hub station. The schedule uses 148 teams of crew scheduled and 37 aircraft (single fleet). The schedule was generated using a first in first out approach with stochastic parameters controlling the rate of crew aircraft changes (0.3) and the 60^{th} percentile journey time from each destination's cumulative journey time distribution. These parameters influence the likelihood of delay propagation and the occurrence of delayed connecting crew. The following experiments investigate the effect of the number of reserve crew available for scheduling for each solution approach, and for the $MIPSSM$ based approaches the effect of the number of input disruption scenarios on solution time and solution quality.

5.3 Investigating the effect of varying the number of reserve available for scheduling

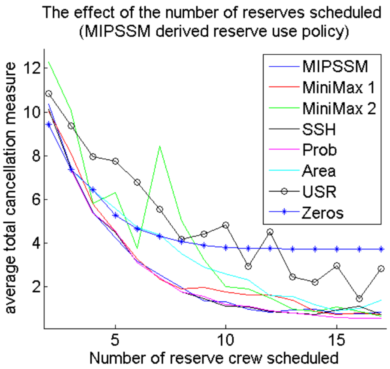

Fig. 1 The effect of the number of reserves which are scheduled on the solution quality of different solution approaches

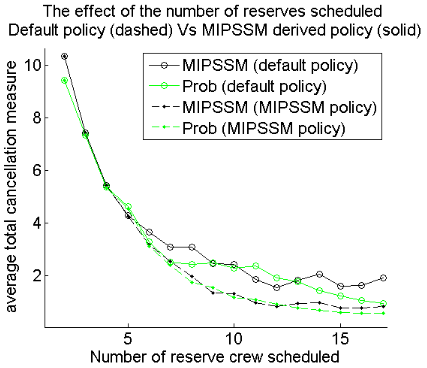

Fig. 2 The effect of the $MIPSSM$ derived reserve use policy

The results in Figure 1 show the effect on the average cancellation measure of varying the number of reserve crew available for scheduling, using 20000 repeat validation simulation tests for each reserve crew schedule for each solution approach. The $MIPSSM$ based approaches are restricted to 50 input disruption scenarios and a maximum of 1 hour to find a solution.
Figure 1 shows how the various reserve crew scheduling approaches compare for different numbers of reserve crew available for scheduling.
The $MIPSSM$, SSH and $Prob$ obtain the lowest average cancellation measures for all numbers of reserve crew available for scheduling of those tested. The $Prob$ model gives a smooth curve of average cancellation measures, whereas $MIPSSM$ and SSH have small fluctuations in average cancellation measure as the number of reserve crew available for scheduling changes. This fluctuation can in part be attributed to the limited number of disruption scenarios used as input for these methods. The $MiniMax1$ modification generally lead to higher average cancellation measures especially when between 9 and 12 reserve crew were available for scheduling. $MiniMax2$ gave the unexpected result that scheduling more reserve crew can lead to a higher average cancellation measure. This fluctuating behaviour of the $MiniMax2$ modification was also observed to a lesser extent in the other methods based on the $MIPSSM$ and can be explained by the fact that the objective of the $MiniMax2$ modification is to suppress the single largest delay or cancellation disruption that can

Method name	**Average cancellation measure**	**Average delay**	**Probability of delay** > 30**mins**	**Cancellation rate**	**Reserve Utilisation rate**	**solution time /mins**
NoRes	15.009	11.147	0.00682	0.03925	-	-
MIPSSM	1.159	12.180	0.00898	0.00140	0.674	28.688
*MiniMax*1	1.246	12.393	0.00938	0.00154	0.666	7.060
*MiniMax*2	1.724	13.874	0.01114	0.00171	0.656	2.259
SSH	1.066	11.870	0.00871	0.00141	0.667	2.871
Prob	1.077	11.518	0.00818	0.00166	0.690	0.443
Area	2.399	14.001	0.01130	0.00353	0.589	0.060
USR	2.925	14.970	0.01336	0.00438	0.555	<0.001
zeros	3.756	11.167	0.00725	0.00902	0.571	<0.001

Table 1 Performance measure averages from 20 repeats

occur and is not to minimise the average cancellation measure. This fluctuation is due to the resultant schedules being designed for worst case disruptions as opposed to the average outcomes. The *Area* under the graph approach lead to average cancellation measures similar to those from the *MiniMax*2 modification without the fluctuations. The *USR* approach lead to the highest average cancellation measures when 10 or fewer reserve crew are available for scheduling, for more than 10 reserve crew the *zeros* approach gave the highest cancellation measures. The *zeros* approach also gave the best results when fewer than 4 reserve crew were available for scheduling, this is because most crew absences are realised at the start of the first day, so scheduling reserve crew at that time prevents cancellations due to crew absence from the outset. The difference between the various solution approaches is clearest when there are around 10 to 12 reserve crew available for scheduling, which also appears to be the most sensible number of reserve crew to schedule (due to diminishing returns). Between 10 and 12 reserve crew, Figure 1 shows that the best performing solution approach was the *SSH*. 10 to 12 reserves for the given problem instance is approximately proportionate to the number of reserve crew scheduled in reality.

Figure 2 shows the effect of using the *MIPSSM* derived reserve use policy described in Section 4.3 compared to the default policy of using reserve crew as demand occurs. Using the *MIPSSM* derived policy had the effect of reducing the average cancellation measure.

5.4 Other performance measure and solution stability comparison of methods

Table 1 gives average performance measures when each method is applied to the same problem instance 20 times, for the *MIPSSM* approaches the simulation generated scenarios differ in each of the 20 repeats as they start with a different random seed. The results show that on average the *MIPSSM* performs best on cancellation rate, however the *MIPSSM* is also the slowest method with average solution times of an hour. The average cancellation measure can be interpreted as the number of cancellations expected in each three

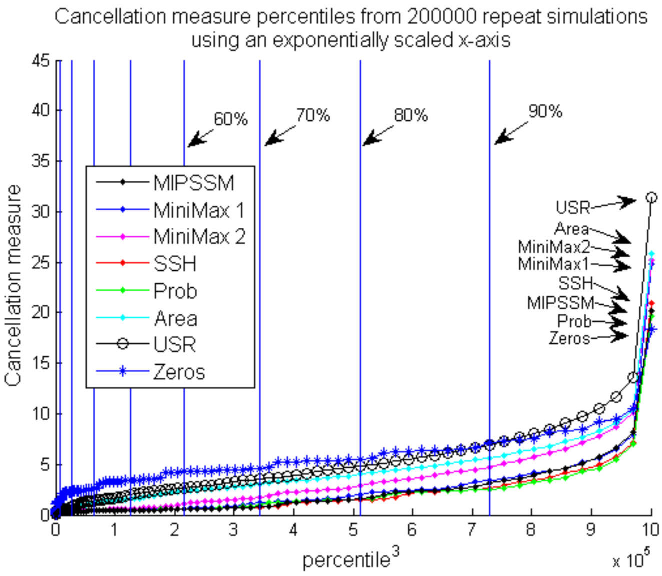

Fig. 3 Percentile cancellation measures

day simulation, but this also includes delays which have been converted to a cancellation measure using Equation 1 of Section 3.2. In terms of all round performance the SSH is a highly efficient approach with the lowest cancellation measure and also, a low average delay, the SSH is also much faster than the $MIPSSM$ with an average solution time of just under 3 minutes. The solution time of the SSH is a result of the termination criteria being satisfied before more than 10 disruption scenarios are added to the master problem. The $Prob$ approach has the second highest average cancellation measure, good average delay performance and a solution time much quicker the those of the $MIPSSM$ based approaches.

The results in Table 1 suggest there is merit in both the probabilistic and $MIPSSM$ approaches to scheduling airline reserve crew under uncertainty. Table 1 also includes performance measures when no reserve crew are scheduled at all as a point of reference. Counter to expectation the probability of delay over 30 minutes is lower without reserve crew, as is the average delay, however this can be attributed to the high cancellation rate, since cancelled flights do not count as delays. The Objectives $MiniMax1$ and $MiniMax2$ are aimed at minimising worst case scenarios, however if the probability of delay over 30 minutes is treated as a measure of worst case scenarios, it does not support this. Reserve utilisation rates are also given in Table 1 and are loosely correlated with the cancellation measures.

Figure 3 displays cancellation measure percentiles, the 100^{th} gives the worst cancellation measure from each approach, and this is the most appropriate validation criteria for the $MiniMax2$ modification. The $MiniMax2$ modification does not have the lowest cancellation measure for the 100^{th} percentile,

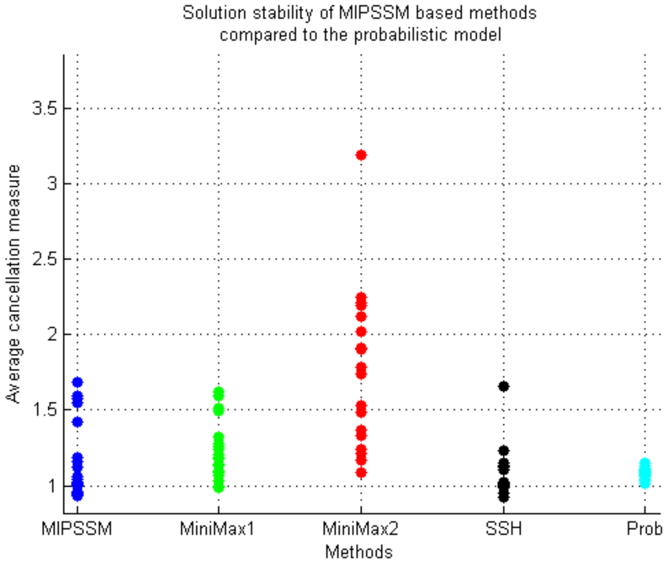

Fig. 4 Solution stability of $MIPSSM$ based methods compared to $Prob$

so it appears that this modification does not achieve its objective. Figure 3 shows the spread of cancellation measures corresponding to each method over the 20 repeats of each method, with each being tested in 20000 repeat validation simulations. Figure 3 demonstrates that for each given percentile the ordering of the methods supports the results given in Table 1 except for the $zeros$ approach which has the lowest worst case cancellation measure. This result suggests that the worst scenario is, for a very large number of crew to be absent at the start of each day, which is precisely the situation the $zeros$ approach can cope with. Figure 4 shows that the $MIPSSM$ based methods have a solution stability issue. Each point on Figure 4 represents a solution to the given method starting from a different random seed in the simulation used to generate the set of disruption scenarios over which the method is solved. Figure 4 shows that the $MIPSSM$ based methods have the potential to give solutions of higher quality that the Probabilistic method, but this depends on the selection of disruption scenarios which used as input for the given $MIPSSM$ based method.

6 Conclusion

In conclusion, a simulation based mixed integer programming approach to airline reserve crew scheduling has been introduced. The main idea is to schedule reserve crew using information from repeat simulations of an airline network where reserve crew are not available, and then scheduling reserve crew in a hindsight fashion in such a way that had they been available, the level of delay and cancellation that was related to disrupted crew would have been minimised. The $MIPSSM$ formulation also took potential knock-on delays into account.

Another feature of the general approach is that it does not depend on the details of the simulation of the airline (i.e. the simulation is almost a black box). In the example problems the simulation had the capability to recovery from delays using airline resource swaps, such swaps were applied before disruption scenario data was derived. In effect the approach would work with any airline schedule simulator, provided the assumption that the use of reserve crew is a last resort recovery action is valid.

The SSH approach showed that the individual scenarios included in the model is at least as important as the number of scenarios, as this heuristic scenario selection approach yielded solutions of higher quality on average compared to the $MIPSSM$ approach, with only a fraction of the input disruption scenarios. The Probabilistic model ($Prob$) represented an entirely different approach to the $MIPSSM$ and gave comparable results, suggesting both approaches have their own merits. In general it was found that the $MIPSSM$, SSH and $Prob$ approaches gave results that were very similar on average, however the $MIPSSM$ based approaches had lower solution stability from one run to the next due to the stochastic nature of these approaches, but significantly outperformed the $Prob$ approach in some cases.

7 Future work

The $MIPSSM$ based approaches rely on stochastic inputs, this is both the greatest strength and weakness of these approaches. Future work includes investigating how to increase solution stability by improving the process of selecting which disruption scenarios to include in the solution phase of the $MIPSSM$ based approaches.

References

1. Christopher Bayliss, Geert De Maere, Jason Atkin, and Marc Paelinck. Probabilistic Airline Reserve Crew Scheduling Model. In Daniel Delling and Leo Liberti, editors, *12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, volume 25 of *OpenAccess Series in Informatics (OASIcs)*, pages 132–143, Dagstuhl, Germany, 2012. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
2. Alexandre Boissy. Crew reserve sizing. In *Agifors syposium.*
3. Jeffrey E. Dillon and Spyros Kontogiorgis. Us airways optimises the scheduling of reserve flight crews. *Interfaces*, 29(5):pp123, September/October 1999.
4. Christian Liebchen, Marco Lubbecke, Rolf H. Mohring, and Sebastian Stiller. Recoverable robustness. Technical report, August 2007.
5. Marc Paelinck. KLM cabin crew reserve duty optimisation. In *Agifors proceedings*, 2001.
6. Sergey Shebalov and Diego Klabjan. Robust airline crew pairing: Moveup crews. *Transportation science*, 2006.
7. Milind G. Sohoni, Ellis L. Johnson, and T. Glennn Bailey. Operational airlines reserve crew scheduling. *Journal of Scheduling*, 2006.
8. H. Paul. Williams. *Model building in mathematical programming.* Wiley, 2002.