Airline Reserve Crew Scheduling Under Uncertainty

by Christopher Bayliss, BSc

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

August 2015

Contents

Acronyms **3**

1 Introduction **4**
1.1 Background information . . . 5
1.2 Goals . . . 7
1.3 Contributions . . . 7
1.4 Publications . . . 8
1.4.1 Journal papers . . . 9
1.4.2 Conference papers . . . 9
1.4.3 Conference abstracts . . . 9
1.5 Thesis structure . . . 10
1.6 Project flow diagram: flow of ideas . . . 12
1.7 Chapter summary . . . 14

2 Literature review **15**
2.1 Airline scheduling . . . 16
2.1.1 Schedule design . . . 17
2.1.2 Fleet assignment . . . 18
2.1.3 Maintenance routing . . . 19
2.1.4 Integrated airline scheduling . . . 20
2.1.5 Crew scheduling . . . 20
2.2 Robust airline scheduling . . . 24
2.2.1 Robust fleet assignment . . . 25
2.2.2 Robust crew scheduling . . . 26
2.2.3 Other approaches to increasing airline schedule robustness . . . 27
2.3 Reserve crew scheduling . . . 30
2.4 Airline operations . . . 32
2.4.1 The crew recovery problem . . . 33
2.4.2 Aircraft re-routing . . . 35
2.4.3 Integrated airline recovery . . . 35
2.4.4 Disruption management . . . 36
2.4.5 Airline operations control . . . 37
2.5 Modelling uncertainty . . . 37
2.5.1 Introduction to probability theory . . . 37
2.5.2 Delay propagation . . . 38
2.5.3 Simulation . . . 39
2.5.4 Statistical distributions . . . 41

2.6 Solution methodologies for deterministic problems 42
2.6.1 Exact methods: Linear and Integer programming . . 42
2.6.2 Meta-heuristics 44
2.6.3 Hybrid approaches 46
2.7 Solution methodologies for stochastic problems 46
2.7.1 Stochastic programming 46
2.7.2 Robust optimisation 48
2.7.3 Variants of robust optimisation 48
2.7.4 Methods for multi-stage decision making 50
2.8 Aspects of problem solving 52
2.8.1 Modelling 52
2.8.2 Solution space 53
2.8.3 Analogies with other problem domains 53
2.9 Chapter summary 54

3 Problem description and definitions **55**
3.1 General problem formulation 55
3.2 Offline reserve crew scheduling 57
3.3 Online reserve policy 60
3.4 KLM specific problem 61
3.5 Definitions 63
3.5.1 Cancellation measure of a delay 63
3.5.2 Rule of thumb reserve policies 64
3.5.3 Heuristic reserve crew scheduling approaches 65
3.5.4 Frequently used solution methodologies 65
3.6 Chapter summary 67

4 Simulation model and methodologies **68**
4.1 Goals 69
4.2 Simulation assumptions and assertions 69
4.3 Simulation flow chart 75
4.4 Airline recovery 76
4.4.1 Delay recovery 76
4.4.2 Reserve teams constructed and used to replace delayed connecting crew 77
4.4.3 Absence recovery 78
4.5 Simulation input 78
4.5.1 Input schedules 78
4.5.2 Stochastic elements of the simulation 79
4.6 Simulation output 79
4.7 Simulation based methodologies 81
4.7.1 The *area under the graph* approach to reserve crew scheduling 81
4.7.2 Simulation reserve policy: SIM 82
4.7.3 Look up table policy: LUT 83
4.8 Chapter summary 84

5 Probabilistic crew absence model **85**
5.1 The simplified problem 86
5.1.1 Assumptions 86
5.1.2 Fundamental equations 88
5.2 The model 89
5.2.1 Calculating the effects of a reserve crew schedule on the probabilities of crew unavailability 89
5.3 Surrogate objective function investigation 90
5.4 Solution methodology investigation 91
5.4.1 Description of solution methods 92
5.4.2 Solution method results 95
5.5 Possible model improvements 97
5.6 Chapter summary 97

6 Improved probabilistic crew absence model **99**
6.1 New probabilistic crew absence model formulation 100
6.1.1 Crew pairings 102
6.1.2 Reserve crew feasibility 102
6.1.3 Reserve use induced delay 103
6.1.4 Evaluating expected cancellations associated with a given reserve crew schedule 104
6.1.5 Enumerating feasible combinations of reserve crew and associated probabilities 105
6.1.6 Solution space 109
6.1.7 Improved model 109
6.2 Experimental results 111
6.2.1 Test instance 111
6.2.2 Experiment design 112
6.2.3 Cancellation prediction accuracy 113
6.2.4 Reserve crew scheduling application 114
6.2.5 Extra performance measures and alternative approaches 115
6.2.6 Results summary 118
6.3 Extended formulation: Including aircraft fleet types and crew ranks and qualifications 118
6.3.1 Required modifications 119
6.4 Generalised reserve policy 122
6.4.1 Generalised reserve policy parameters 123
6.4.2 Generalised reserve policy parameter space 125
6.4.3 Experimental design 125
6.4.4 GRP parameter experiment results 126
6.5 Chapter summary 129
6.6 The CAM used in subsequent chapters 129

7 Probabilistic crew delay model **131**
7.1 Motivation for a probabilistic model of crew-related delay . . 132
7.2 Model overview 132
7.3 The probabilistic crew delay model 133
7.3.1 Phase 1: Input Generation 134
7.3.2 Phase 2: Probabilistic Crew Delay Optimisation . . . 136

7.3.3 Phase 3: Validation 139
7.4 Results 139
7.4.1 Data Instances 139
7.4.2 Convergence 140
7.4.3 Accuracy of the model 141
7.4.4 Comparison with other approaches 141
7.5 Chapter summary 144

8 Statistical delay propagation model 145
8.1 The Statistical Delay Propagation Model 146
8.1.1 Assumptions 148
8.1.2 Overview of the $SDPM$ 149
8.1.3 ETA matrices 153
8.1.4 Modelling aspects of the $SDPM$ 154
8.1.5 Calculating the probabilities of resource availability and resource swap availability 158
8.1.6 Calculating the cumulative probability of departure during a given time interval 160
8.1.7 Updating departure distributions 161
8.1.8 Calculating an arrival time distribution for a hub-spoke-hub cycle 163
8.2 Experimental results 164
8.2.1 Description of $SDPM$ applications 164
8.2.2 Test instance 165
8.2.3 Modelling accuracy of the $SDPM$ 165
8.2.4 The effect of interval size 169
8.2.5 Scheduling and policy applications of the $SDPM$ . . 171
8.3 Including Aircraft fleet types, crew ranks and qualifications . 176
8.3.1 Experiment results for the case of multiple fleet types, crew ranks and qualifications 176
8.3.2 The effect of interval size on solution quality 176
8.4 Chapter summary 177

9 Mixed integer programming simulation scenario model 179
9.1 Overview of the MIPSSM 181
9.1.1 Stages of the MIPSSM approach 181
9.1.2 Cancellation measure of a delay 183
9.1.3 Disruption scenarios 184
9.1.4 Feasible reserve instances 184
9.2 Disruption scenario generation simulation 185
9.2.1 Simulation 185
9.2.2 Simulation derived scenarios 187
9.3 The MIPSSM's Mixed Integer Programming formulation . . 191
9.3.1 MIPSSM formulation 191
9.4 MIPSSM modifications 193
9.4.1 Alternative objectives for the MIPSSM 193
9.4.2 Scenario Selection Heuristic 194
9.5 Optimal reserve use policy derivation 195
9.6 Experimental results 197

9.6.1 Experiment design 197
9.6.2 Investigating the effect of varying the number of reserve crew available for scheduling 197
9.6.3 Other performance measures and solution reliability . 199
9.7 The effect of scenario sets on reserve crew schedule quality . 202
9.7.1 Attributes of sets of scenarios 203
9.7.2 Testing pools of scenarios 204
9.7.3 Algorithms based on the results of the scenario set investigation 206
9.8 Extending the MIPSSM approach for the case of fleets, crew ranks and qualifications 207
9.8.1 Example problem 207
9.8.2 Extended simulation generation of disruption scenarios 207
9.8.3 Extended MIPSSM formulation 209
9.8.4 Results 210
9.9 Future work 210
9.9.1 Iterative solution approach to $MIPSSM$ 210
9.10 Chapter summary 211

10 Comparison of all reserve crew scheduling and policy approaches 212
10.1 The test instances 212
10.2 Description of the reserve crew scheduling approaches being compared 214
10.2.1 Probabilistic approaches 214
10.2.2 MIPSSM based approaches 216
10.2.3 Other approaches 216
10.3 Experimental design 216
10.4 Phase 1: Reserve crew scheduling results 217
10.4.1 Probabilistic model results 217
10.4.2 MIPSSM results 222
10.5 Phase 2: Comparing all approaches for reserve crew scheduling and reserve policies 225
10.5.1 Applying policies in instances of crew absence 228
10.5.2 Results tables 229
10.6 Results summary 235

11 Conclusion 237
11.1 Probabilistic reserve crew scheduling 237
11.2 Scenario-based reserve crew scheduling 240
11.3 Comparison of the probabilistic and scenario-based approaches 241
11.4 Online reserve policies 242
11.5 General insights gained 243
11.6 Advice for KLM 244
11.7 Thesis summary 244

12 Potential future extensions 245
12.1 Reserve crew scheduling 245
12.1.1 Probabilistic based approaches 245
12.1.2 Scenario-based approaches 246
12.1.3 Hybridised approaches 246
12.1.4 Extended formulations 247
12.1.5 Integrated crew scheduling and reserve crew scheduling 248
12.1.6 Other applications 248
12.2 Reserve Policies 249
12.2.1 Approximate dynamic programming 249
12.2.2 Hybridised reserve policies 250
12.3 Integrated reserve crew scheduling and reserve policy optimisation 250

A Extra results for the probabilistic crew delay model 259
A.1 Experimental results representative of the three main types of methods of reserve crew scheduling for each of the 25 data instances 259
A.1.1 *Prob 1* results 259
A.1.2 *Area 1* results 260
A.1.3 *Results for uniform distribution* 262

B GRP parameter experiment results: Average cancellation measure sensitivity to policy parameter sets used in reserve crew scheduling and online 264

C Conditions for a delay-reducing resource swap 267

D Additional experimental results for the *SDPM* 269
D.1 Modelling accuracy of the *SDPM* 269
D.2 Scheduling and policy applications of the *SDPM* 272
D.3 Reserve policy application comparison 276
D.3.1 Reserve crew policies under consideration 276
D.3.2 Reserve policy application results 276

E Additional statistical delay propagation model test results 279
E.1 Additional initial prediction tests 279

F Extra results for the statistical delay propagation model 283

G Additional results for the comparison of all approaches 288
G.1 10 fold cross-validation for test instances 1 to 6 288
G.2 Average cancellation measure plots 289
G.3 Delay and cancellation performance measures of the probabilistic models 291

H Variable cancellation threshold 293

List of Figures

1.1 Flow of ideas 13

2.1 Reserve crew scheduling centred view of airline scheduling and operations 15
2.2 Delays can spread to multiple flights if "resources split" causing a "switch delay" (dashed lines). This occurs when the the crew and aircraft of a delayed arrival are assigned to different subsequent flights. 24

3.1 Block times and their constituent parts 57
3.2 The effect of the delay exponent (used to penalise delay) on average delay and cancellation rate 64

4.1 Simulation flow chart 75
4.2 Airline schedule visualisation, the flight legs to the hub (yellow) or to a spoke (green) which are assigned to aircraft (rows) and crew (blue lines) 80
4.3 Example recovery actions graph 81
4.4 Reserve demand graph and equal total demand intervals . . 82

5.1 Range of cancellation rates provided by the dynamic programming solution, the even distribution heuristic and with no reserves allocated 96

6.1 A growing reserve crew combination tree 106
6.2 Cancellation predictions from evaluations of the $SPCAM$, CAM and refined probabilistic crew absence model ($CAM*$) compared to those derived from repeat simulations 110
6.3 Cancellation predictions from evaluations of the $SPCAM$, CAM and $CAM*$ compared to those derived from repeat simulations 114
6.4 The effect of each variant of the probabilistic crew absence model and the search heuristics used to schedule reserve crew on the resultant simulation derived average cancellation measure 115
6.5 Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 1 127

6.6 Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 6 128

7.1 Average root mean squared error of P elements derived from simulation using 10-fold cross validation for a variety of sample sizes 140
7.2 CDM predicts crew-related delay observed in validation simulations 141

8.1 Delay propagation cycle 146
8.2 Propagating delay distribution 147
8.3 Resource swaps in an ETA matrix context 155
8.4 $SDPM$ delay predictions compared to those derived from simulation 166
8.5 The predicted average delay reductions due to allowing swap recovery actions 166
8.6 The predicted average delay reductions due to using reserve crew to absorb delays 167
8.7 The predicted average delay reductions due to using reserve crew to absorb delays 168
8.8 The effect of interval size (W) on prediction accuracy and evaluation time 169
8.9 The effect of interval size (W) on reserve crew schedule quality and reserve policy quality 171

9.1 Sequential stages of the $MIPSSM$ approach to scheduling airline reserve crew 182
9.2 Feasible reserve instance attributes 184
9.3 Flow chart of the simulation used to derive disruption scenarios 186
9.4 The effect of the number of reserve crew which are scheduled on the solution quality of different solution approaches . . . 198
9.5 The effect of the $MIPSSM$ derived reserve use policy . . . 199
9.6 Percentile cancellation measures 201
9.7 Solution reliability of $MIPSSM$ based methods compared to $Prob$ 202
9.8 Flowchart of the population of three pools of scenarios . . . 203
9.9 The effect of the pool from which scenarios are selected and the number of scenarios selected on the average cancellation measure associated with the reserve crew schedule derived from the given set of scenarios 205

10.1 The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 1 with the absence only policy 218

10.2 Cancellation measures for each solution methodology for all combinations of test schedule, reserve policy and evaluator, in descending order of the average cancellation measure for each combination of test schedule, reserve policy and evaluator 219
10.3 Cancellation measures for each probabilistic evaluator for all combinations of test schedule, reserve policy and solution methodology, in descending order of the average cancellation measure for each combination of test schedule, reserve policy and solution methodology 220
10.4 Average delay and cancellation performance of the reserve crew schedules derived using the SDM, CDM and SDPM probabilistic evaluators 221
10.5 Average cancellation measures for MIPSSM based approaches 224
10.6 Average cancellation measures for all tested approaches to reserve crew scheduling and reserve policies, for each test instance 226

B.1 Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 2 264
B.2 Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 3 265
B.3 Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 4 265
B.4 Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 5 266

D.1 Straight forward delay propagation 270
D.2 The predicted delay reduction due to swap recovery actions from simulation and the $SDPM$ 270
D.3 Predicted decrease in average delays when reserve crew can be used to cover for delayed crew as well as covering for absent crew 271
D.4 Predicted increase in cancellation probabilities when reserve crew can be used to cover for delayed crew as well as covering for absent crew 271
D.5 The effect of interval size on prediction accuracy and evaluation time 272
D.6 Average cancellation measures when reserve schedules are used in conjunction with different reserve policies in schedule 1274
D.7 Average cancellation measures when reserve schedules are used in conjunction with different reserve policies in schedule 2274

D.8 Average cancellation measures when reserve schedules are used in conjunction with different reserve policies in schedule 3 275
D.9 Average cancellation measures when reserve schedules are used in conjunction with different reserve policies in schedule 4 275

E.1 Predicted average delays and cancellation rates for test instance 1 derived from repeat simulations and the SDPM . . 279
E.2 Predicted average delays and cancellation rates for test instance 2 derived from repeat simulations and the SDPM . . 280
E.3 Predicted average delays and cancellation rates for test instance 3 derived from repeat simulations and the SDPM . . 280
E.4 Predicted average delays and cancellation rates for test instance 4 derived from repeat simulations and the SDPM . . 281
E.5 Predicted average delays and cancellation rates for test instance 5 derived from repeat simulations and the SDPM . . 281
E.6 Predicted average delays and cancellation rates for test instance 6 derived from repeat simulations and the SDPM . . 282

F.1 The effect of interval size on solution quality in test instance 1 283
F.2 The effect of interval size on solution quality in test instance 2 284
F.3 The effect of interval size on solution quality in test instance 3 285
F.4 The effect of interval size on solution quality in test instance 4 285
F.5 The effect of interval size on solution quality in test instance 5 286
F.6 The effect of interval size on solution quality in test instance 6 286

G.1 Convergence of the average RMSE of the average cancellation measure for different fold sizes 288
G.2 The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 2 with the absence only policy 289
G.3 The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 3 with the absence only policy 289
G.4 The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 4 with the absence only policy 290
G.5 The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 5 with the absence only policy 290
G.6 The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 6 with the absence only policy 291
G.7 Average delay and cancellation performance of the reserve crew schedules derived using the SDM, CDM and SDPM probabilistic evaluators (3 day test instances) 291
G.8 Average delay and cancellation performance of the reserve crew schedules derived using the SDM, CDM and SDPM probabilistic evaluators (7 day test instances) 292

List of Tables

2.1 Airline scheduling terminology . . . 17
2.2 Classification of the existing literature on reserve crew scheduling . . . 30

3.1 Problem description notation . . . 56

5.1 Notation . . . 87
5.2 Objective functions . . . 90
5.3 Reserve utilisation and cancellation rates in 2000 simulations (50000 flights) . . . 91
5.4 Objective values, simulation coverage levels and solution times for a variety of solution methods . . . 95

6.1 Notation . . . 101
6.2 Comparison of approaches to reserve crew scheduling using more performance measures . . . 116
6.3 Reserve crew qualification groups and the fleets they are qualified for . . . 119
6.4 Fleet crew requirements (FCR) . . . 119
6.5 Generalised reserve policy notation . . . 123
6.6 The GRP policy parameter combinations that minimises average cancellation measure . . . 127

7.1 Notation . . . 133
7.2 Simulation derived performance measures for a variety of solution methods . . . 143

8.1 Definitions . . . 150
8.2 Overall delay and cancellation prediction accuracy results for various simulation and $SDPM$ configurations . . . 168
8.3 Average cancellation measures for different combinations of configurations of the $SDPM$ used for reserve crew scheduling and as a reserve policy . . . 173
8.4 Extra performances measures for the interesting results of Table 8.3 . . . 173
8.5 Optimal trade-off interval sizes for the Chapter 10 test instances 177

9.1 Delay cancellation measure related notation . . . 183
9.2 Schedule notation . . . 185
9.3 Disruption scenario notation . . . 187
9.4 $MIPSSM$ formulation notation . . . 191

9.5 Notation for the $MIPSSM$ derived policy 195
9.6 Performance measure averages from 20 repeats 200
9.7 Notational changes required for extending the $MIPSSM$ approach for the case of multiple fleets, crew ranks and qualifications 208

10.1 Test instance properties 213
10.2 Details of the probabilistic models used as evaluators in various search methodologies 214
10.3 Average cancellation measures for each solution methodology and probabilistic evaluator 223
10.4 Number of times two policy variants outperform one another 228
10.5 1 day test instance performance measures for the best probabilistic and MIPSSM derived reserve crew schedules used in conjunction with each reserve policy 230
10.6 3 day test instance performance measures for the best probabilistic and MIPSSM derived reserve crew schedules used in conjunction with each reserve policy 232
10.7 7 day test instance performance measures for the best probabilistic and MIPSSM derived reserve crew schedules used in conjunction with each reserve policy 234

A.1 Prob 1 cancellation rate results from 25 schedule instances . 259
A.2 Prob 1 reserve utilisation rate results from 25 schedule instances 260
A.3 Prob 1 average crew delay results from 25 schedule instances 260
A.4 Prob 1 average total delay results from 25 schedule instances 260
A.5 Prob 1 probability of delay over 30 minutes results from 25 schedule instances 260
A.6 Area 1 cancellation rate results from 25 schedule instances . 261
A.7 Area 1 reserve utilisation rate results from 25 schedule instances 261
A.8 Area 1 average crew delay results from 25 schedule instances 261
A.9 Area 1 average total delay results from 25 schedule instances 261
A.10 Area 1 probability of delay over 30 minutes results from 25 schedule instances 262
A.11 Uniform cancellation rate results from 25 schedule instances 262
A.12 Uniform reserve utilisation rate results from 25 schedule instances 262
A.13 Uniform average crew delay results from 25 schedule instances 262
A.14 Uniform average total delay results from 25 schedule instances 263
A.15 Uniform probability of delay over 30 minutes results from 25 schedule instances 263

D.1 Test instance properties 269
D.2 Average cancellation measures for different combinations of configurations of the $SDPM$ used for reserve crew scheduling and as a reserve policy averaged over 10 repeats for each configuration used for reserve crew scheduling 273

D.3 Actual event times version of the 2 day test instance. Average cancellation measures for different combinations of configurations of the $SDPM$ used for reserve crew scheduling and as a reserve policy averaged over 10 repeats for each configuration used for reserve crew scheduling 273
D.4 Reserve policy performance measures: Test instance 1 277
D.5 Reserve policy performance measures: Test instance 2 277
D.6 Reserve policy performance measures: Test instance 3 277
D.7 Reserve policy performance measures: Test instance 4 278

Abstract

This thesis addresses the problem of airline reserve crew scheduling under crew absence and journey time uncertainty. This work is primarily concerned with the allocation of reserve crew to standby duty periods. The times at which reserve crew are on duty, determine which possible crew absence or delay disruptions they can be used to absorb. When scheduling reserve crew, the goal is to minimise the expected levels of delay and cancellation disruptions that occur on the day of operation. This work introduces detailed probabilistic models of the occurrence of crew absence and delay disruptions and how reserve crew are used to absorb such disruptions. Firstly, separate probabilistic models are developed for crew absence and delay disruptions. Then, an integrated probabilistic model of absence and delay disruptions is introduced, which accounts for: delays from all causes; delay propagation; cancellations resulting from excessive delays and crew absence; the use of reserve crew to cover such disruptions given a reserve policy; and the possibility of swap recovery actions as an alternative delay recovery action. The model yields delay and cancellation predictions that match those derived from simulation to a high level of accuracy and does so in a fraction of the time required by simulation. The various probabilistic models are used in various search methodologies to find disruption minimising reserve crew schedules. The results show that high quality reserve crew schedules can be derived using a probabilistic model.

A scenario-based mixed integer programming approach to modelling operational uncertainty and reserve crew use is also developed in this thesis and applied to the problem of reserve crew scheduling. A scenario selection heuristic is introduced which improves reserve crew schedule quality using fewer input scenarios.

The secondary objective of this thesis is to investigate the effect of the reserve policy used on the day of operation, that is, determining when and which reserve crew should be utilised. The questions of how reserve policies can be improved and how they should be taken into account when scheduling reserve crew are addressed. It was found that the approaches developed for reserve crew scheduling lend themselves well to an online application, that is, using them to evaluate alternative reserve decisions to ensure reserve crew are used as effectively as possible. In general it is shown that 'day of operation' disruptions can be significantly reduced through both improved reserve crew schedules and/or reserve policies. This thesis also points the way towards future research based on the proposed approaches.

Acknowledgements

The research presented in this thesis originated from a previous collaboration between Geert De Maere of the University of Nottingham and Marc Paelinck of Air France-KLM, to them I owe thanks for this great opportunity they have provided me with. This research was performed under the supervision of Jason Atkin, Geert De Maere and Marc Paelinck. I offer my sincerest gratitude to my supervisors for the time and effort they spent providing feedback on the work I have done during this research project. Without which my writing style would be much worse off. To Jason Atkin and Geert De Maere, special thanks are owed for the hours of interesting discussions and sometimes heated debates on various aspects of this project and for providing continued support and guidance throughout this research project.

I would also like to extend my gratitude to my brothers, Kevin for his comments and questions when providing an audience when I was practising for conference presentations and to my brother Anthony for always being there for me. My parents are also owed a great debt of gratitude for their unending support and encouragement and for their understanding over the course of this research project.

Many thanks are owed to the staff at KLM's crew planning and assignment department and Marc Paelinck in particular for providing data and perspectives based on first hand experience. Thanks are also owed to the ASAP research group for providing a comfortable and friendly working environment. I am also grateful to the technical support group in the computer science department and the administrative staff at ASAP. Finally I would like to thank the EPSRC LANCS initiative (grant ref EP/F033214/1) for the funding which has made this research possible.

Acronyms

CAM Crew Absence Model.

CDM Crew Delay Model.

ETA Estimated Time of Arrival.

FRQ Fleets Ranks and Qualifications.

GRP Generalised default Reserve Policy.

LUT Look Up Table policy.

$MIPSSM$ Mixed Integer Programming Simulation Scenario Model.

$SDPM$ Statistical delay propagation model.

SIM Simulation reserve policy.

$SPCAM$ Simplified Probabilistic Crew Absence Model.

abs only Absence Only policy.

$default$ Default reserve policy.

Chapter 1

Introduction

This thesis presents the findings from a PhD project on the subject of airline reserve crew scheduling under uncertainty. Airlines operate in an uncertain environment due to the effects of weather, congestion (ground and air), unexpected aircraft maintenance and crew unavailability. This project is primarily concerned with disruptions caused by crew unavailability. These occur in the form of absent crew, for example, due to illness and in the form of delayed crew. Both absent crew and crew-related delay disruptions can be absorbed by replacing the affected crew with reserve crew. Reserve crew are spare crew who are on standby at specific locations to cover for delayed or absent crew at short notice. As crew represent the second largest cost to airlines (fuel is the leading cost), using crew efficiently is important to the profitability of an airline. This encourages airlines to operate tight schedules that are susceptible to the propagation of disruptions. Reserve crew scheduling adds a layer of recoverability that is essential for the smooth running of an airline. Reserve crew are also used for operations left unassigned till the last minute (known as open flights). These unassigned flights provide a layer of schedule flexibility which enables the use of unused reserve crew as well as disrupted crew who may have been replaced with reserve crew.

This project introduces models of crew absence and delay propagation and the impact that reserve crew schedules have on absorbing such disruptions. The resultant models are used to search the solution space of reserve crew schedules (the times at which the reserve crew will be on standby duty). The main overall contribution of this thesis is that it advances the level of explicit detail included in the models of crew-related disruptions and the modelling of the impact that a given reserve crew schedule has on absorbing such disruptions (see Section 1.3). Three main approaches to reserve crew scheduling are developed and investigated, including probabilistic, scenario-based and heuristic approaches. Probabilistic approaches use the probabilities of disruptions to calculate the probabilities of reserve crew requirements and then use the probabilities of reserve crew availability to work out the probabilities that crew-related disruptions still occur, and continue in this fashion iteratively. Scenario-based models consider an explicit set of possible disruption scenarios and try to find the reserve crew schedule that absorbs the most disruptions over each of those scenarios simultaneously. Heuristic approaches include rule of thumb based scheduling approaches as opposed to optimisation based scheduling and often provide a good initial benchmark performance level to compare the more advanced approaches to.

As a secondary objective, this project also investigates reserve crew use from an online perspective in the form of reserve policies. Reserve policies determine whether reserve crew should be used for any given crew-related disruption and which combination of individual reserve crew members should be used. Online reserve policies need to be adaptive to the current conditions. They also need to be able to make globally informed decisions, even if this means not absorbing avoidable disruptions now, in order to ensure that larger disruptions that may happen later can be covered. Such a decision is referred to as reserve holding. In this project many of the approaches to reserve crew scheduling can be adapted and used online as reserve policies which ensure that reserve crew are used to absorb the most important/largest disruptions.

The main contributions/findings are discussed in Section 1.3 and include the introduction of: reserve crew scheduling approaches based on highly detailed probabilistic models of absence and delay disruptions and reserve crew used to absorb them; and scenario-based approaches including a scenario selection heuristic solution approach. Additionally, the reserve policy investigations showed that reserve crew use can be improved using adapted versions of the offline reserve crew scheduling approaches online, and that certain rule of thumb policies also lead to good levels of disruption absorption.

Chapter structure

Section 1.1 gives some background information for this research. Section 1.2 states the goals of this research. Section 1.3 summarises the main contributions from this research. Section 1.4 lists the publications resulting from this research. Section 1.5 explains the structure of this thesis. Section 1.6 discusses the flow of ideas throughout this research. Section 1.7 summarises the main points from this chapter.

1.1 Background information

This project originated from a collaboration between the University of Nottingham and KLM (now Air France-KLM[1]), and was funded by the EPSRC LANCS initiative (grant ref EP/F033214/1). The project was initiated with a visit to KLM headquarters in Amstelveen, with a series of meetings with key figures in the KLM crew planning and assignment department. Key personnel at KLM have worked on the problem of reserve crew scheduling, introducing fundamental changes to KLM reserve crew scheduling practices.

[1]KLM and Air France still operate as two separate airlines but are owned by the Air France-KLM group. In this thesis, the described practices refer to the specific case of KLM

Much of the following tries to extend their work and address research questions which were discussed in that first meeting.

Overview of previous work

The vast majority of research on airline scheduling is based on the assumption that everything goes to plan on the day of operation. A recent trend in the literature is to acknowledge operational uncertainty in the scheduling phase of operations, otherwise known as robust scheduling (see Section 2.2). Such work typically shows that optimising operational costs as opposed to planned costs is beneficial both in theory and practice. The reason is that, due to disruptions, operational costs usually exceed planned costs, especially in unrobust solutions.

Previous work on robust scheduling (see Section 2.2), when applied to airline scheduling, typically aims to schedule resources in such a way that the schedule has a minimal potential for developing infeasibilities or so that schedule feasibility can easily be regained in the event of disruptions. Reserve crew can augment the robustness of an airline schedule. Not all disruptions can be addressed in the scheduling phase by modifying certain properties of the schedule. Crew absence is an example of such a disruption. Models for reserve crew scheduling and the use of reserve crew can be used to address a schedule's risk of infeasibility due to crew absence.

Before this project, much work on reserve crew scheduling (see Section 2.3) was based on minimising the requirement for reserve crew by better understanding the causes of disruptions that require reserve crew and forecasting the requirement of reserve crew such that a balance is achieved between the cost of reserve crew provision and the cost of uncovered disruptions. In this thesis, reserve crew are scheduled with the aim of minimising the expected levels of delays and cancellations that result from crew absence and delay propagation after the application of recovery actions.

Planning stages

KLM's approach to crew scheduling consists of four stages, manpower planning, pairing, assignment and tracking and control. Manpower planning is the task of determining the total number of employees required over the mid to long term future to carry out expected future operations. Crew pairing determines sequences of flights that can feasibly be operated by individual members of crew such that all tasks are included in separate crew pairings, a task which is performed approximately a month before the day of operations. Assignment determines which crew operate each crew pairing. The tracking and control group reacts to crew disruptions by rescheduling crew to ensure the crew schedule remains feasible.

Reserve blocks

At KLM, cabin crew are required by contract to perform several reserve blocks per year. Reserve blocks are crew assignments in which crew are stationed at their domicile (home station/airport) and are on standby ready to replace disrupted crew at short notice. The structure of reserve blocks is described in more detail in Section 3.4 of the problem description chapter. One of the main goals of this research is to investigate approaches to the scheduling of reserve blocks, i.e. specifying the numbers of reserve blocks starting each day and the standby duty times associated with each of those reserve blocks.

Current approach

KLM's current approach to reserve crew scheduling involves extrapolating historical data on reserve crew requirements into the future, they have equal numbers of starting reserve crew on each day and state that currently no advanced approaches are used to schedule reserve crew duty start times. The actual duty patterns are determined manually through trial an error, experience and intuition.

1.2 Goals

The goals of this thesis are as follows:

- Develop methods of modelling the occurrence of crew-related disruptions including how reserve crew are used in response to these.
- Use the resultant models of airline crew disruption uncertainty and possible recovery actions to schedule reserve crew in a way that minimises:
  - The chance that the crew schedule becomes infeasible.
  - The expected number of cancellations and the expected level of delay propagation
- Advise KLM on any insights and recommendations for best practices for reserve crew scheduling.
- Investigate alternative reserve policies, the interaction between the reserve crew schedule and the reserve policy and how best to take the reserve policy into account during reserve crew scheduling.

1.3 Contributions

The contributions of this thesis include the following:

- (Main overall contribution). The proposed approaches make use of all of the available information in order to determine the reserve crew schedule.
- Detailed probabilistic models of disruptions and reserve crew use are developed (Chapters 5 to 8).
  - Reserve demand modelled at a per flight resolution.

  - Probabilistic models were developed for reserve crew demand and usage, with separate models for crew absence and delay disruptions.
  - Development of a statistical model of delay propagation which models event time uncertainty as a function of the: airline's schedule; reserve crew schedule; recovery policy (including reserve crew and resource swap recovery actions) and journey time uncertainty.
  - Integration of probabilistic crew absence and delay propagation models, which are capable of rapid and highly accurate delay and cancellation predictions.
  - The probabilistic models are used to derive high quality reserve crew schedules.
  - The probabilistic models are applied in an online context to evaluate alternative reserve decisions to make globally informed recovery decisions.

- Scenario-based models are developed (Chapter 9).
  - A framework for disruption scenario generation is developed which provides input scenarios for a mixed integer programming approach to reserve crew scheduling.
  - A scenario selection heuristic is introduced which results in higher quality reserve crew schedules with fewer input scenarios than the standard formulation.
  - Investigations are carried out into the most effective objective functions and the effect of different types of input disruption scenarios on the quality of the derived reserve crew schedules.

- Investigation of online reserve policies (parts of Chapters 6, 8 and 9).
  - Development and investigation of reserve crew selection policies and reserve crew holding policies.
  - The encoding of reserve selection policies within reserve crew scheduling models.
  - The application of probabilistic models for reserve crew scheduling in an online context to provide globally informed decision support for reserve crew use.

- Comparison of all considered approaches to reserve crew scheduling and reserve policies when applied to a range of realistic test instances (Chapter 10) leading to profound insights into the effectiveness of the different approaches when applied in a range of situations.

1.4 Publications

The publications resulting from this work are as follows.

1.4.1 Journal papers

- Revised and resubmitted to Annals of Operations Research
  C. Bayliss, G. De Maere, J.A.D. Atkin, M. Paelinck "A Simulation Scenario Based Mixed Integer Programming Approach to Airline Reserve Crew Scheduling Under Uncertainty"

- Revised and resubmitted to the journal of Computers and Operations Research
  C. Bayliss, G. De Maere, J.A.D. Atkin, M. Paelinck "Extended Probabilistic Crew Absence and Reserve Crew Recovery Model"

- Submitted to Transportation Research Part B: Methodological
  C. Bayliss, G. De Maere, J.A.D. Atkin, M. Paelinck "A Statistical Delay Propagation Model for Airline Reserve Crew Scheduling and Decision Support"

1.4.2 Conference papers

- C. Bayliss, G. De Maere, J.A.D. Atkin, and M. Paelinck "Probabilistic airline reserve crew scheduling" in Proceedings of the 12th Workshop on Algorithmic Approaches for Transportation Modelling Optimization, and Systems (ATMOS 2012), 2012,
  doi: 10.4230/OASIcs.ATMOS.2012.132

- C. Bayliss, G. De Maere, J.A.D. Atkin, M. Paelinck "Scheduling Airline Reserve crew to Minimise Crew Related Delay Using Simulated Airline Recovery and a Probabilistic Optimisation Model" in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, p1944-1950, 2013, doi: 10.1109/SMC.2013.334

- C. Bayliss, G. De Maere, J.A.D. Atkin, M. Paelinck "A Simulation Scenario Based Mixed Integer Programming Approach to Airline Reserve Crew Scheduling Under Uncertainty" in Proceedings of the 10th International Conference of the Practice and Theory of Automated Timetabling , p62-81, 2014

1.4.3 Conference abstracts

- C. Bayliss, G. De Maere, J.A.D. Atkin, M. Paelinck "Probabilistic Airline Reserve Crew Scheduling Model" 3rd Student Conference on Operational Research, Nottingham, 2012

- C. Bayliss, G. De Maere, J.A.D. Atkin, M. Paelinck "A Graph Approach to Modelling the Uncertain Flow of Airline Resources Through an Airline Network" 4th Student Conference on Operational Research, Nottingham, 2014

1.5 Thesis structure

The remainder of this thesis is structured as follows.

Chapter 2 gives a review of the literature related to issues surrounding the problem of airline reserve crew scheduling under uncertainty. The literature review covers previous work on airline scheduling problems, airline operations, work in related fields, the modelling of uncertainty and search methodologies.

Chapter 3 gives a formal definition for the problem of airline reserve crew scheduling under uncertainty as well as a number of key definitions for conventions used throughout this thesis.

Chapter 4 introduces the simulation developed in this project as a tool for investigating the problem of airline reserve crew scheduling, which is also used for validating reserve crew schedules and reserve policies. The simulation is introduced early on as it is used in many of the subsequent chapters for a number of different purposes.

Chapters 5 to 8 cover the probabilistic models of crew-related disruptions and reserve crew used to absorb them.

Chapter 5 introduces a probabilistic model of crew absence and reserve recovery in its simplest form, with simplifying assumptions covering all issues considered to be mere details obscuring the underlying problem. In particular, Chapter 5 introduces an iterative formulation for calculating probabilities of cancellation due to crew absence for a given reserve crew schedule. Search heuristics then use the model to search for the reserve crew schedule that minimises the overall expected number of cancellations. The work of Chapter 5 also resulted in a full conference paper [17].

Chapter 6 presents a reformulation of the simplified probabilistic crew absence model where most of the simplifying assumptions have been removed. The extended formulation allows for the possibility of multiple crew being absent simultaneously from crew pairings and the probabilities that feasible combinations of reserve crew are simultaneously available to cover for such disruptions. Chapter 6 also explicitly models the structure of crew pairings, the logic regarding which combinations of reserve crew can possibly be used for different disruptions as well as acknowledging the fact that the total expected number of absent crew is best described using a probability distribution. Delays which are caused by waiting for reserve crew to begin standby duties are also taken into account in the expected cancellations objective function, using a function which maps delays to a measure of cancellation. The extended probabilistic crew absence model is also applied to the case where reserve crew have ranks and qualifications which limit the roles they can perform and the fleet types they can operate on.

Chapter 7 presents a probabilistic model of crew-related delay disruptions and reserve crew recovery which is in many ways analogous to the simplified probabilistic crew absence model of Chapter 5. Chapter 7 introduces the matrix structure which is used to model delay propagation. Then, the simulation learning phase is described in which the delay propagation matrix is populated. The procedure is also introduced for evaluating reserve crew schedules in terms of their associated expected overall delay absorption, which involves calculating the effect that the given reserve crew schedule has on the probabilities of delay propagation in the matrix. The work of Chapter 7 resulted in a full conference paper [15].

Chapter 8 introduces a statistical model of delay propagation in an airline network, which calculates departure and arrival time distributions for all scheduled flights, whilst taking journey time uncertainty and the airline's recovery policy (including the use of reserve crew and swap recovery actions) into account. Chapter 8 introduces the required operations for calculating how delay distributions propagate through an airline's schedule whilst allowing for the effects of airline recovery actions, scheduled slack and journey time uncertainty. The statistical delay propagation model accounts for both crew absence and journey time uncertainty. Crew absence uncertainty is incorporated into the model using the input parameters derived from the improved probabilistic crew absence model of Chapter 6. The statistical delay propagation model is validated in terms of prediction accuracy, in applications to reserve crew scheduling and when applied as a reserve holding policy.

Chapter 9 introduces an alternative reserve crew scheduling approach to that of the probabilistic models, in the form of a scenario-based mixed integer programming approach. Chapter 9 describes how to derive the disruption scenarios which are the inputs for a mixed integer programming formulation which is solved to find a reserve crew schedule that performs well in a wide range of scenarios. Chapter 9 also investigates the effect of the number and type of scenarios included in the formulation. Several scenario selection algorithms are also given. The work of Chapter 9 resulted in a full conference paper [16].

Chapter 10 compares all approaches to reserve crew scheduling and reserve policies developed in the previous chapters when applied to the most detailed problem instances considered in this project. These problem instances allow for: reserve crew that come in a range of rank and qualification combinations; aircraft that come in a range of fleet types; the crew absence uncertainty; the journey time uncertainty; individual fleet crewing requirements and the airline's recovery policy including swap recovery actions as well as reserve crew use.

Chapter 11 concludes this thesis with a summary of the main insights gained from this work.

Chapter 12 considers the future research directions.

1.6 Project flow diagram: flow of ideas

Figure 1.1 illustrates the flow of ideas during this project. The arrows indicate how the different approaches considered in the project evolved over time. Going from left to right corresponds, approximately, to a chronological account of the ideas explored in the project. The top row of the flow chart shows how the project started by considering a simplified version of the problem. The bottom row draws together the work on the secondary objective of the project, that of investigating online reserve policies. The central band of the flow chart shows how the work on the main objective of the project, that of investigating approaches to offline reserve crew scheduling, developed over the course of the project.

Figure 1.1 shows that the main approaches to offline reserve crew scheduling investigated were probabilistic approaches and scenario-based approaches. Initially probabilistic models were developed separately (to be combined later on) for each type of crew-related disruption (absence and delay), the simplified probabilistic crew absence model was developed first. Then there was an initial attempt at a statistical model of delay propagation, the structure of the initial attempt did not allow for the modelling of swap recovery actions. As a result, the probabilistic crew delay model was developed, which was a learning based approach rather than a theoretical model, and was in many ways analogous to the original simplified probabilistic crew absence model. An alternative approach to modelling the problem was also under consideration, that of scenario-based modelling, this approach motivated the development of the single hub airline simulation tool, which would later be used for solution validation and in simulation based algorithms for offline reserve crew scheduling and as an online reserve policy. The mixed integer programming simulation scenario model (or the $MIPSSM$) used the simulation to provide input disruption scenarios, which were used to form the objective and constraint coefficients of a MIP formulation, which was then solved to find a reserve crew schedule. The $MIPSSM$ was the first model to include both types of crew-related disruptions for which reserve crew can be used to absorb. To do this, the delay cancellation measure function was developed to capture the perceived disruption of a delay of a given size relative to a flight cancellation. The delay cancellation measure function meant that the $MIPSSM$ could be formulated as a single objective problem, the issue of the trade-off between the level of protection against delay and cancellation disruptions was captured by a decision maker parameter in the delay cancellation measure function. The $MIPSSM$ also allowed for the possibility of multiple crew being simultaneously absent from a given crew team (see typical assumptions, top).

At around the same time the probabilistic crew absence model was also extended to allow for the possibility of multiple crew absence per crew pairing. Additionally, the delay cancellation measure function was used in

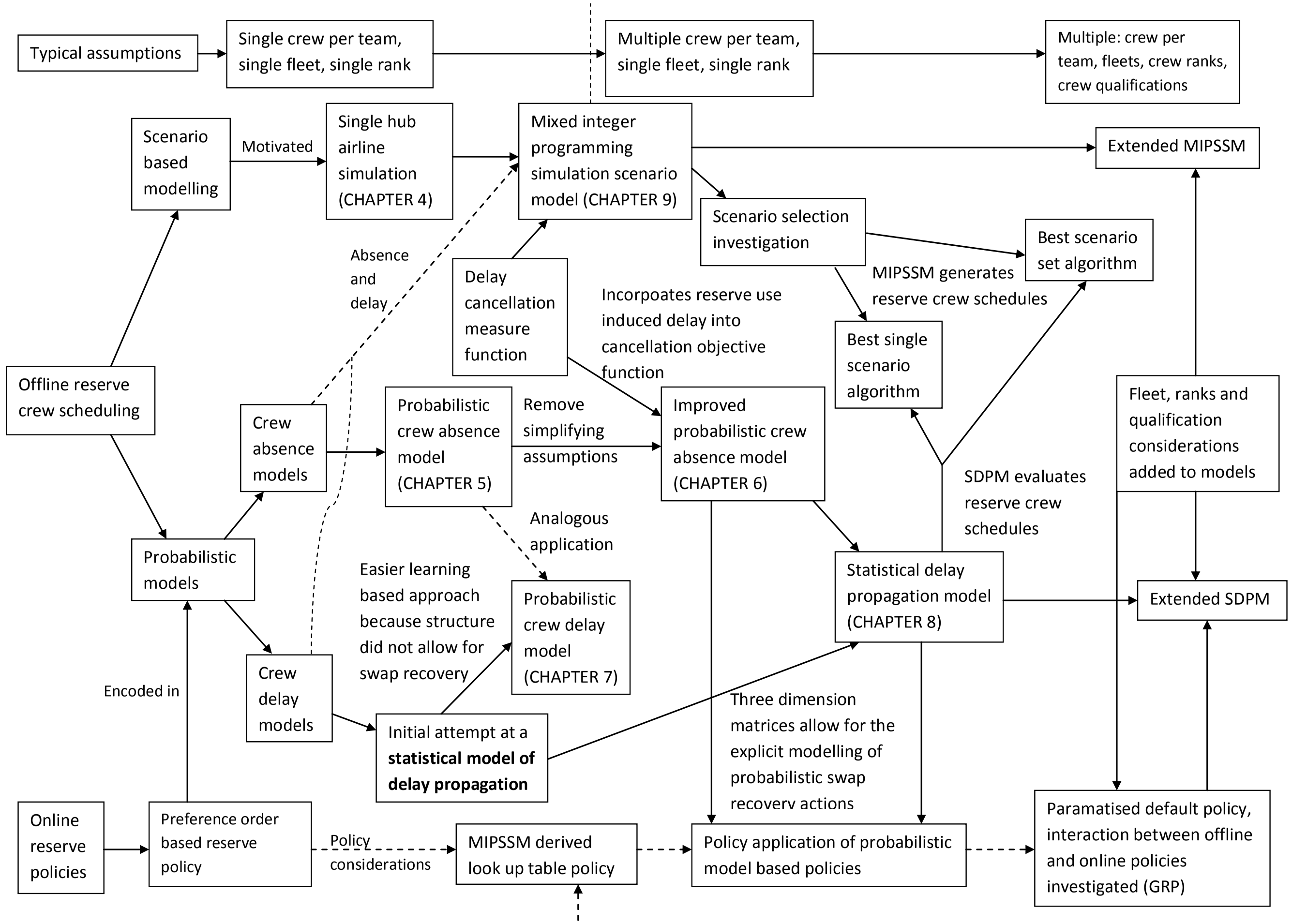

Figure 1.1: Flow of ideas

the improved probabilistic crew absence model to include reserve-induced delays in the cancellation based objective function. The improved probabilistic crew absence model was an entirely theoretical model of crew absence disruptions and reserve crew used to absorb them, this was not the case for the probabilistic crew delay model. This motivated a fresh attempt at a statistical model of delay propagation, an extended three dimensional matrix structure allowed for the explicit modelling of probabilistic swap recovery actions. The statistical delay propagation model ($SDPM$) was a fully theoretical model of crew absence and delay disruptions in contrast to the simulation based $MIPSSM$. Both the $MIPSSM$ and $SDPM$ led naturally to online reserve policy applications. The $MIPSSM$ and $SDPM$ were also combined to produce reserve crew scheduling algorithms that exploited the strengths of both approaches. The $MIPSSM$ excels at generating reserve crew schedules given a set of disruption scenarios, whilst the $SDPM$ is a fast and accurate evaluator of reserve crew schedules.

The initial justification for assuming a single aircraft fleet and a single crew rank was that the problem is largely decomposable according to these divisions in airline resources, and that after developing single fleet/rank models the full problem could then be tackled without major modifications to the models. If an airline uses pure fleet pairings (see Section 2.1.5) crew scheduling decomposes according to fleet types (as in Romer and Mellouli [84]). This was found to be largely the case and the $MIPSSM$ and $SDPM$ approaches were extended, with only relatively minor modifications, to the case where multiple aircraft fleets and multiple crew ranks and qualifications were considered in the same problem. A consequence of the consideration of multiple crew ranks and qualifications was that the default reserve policy needed to be parameterised, as it generated many possible combinations of reserve crew with different combinations of ranks and qualifications all feasible for the same disruption. This parameterised policy was encoded within the improved probabilistic crew absence model, and as a result, in the $SDPM$ as well.

1.7 Chapter summary

This chapter has set the scene for a thesis on the topic of reserve crew scheduling under uncertainty. The goals of this thesis include the development of models of the uncertainty of the occurrence of crew-related disruptions and reserve crew used to absorb them, such models can be used for reserve crew scheduling. The main contributions of this thesis were listed, of those the development of several probabilistic models and a scenario-based approach represent the main contributions. Several journal papers have been submitted and a number of conference papers have been published during this research. This chapter also outlined the flow of ideas throughout this research using a flow diagram.

Chapter 2

Literature review

Airline reserve crew scheduling is a critical task that can, in a number of ways, be viewed as a problem that is centrally located within the research field of domestic aviation. Firstly, the crew schedule is the middle layer of an airline's schedule, surrounded by aircraft routings and passenger itineraries. Secondly, although reserve crew scheduling is a scheduling task, how the reserve crew are used on the day of operation is a task for the airline operations recovery department (or desk). So the process of scheduling and using reserve crew spans both airline scheduling and airline operations.

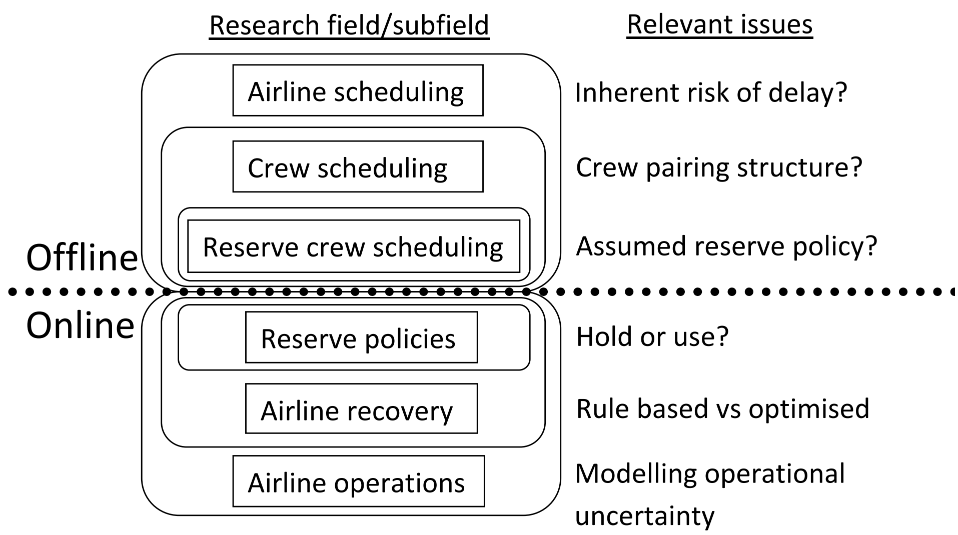

Figure 2.1: Reserve crew scheduling centred view of airline scheduling and operations

Figure 2.1 illustrates how reserve crew scheduling and reserve policies are scheduling and operational tasks respectively and how the overarching research areas of airline scheduling and airline operations relate to reserve crew scheduling via respective levels of subfields of those research areas. The comments/questions to the right hint at the relevant issues with respect to the problem of airline reserve crew scheduling.

Crew schedules are also the most constrained aspect of the airlines schedule, in the sense that human resources require long periods of rest in comparison to the mechanical resources such as aircraft. So crew have to be protected from fatigue for safety reasons and many rules and regulations exist for this purpose (see Banks et al. [9] for a survey of these). Reserve crew provide a layer of slack which protects the feasibility of the crew schedule, and since there are dependencies between the various layers of the airline schedule, reserve crew also have the effect of increasing the robustness of the overall airline schedule.

Chapter structure

As depicted in Figure 2.1 a study of reserve crew scheduling requires knowledge of both airline scheduling (see Section 2.1) and airline operations (see Section 2.4), with a particular emphasis on crew scheduling considerations and airline disruption management. Section 2.1.5 is devoted to previous work on airline crew scheduling. Since reserve crew scheduling is a method of augmenting the robustness of a crew schedule, Section 2.2 considers existing work on robust airline scheduling. Section 2.3 covers previous work on the problem of reserve crew scheduling. Section 2.4 describes the existing work on airline operations, including the crew recovery problem. An important element of this project is the modelling of uncertainty and as such Section 2.5 is devoted to this topic. Section 2.6 concerns solution methodologies for deterministic problems. Section 2.7 concerns solution methodologies for problems with uncertain input parameters. Section 2.8 considers general aspects of problem solving. Section 2.9 summarises this chapter.

In addition to providing a review of the existing literature, this chapter also discusses the existing literature in comparison to the approaches that are proposed in the subsequent chapters of this thesis.

2.1 Airline scheduling

The airline scheduling problem is a very large complex problem and because of this the problem is usually tackled sequentially in several steps, as described by Barnhart et al. [10]. The overall sequential structure is summarised by Bazargan [18] as: the fleet assignment requires the schedule design, maintenance/aircraft routing requires fleet assignment to generate aircraft rotations, crew pairing requires aircraft routing (as its input) and crew rostering requires (the result from) crew pairing. see Table 2.1 for the definitions of these airline scheduling terms.

The crew scheduling problem is further broken down into crew pairing and crew assignment [10, 12, 55, 58, 89, 98]. Crew pairing is the task of generating sequences of flight duties consisting of several days of work that can be assigned to an individual member of crew. Crew pairings begin and end at crew domiciles (home airports). Additionally, other aspects of the sequential approach to the airline scheduling problem are also further decomposed into smaller more manageable sub-problems. Note also

Term	Definition
Schedule design	determines which flight legs should be flown, schedule design can be seen as a marketing task
Flight leg	a single journey from one station to another
Fleet assignment	assigns aircraft types (fleets) to flight legs.
Maintenance routing	assigns specific aircraft to flight legs, such that all aircraft receive sufficient maintenance
Crew duty	a sequence of flights to be performed by a crew member in one day
Domicile	the home station of a crew member
Crew pairing	generates sequences of crew duties that start and end at crew domiciles
Crew assignment	assigns specific crew to crew pairings

Table 2.1: Airline scheduling terminology

that there is usually a negotiation period in which requests for changes in a higher level schedule can be made [75]. Many of the overview papers spend time emphasising the value of putting time and effort into the airline scheduling problem, Yen and Birge [106] call it the billion dollar problem, others including Gopalan and Talluri [44] and Rushmeier et al. [89] state that "the schedule is the primary product of an airline and is the single determining factor affecting airline profitability". Rushmeier et al. [89] also state that the airline industry has the largest scheduling problems of all industries. Another interesting way of summarising airline scheduling, which was given by Ball et al. [7] is to consider the schedule as consisting of 3 layers: aircraft, crew and passengers. A feasible schedule must satisfy the flow constraints for each layer. Additionally, Ball et al. [7] provide a detailed overview of the intricacies of the commercial air transportation network from an air carrier's, an airport's and air traffic control perspectives. The complexity of the air transportation network arises from the competing interests of airlines with respect to finite airport and airspace capacity, and the sensitivity to disruptions caused by the interaction between the different schedule layers.

2.1.1 Schedule design

In schedule design, airlines determine the set of origins and destinations and the corresponding departure and arrival times of the flights they will operate. The airline tries to generate the most profitable flight legs possible. According to the thesis of De Maere [67] schedule design consists of three subproblems, route development, frequency planning and timetable development.

Airline network types

Different airlines use different types of networks. The two basic network types are point-to-point and hub-and-spoke networks [66]. Airlines that operate hub-and-spoke networks have one or multiple stations (hub airports) through which all traffic passes between spoke stations. For the case of multiple hub-and-spoke networks there may also be high frequency flights between the hub stations. The advantage of a hub-and-spoke network is that the airline can offer passengers many origin-destination pairs with relatively few flights, however such trips often involve a connection at one of the hub stations. In contrast, airlines that operate a point-to-point network have no hub station. Instead, they offer direct flights between a range of stations. Point-to-point networks are often operated by "budget airlines" or "low cost carriers" and offer flights on high demand routes. Due to the large traffic volumes at hub stations they have the largest reserve crew scheduling problems. The reserve crew scheduled at a hub station are the main focus of this thesis.

Li et al. [62] give an analysis of a Japanese airline that operate a dual hub network. They define a connection quality measure for the connections the airline offers to its customers. The measure takes total flight time, the number of connections and the existence of direct flights that may be offered by competitor airlines into account. It was found that some of the connections the Japanese airline operated were of low quality due to other airlines offering direct flights for the same origin-destination pairs. Their analysis led them to conclude that there was a case for increasing the number of international flights from one of the airports of the airline's dual hub network.

The focus in this thesis is on a single hub network, the reserve crew being scheduled are those based at the hub station. A multiple hub application of the proposed approaches would involve solving a set of reserve crew scheduling subproblems, one for each hub. However, the latter is beyond the scope of this thesis, but is discussed as a possible direction for future work in Chapter 12.

2.1.2 Fleet assignment

The goal of fleet assignment is to match aircraft capacity with demand in order to avoid passenger spills (where demand exceeds the number of seats) or under utilisation of capacity. Schedule design and fleet composition are the input for fleet assignment. The fleet assignment problem is often modelled as a network flow problem [44]. In such network flow models, nodes represent connection possibilities at particular stations and arcs represent flight legs and also overnight legs. The advantage of this approach is that equipment balance constraints are then easy to enforce. A current issue in fleet assignment (raised by Barnhart in [10]) relates to the ordering of the sequential airline scheduling problem. The problem is that sequentially solving the fleet assignment followed by the routing problem, can result in maintenance requirement violations. Integrated fleeting and routing models avoid this. Gopalan and Talluri [44] describe a periodic approach that assumes the same schedule every day and the problem is to find aircraft rotations such that maintenance requirements are satisfied. In [44] Gopalan and Talluri also state that in the fleet assignment problem "throughs" are highly profitable and the number of these is often maximised. A "through" corresponds to an aircraft stopping at the hub station as an intermediate destination, allowing passengers to stay on the same aircraft in a trip from one spoke station to another spoke station.

The relevance of fleet assignment from a reserve crew scheduling perspective is slightly indirect. The fleets assigned to each flight leg determine the necessary qualifications of the crew that will be assigned to those flight legs. Reserve crew also have qualifications which determine the fleets they are qualified to operate on. Additionally, the fleet assigned to a flight leg will influence the probability that flight is delayed, which is because different fleets have different airspeeds at which overall fuel consumption is minimised [60]. Therefore fleet assignment influences reserve crew feasibility and the occurrence of delays for which reserve crew may be required to replace delayed crew.

2.1.3 Maintenance routing

For completeness, the maintenance routing phase of airline scheduling is described. For safety reasons aircraft must undergo regular maintenance checks. The various types of maintenance checks include high frequency visual checks and lower frequency more rigorous checks. Gopalan and Talluri [44] give details of the types of maintenance checks required by Federal Aviation Administration regulations. Checks are labelled A,B,C and D. A checks are high frequency routine visual inspections. If A,B,C or D checks are not performed on time, the aircraft is prohibited from flying. The B, C and D checks are less frequent and more rigorous, in these checks aircraft can be taken out of service for a week or more and may be completely dismantled. Maintenance carried out by airlines is usually more frequent than those specified by civil aviation authorities. This adds a little flexibility which means that missing a single maintenance check need not result in the aircraft immediately being prohibited from flying. For an example of work on maintenance routing see [60]. Lapp and Wilkenhauser [60] introduce a tail assignment model which is designed to minimise the total amount of fuel required to implement a schedule. They use aircraft efficiency ratings and assign lines of flight to specific aircraft with the objective of minimising fuel consumption. They also introduce a model where new lines of flight can be generated, however they found that this leads to over utilisation of efficient airframes and also has the knock-on effect of increasing crew costs and increased delay propagation.

Planning rules: Equal utilisation of airframes

Airlines use planning rules that are known to incorporate beneficial features into aircraft routings including the equal utilisation of airframes [89]. This simplifies the task of maintenance routing because flying hours will be more evenly balanced across the airframes within each fleet. Another planning rule stated by Rushmeier et al. [89] which is used by US Airways is that all aircraft must pass through an overnight station every three days so that maintenance checks can be performed overnight.

2.1.4 Integrated airline scheduling

Section 2.1 stated that airline scheduling is performed in a sequential manner, it is common wisdom that sequential approaches to problems do often preclude truly optimal solutions. The reason for breaking the airline scheduling problem into stages is its combinatorial nature, focussing on a small part of the problem at a time makes the problem more manageable. Unfortunately this can also have the effect of removing regions of the solution space, that are available in the following subproblem, possibly including the optimal solution. A recent trend in the research literature for airline scheduling is the integration of two or more subproblems of the airline scheduling problem.

Weide [101] integrates fleet assignment and crew pairing in an algorithm which iteratively increases the robustness of an input airline schedule. The work of Weide is discussed in more depth in Section 2.2.3 in the context of robust scheduling.

Pita et al. [78] integrate schedule design and fleet assignment under airport congestion. Their mixed integer programming model takes airline competition and cooperation into account using the frequency of competitor's flights to estimate the market share gained by attaining particular airport gate slots to serve demand for different origin destination pairs. They model demand for origin destination pairs according to early morning, late morning and early afternoon time periods. The authors validate their approach using data from a real national air transportation network.

In [45] Grunert presents an approach for tackling an integrated network design and fleet assignment as discussed in Section 2.1.2. The work of Grunert [45] is discussed in the context of a hybridised solution approach involving integer programming and a tabu search algorithm.

In [11] Barnhart et al. introduce a flight string model, in which flight strings are defined as sequences of flights, that can feasibly by assigned to an aircraft, which begin and end at maintenance stations. As a result, any aircraft assigned to such a flight string is therefore maintenance feasible. The flight string model of Barnhart et al. [11] allows fleet assignment and maintenance routing to be solved in a single problem, thus integrating fleet assignment and maintenance routing problems. The trend of integrating airline scheduling sub-problems is mirrored in the airline recovery problem. Peterson [76] addresses the full integrated airline recovery problem from schedule recovery through to passenger re-routing. The work of Peterson is discussed in more depth in Section 2.4.3 in the context of airline recovery.

2.1.5 Crew scheduling

As described in Section 2.1 the crew scheduling process is typically decomposed into two sub-problems: crew pairing and crew assignment [12]. Additionally, crew pairing and crew assignment each have their own unique types of constraints. The crew pairing problem considers work rules, whilst the crew assignment problem considers individual crew needs. The work of Rosenberger et al. [87] contains a detailed account of the different types of crew constraints that have to be taken into account on the day of operation.

Crew pairing

Crew pairing finds minimum cost generic (anonymous) strings of duties that begin and end at crew bases. In this problem possible pairings are assessed in terms of the costs of layovers (overnight stays at outstations), time away from base and other incremental costs. Pairings are made of duty periods which are essentially shifts or a day's work. Duties must begin from where a previous duty ended, unless dead-heading (crew transported as passengers) is used. The crew pairing problem is important from a reserve crew scheduling perspective because it determines how damaging uncovered crew-related delays can be, both in terms of cancellations due to crew absence and how uncovered delays will propagate through the schedule. Barnhart et al. [12] present the crew pairing problem in the form of a set partitioning problem which is given as follows:

$$Minimise : \sum_{p \in P} c_p y_p \tag{2.1}$$

$$\sum_{p:i \in p} y_p = 1 \ \forall i \in F \tag{2.2}$$

$$y_p \in \{0, 1\} \ \forall p \in P \tag{2.3}$$

In this model P is a matrix of all feasible pairings, each column of P (p) corresponds to a crew pairing. Each column states which flights are contained in that pairing with a 1 in row i to indicate inclusion of flight i in that pairing and 0 otherwise. F is the set of all flights to be covered. The objective function (Equation 2.1) is to minimise the cost of a set of pairings that cover all flights exactly once. Constraint 2.2 ensures that all flights are covered by exactly one pairing. Constraint 2.3 asserts that y_p is a binary decision variable. Barnhart et al. [12] acknowledge that this approach requires all feasible pairings to be enumerated and therefore is not a tractable approach for anything but the smallest of crew pairing problems. Barnhart et al. [12] also demonstrate, with a numerical example, how a small instance of a crew pairing problem can be solved with this model. Typically, methods such as column generation are used to solve the problem in which only a subset of feasible pairings are considered at any one time.

Beasley and Cao [19] introduce a tree search algorithm for the generic crew scheduling problem with N tasks and K crew with constraints for duty length and the temporal and spatial feasibility of the overall schedule. They use a branch and bound approach involving Lagrangian relaxation and subgradient optimisation to provide lower bound estimates of the value of solutions, which are used to guide the algorithm.

Ball and Roberts [8] introduce a graph partitioning approach to the airline crew scheduling problem. In their model nodes correspond to flights and paths visiting a sequence of nodes correspond to crew pairings. Their algorithm proceeds by extending a set of pairings with an extra flight at each iteration, followed by a pairing improvement phase and a feasibility check for accepting the set of pairings that are considered in the next iteration. Their approach takes deadheading into account as well as multiple domiciles.

Klabjan et al. [56] tackle the problem of airline crew scheduling with regularity. They define flight legs according to a measure of regularity, which is defined as the number of days on which the same flight is repeated in the pairing. Any given crew pairing belongs to a regularity group that corresponds to the number of days the same set of flights is repeated. Their objective is to minimise a weighted sum of the number of pairings that are irregular and the total number of pairings generated to cover all flights. Their computational results demonstrate that their approach outperforms the methods used by most large airlines on the criteria of deadheads, flight time credit (see Section 2.1.5 for a detailed definition of flight time credit) and regularity.

Anbil et al. [5] give an account of the success of a commercial airline crew pairing optimiser known as TRIP (1991). They highlight the financial importance of crew scheduling at what was Americas biggest airline, with annual crew costs exceeding $1.3 billion, a cost second only to fuel costs. The TRIP optimiser was the result of improvements in computer architecture and interior point methods for solving integer programming problems.

Romer and Mellouli [84] consider a crew pairing problem where explicit consideration is given to the possibility of cabin crew of high rank being assigned to roles for which lower ranked cabin crew are also feasible. They state that crew pairing optimisation is usually decomposed according to crew-roles/ranks and fleet types. They show that when crew can be assigned to roles which are below their rank, better results can be obtained by solving an integrated crew pairing model of hierarchical crew. The work of Romer and Mellouli tackles cabin crew scheduling, whilst this thesis is primarily aimed at the scheduling of reserve cabin crew. The ranks and qualifications of reserve crew determines which roles they can undertake and which fleets they can operate on in the event of crew related disruptions. This thesis considers the possibility of "flying below rank" (as well as crew fleet qualifications) at the end of Chapters 6 and 9. Romer and Mellouli do not consider fleet qualifications and instead focus on the case where the problem decomposes according to fleet types. This is reasonable from a crew scheduling perspective as crew pairings usually consist of flights on a single fleet type (or pure fleet pairings). However, reserve crew can be used to absorb crew related disruptions affecting any fleet they are qualified for.

The typical constraints in the crew pairing problem are described by Barnhart et al. [12] and include minimum and maximum connection times and maximum flight time (per day, week, month and year). Banks et al. [9] list the rule types relating to flight attendant fatigue, which includes maximum flight time, minimum rest times and jet lag (circadian rhythms). They also point out that very few rules exist concerning at what time the rest stopwatch is started, e.g. is it when the duty finishes or when the crew reach the rest facility. In general, airlines have bargaining agreements with unions that govern the maximum flying time per month and the minimum and maximum rest times used during in the planning phase. Collective bargaining agreement rules are usually set well within the legal limits set by civil aviation authorities, which means that there is some rule flexibility on the day of operation.

At KLM [57] each possible crew pairing has a corresponding required minimum numbers of days off. For example, a long haul pairing from Amsterdam to Manilla requires 9 days off on its completion. Section 3.4 describes KLM crew scheduling constraints in more detail, including: fleet crewing requirements; the rank structure for cabin crew; and how on the day of operation some exceptional circumstances permit “flying minus one” or “flying above rank”. In this thesis, fleet crew requirements are taken into account in Chapters 6 and 9 when the probabilistic models and mixed integer programming approaches to reserve crew scheduling are extended from single fleet models to multiple fleet models.

In this thesis, the most important aspect of crew scheduling is the structure of crew pairings. The structure of crew pairings, determines which reserve crew will be feasible to cover disrupted crew affecting that crew pairing and also determines which subsequent flights will be disrupted if a crew-related disruption cannot be covered for. As a result of this a possible area for future research is to integrate crew scheduling and reserve crew scheduling, the crew schedule can be manipulated to make reserve crew scheduling easier or the crew recovery problem easier. In Chapter 10 the test instances are based on real aircraft routings and crew schedules which are generated using a set partitioning formulation similar to that described above.

Crew assignment

Crew assignment follows crew pairing. In crew assignment pairings are assigned to individual crew. Current literature [98, 89, 58] also suggests that different airlines perform crew assignment in different ways, including bidline, preferential bidding and rostering. Rushmeier et al.[89] describe bidline rostering as an approach where anonymous schedules are created which are then bid for by crew. In the bidline approach, crew who are not assigned to their preferred schedules are awarded points, which give them prioritisation in the subsequent month’s schedule. Other airlines perform crew assignment using an automated approach in which all crew schedules have to satisfy quality of life constraints (such as regularity). This is the case for rostering and preferential bidding, which generate personalised schedules. Rostering takes vacation and training into account, whilst preferential bidding meets the requirements of more senior crew first.

Kohl and Karish [58] address the crew rostering problem, giving details of all of the different types of constraints that occur in this problem. The authors use an objective function in their crew rostering model which includes cost and quality of life considerations. They use constraints that ensure that each individual crew member does not violate duty time regulations, and that the qualification requirements for specific pairings are satisfied by the assigned crew.

Pay structures

The work of Rosenberger et al. [87] goes into the details of how the costs (or flight time credit) of crew schedules are typically determined at airlines based in the United States. The cost of duties within a particular crew pairing are proportional to the maximum of three quantities. These include: the fraction of the elapsed duty time; the minimum guaranteed time; and the sum of block times. Pairing costs are proportional to the maximum of three quantities: the time away from base; the minimum guaranteed pay per duty multiplied by the number of duties in the pairing; and the sum of the duty costs as determined from the duty cost structure. The aforementioned pay structure does not apply to KLM [57] who have a fixed salary pay structure, in which staff have target numbers of flying hours to be worked each month. KLM in particular have staff with different percentage contracts. Pay at KLM increases with rank.

2.2 Robust airline scheduling

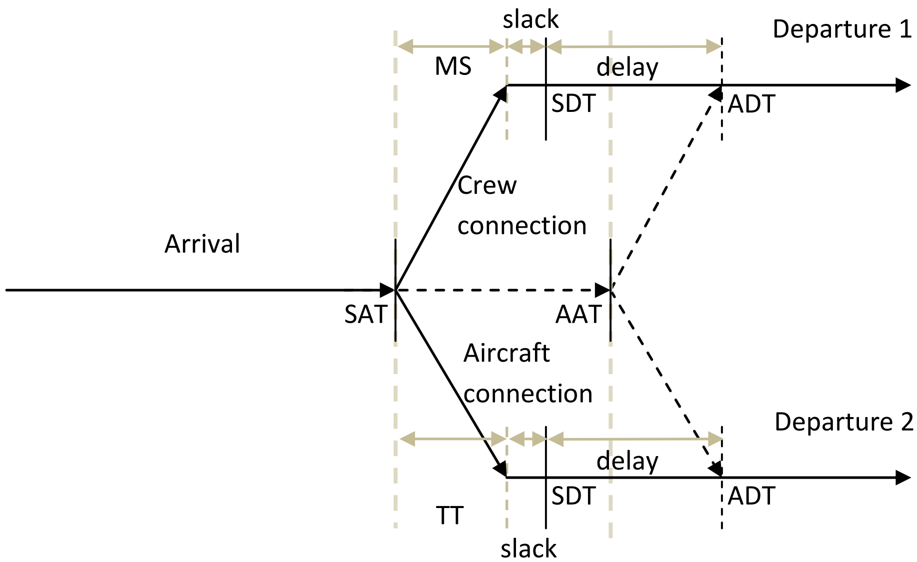

Figure 2.2: Delays can spread to multiple flights if "resources split" causing a "switch delay" (dashed lines). This occurs when the the crew and aircraft of a delayed arrival are assigned to different subsequent flights.

Reserve crew scheduling can be viewed as a method of increasing the robustness of an airline schedule. In the literature, there does not seem to be a single definition of robustness, but most agree that robust schedules are designed to be less sensitive to disruptions. Another common theme in the literature is that robustness is usually something that can be increased in the planning phase by considering what happens to a schedule when disruptions occur. An airline schedule can be made more robust by making it delay resistant. Delay resistance can be achieved by: adding slack (extra time between consecutive flights) in to the schedule; strategically scheduling resources to ensure that delays do not propagate to multiple flights by avoiding resource splitting (see Figure 2.2); or by scheduling spare airline resources (reserve crew or spare aircraft) that can by used to replace disrupted airline resources. Bertsimas et al. [22] provides a general introduction to robust optimisation and that of Gabreal et al. [41] for a general overview of work on robust optimisation. The following gives a survey of some of the existing approaches to robust airline scheduling.

2.2.1 Robust fleet assignment

Smith [96] tackles the robust fleet assignment problem using the concept of station purity to maximise the number of possible aircraft swaps. Station purity here refers to scheduling aircraft to limit the number of fleet types at each station, so that aircraft of the same fleet type have common ground time at each airport. The benefits from station purity are two fold. Firstly, maximising the common ground time of aircraft from the same fleet, also maximises the opportunity for crew swaps, because the crew will be qualified for each other's fleet. Secondly, the increased availability of possible aircraft swaps brought about by station purity, increases the possibility that aircraft can be swapped after the publication of the schedule to match aircraft capacity with the latest demand figures, but only if the crew assigned to these flights are swappable. Smith [96] extends the typical fleet assignment model (FAM) to a station decomposition model (SDM). The approach is aimed at hub-and-spoke networks where flights are either to or from a hub station. Smith's SDM involves iteratively solving a master problem for the hub, and sub-problems for spoke groups. The aim is to maximise the station purity of the spoke station groups. In this thesis, aircraft swap and crew swap recovery actions are considered. In general it is assumed that a crew swap recovery action is cheaper and therefore preferable to using a team of reserve crew to replace delayed crew. So when more crew swaps are available the expected demand for reserve crew at that time is reduced. The general point here is that different approaches to increasing schedule robustness can be complementary as opposed to conflicting.

Other research on robust fleet assignment includes that of Rosenberger et al. [86] which introduces a flight string (see Section 2.1.4) based fleet assignment model which minimises the number of different stations in strings assigned to fleets as to maximise the number of short cancellation cycles. The purpose of this is to reduce the number of flights that have to be cancelled (cancellation cycle), if one flight in the string is cancelled. In short cancellation cycles, the flight string can be resumed after a smaller number of cancellations. Rosenberger et al. refer to this as minimising the hub-connectivity of flight strings. In a single hub-and-spoke network hub cancellation cycles typically involve two flights, the outward and return legs.

2.2.2 Robust crew scheduling

Yen and Birge [106] describe an iterative approach to increasing crew schedule robustness. In each iteration the traditional crew pairing model is solved, followed by a recourse problem that allows or disallows crew connections that may lead to delay in the next iteration. Their model aims to minimise "switch delays". Switch delays (see Figure 2.2) are delays that spread to 2 flights in a single connection, because the crew and aircraft connect to and delay two separate flights. In the context of Yen and Birge's work [106], pairings that involve aircraft changes hold more potential for delay and are eliminated and not branched upon in their "flight pair branching algorithm". The recourse problem which calculates "switch delays" does so by considering a set of random scenarios and solving a set of decision variables to compute how delay propagates through the schedule given the most recent solution to the crew pairing problem. Their experimental results indicate that their approach leads to solutions with crew schedules involving fewer aircraft changes and as a result reduces switch delays. This increased robustness of the crew schedule comes at the cost of an increase in the crew pairing cost compared to the initial cost optimal input solution to the crew pairing problem.

Schaefer et al. [91] use a simulation package called SimAir to estimate operational costs of crew pairings, which are then used when solving the crew pairing problem. They make several simplifying assumptions including: no flight cancellations; the only method of recovery is pushback (delaying flights); and the operational costs of pairings can be accurately estimated in isolation from the effects of other pairings. An alternative approach is described called the penalty method. This approach does not use simulation to derive operational pairing costs, instead the features of pairings such as the rest periods between flights, the difference between the planned pairing finish time and the legal limitations on maximum work hours are used to penalise pairings which are likely to perform badly in operations. They propose a local search method to identify the most representative weights for penalising the various features of pairings. Their experiment results indicate that crew schedules based on operational costs perform better than those based on planned costs.

Shebalov and Klabjan [94] investigate the concept of move-up crews, this is where crew pairing is performed in such a way that creates crew swap opportunities by encouraging common ground time between swappable crew. They found that increasing the number of move-up crews towards the end of pairings is beneficial as this is when most delay disruptions occur. The move-up crew formulation is solved using a column generation strategy. Their experimental results show that their approach minimises operating costs for days involving 1 or more disruptions due to the reduced recovery cost caused by increasing the availability of crew swaps. But when no disruptions occur operational costs are increased by their approach.

Robust railway crew scheduling

Nielson [73] investigates the effect of planning rules on the (delay) absorption capacity of railway schedules, the planning rules investigated each have analogies with airline crew scheduling. They include: slack between consecutive journeys; scheduling crew as teams (drivers and conductors stay together); scheduling crew to stay on the same rolling stock for as many consecutive journeys as possible; reserve crew; crew swap opportunities; and scheduling crews to limited sets of lines (to limit the propagation of crew related delays to different lines). A simulation is used to test the schedules that are built using various degrees of the above mentioned planning rules. Disruption scenarios are generated and total delay propagation is calculated for each simulation. In general their experiments show that the planning rules do increase the absorption capacity of schedules but the increase of schedule costs means that a trade-off exists.

Robust crew scheduling summary

Yen and Birge's work [106] and that of Nielson [73] both evaluate the robustness of solutions by generating random disruptions scenarios. Chapter 9 introduces a scenario-based mixed integer programming approach to reserve crew scheduling. In contrast to the models of Yen and Birge [106] and Nielson [73] which use disruption scenarios to evaluate previously generated solutions, the approach of Chapter 9 uses disruption scenarios to find a robust solution specifically designed for a given set of input scenarios. In summary, the literature suggests that robust crew schedules can be achieved in the strategic planning phase by considering the effects of disruptions from the points of view of crew scheduling, planned departure times, recovery and planning rules. The robustness of a crew schedule can be increased by minimising the dependencies between different crew pairings and aircraft routings and by increasing the availability of swap recovery actions. Both of these approaches reduce the demand for reserve crew, therefore the existing robustness of a crew schedule should be taken into account during reserve crew scheduling. In this thesis explicit consideration is given to the structure of the airline's schedule and the effect that this has on reserve demand (see the seventh bullet point of Section 3.2).

2.2.3 Other approaches to increasing airline schedule robustness

Sohoni [97] et al. introduce stochastic programming models for modifying airline schedule departure times within allowable time windows, with the aim of increasing on-time performance and minimising the probability of passengers missing connections. Metrics for on-time performance and passenger itinerary completion are introduced and are used in the objective functions of their stochastic programs. The on-time performance metric is defined as the probability that a particular departure is delayed by more than 15 minutes. The passenger itinerary completion metric is defined as the probability that a passenger on a flight will be able to connect to any of the flights in the connection set for the given flight. They model block times (gate to gate times) with log Laplace distributions. Two stochastic programs are introduced. The first maximises profit with constraints for the minimum performance levels for on-time and passenger itinerary completion. The second model maximises the performance level with respect to constraints on the maximum cost. The models also feature an explicit set of passenger itineraries with associated demands which are used to calculate the performance measures. Cost calculations are based on the effect that the modified schedule has on crew and aircraft usage. In their model, the use of block time distributions leads to non-linear terms in the model which are incorporated into the stochastic programs using piecewise linear approximations. To solve the resultant models, cut generation (or constraint generation) algorithms are developed. Their experimental results show that increasing performance and service levels increases costs.

Weide et al. [102] introduce an iterative approach to solving the integrated crew pairing and aircraft routing problems with the objective of minimising the overall cost and minimising what are termed restricted aircraft changes. A connection is an arc between the arrival of one flight and the departure of another, if crew and aircraft are assigned to the same connection this implies that they stay together. Restricted connections have a crew sit time which is greater than or equal to the minimum sit time but less than the restricted time (for that connection). When crew change aircraft restricted connections significantly increase the risk of switch delays (see Figure 2.2). Weide et al. state that it is desirable to have crew and aircraft stay together and to avoid crew changing aircraft as much as possible. Firstly, the minimum sit time for crew is smallest when the crew stay on the same aircraft, and secondly, aircraft changes give rise to the possibility of switch delays. Therefore, minimising restricted aircraft changes has the effect of increasing a schedule's robustness. Weide et al. use the objective of minimising the cost of the crew pairing solution plus penalties for assigning crew to restricted connections that aircraft are not scheduled to, whilst assuming the cost of the aircraft routing is fixed. The solution approach employed by Weide et al. starts from the cost optimal crew pairing and aircraft routing solutions. Then, the aircraft routing and the crew pairing problems are solved alternately in an iterative algorithm. In each iteration the aircraft routing problem is solved using the objective of minimising the penalty due to crew assigned to restricted connections for which aircraft are not assigned, this step encourages aircraft to follow crew. The crew pairing problem is then solved to cost optimality given the most recent aircraft routing, with a weighted term for a measure of non-robustness. In this fashion Weide et al. start with a cost optimal solution and iteratively improve the robustness of the crew and aircraft schedules. The experimental results reported by Weide et al. [102] for the iterative airline scheduling algorithm show that increasing schedule robustness increases the schedule's cost.

The master thesis of Ageeva [2] introduces an approach to incorporating robustness in an airline schedule using the concept of overlaps. An overlap is defined as an opportunity for two aircraft to swap schedules where there is also an opportunity for the swap to be undone at a later time before the next scheduled maintenance checks for each aircraft. Ageeva tries to formulate the aircraft routing problem to include overlap maximisation alongside cost minimisation, but encounters non-linearities that lead to an alternative approach based on generating alternative solutions and then externally evaluating the robustness of the solutions. The mathematical framework used defines points as ground time intervals at a specified airport, strings consist of a sequence of points and overlaps exist if strings have two common points which allow a swap to be performed and to be undone later. Ageeva's experimental results indicate that the proposed approach increased aircraft routing robustness by as much as 35%.

The work of Dunbar et al. [35] proposes an integrated approach to aircraft routing and crew pairing that minimises delay propagation. Their approach estimates the expected delay for crew (aircraft) flight sequences whilst taking into account the aircraft (crew) sequences that are assigned to the same flights. Their experimental results show that their integrated approach outperforms the sequential approach to solving aircraft routing and then crew pairing.

The thesis of Lan [92] focusses on increasing the robustness of an airline schedule by minimising the potential for delay propagation and passenger disruptions. Lan formulates two models. The first is an aircraft routing model which minimises delay propagation. The second allows small adjustments to departure times with the goal of minimising passenger disruptions. The resultant models are solved with branch and price [14]. The experimental results indicate that reduced delay propagation and passenger disruptions can be achieved with the approach.

The thesis of De Maere [67] investigates multiple objective approaches to robust airline scheduling. The author considers retiming flights within allowable time windows and changes to the aircraft maintenance routing and their effect on schedule robustness. A simulation study is carried out to validate the trade-off solutions yielded from the proposed approaches.

Duck et al. [33] introduce an integrated approach to the crew pairing and aircraft routing problems. Their objective includes cost terms from the usual crew pairing and aircraft routing formulations plus an expected propagated delay term. The problem of determining the delay propagation associated with a solution being a function of the solution itself is circumvented by decomposing the problem into separate crew and aircraft routing problems each with their own recourse problem. Delays for a given schedule are calculated by considering a stochastically generated set of scenarios, where each scenario specified realised departure and arrival event times. An iterative approach (based on the iterative algorithm of Weide et al. [102]) is used in which crew pairing and aircraft routing subproblems are solved in each iteration. They use a branch and price approach where the solutions of the recourse problems are used to guide the generation of new columns. Their results show that expected reactionary delay can be reduced by up to 6.4% without increasing the crew cost.

In this section several different approaches to increasing the robustness of airline schedules have been discussed. Sohoni et al. [97], Lan [92] and De Maere [67] modified scheduled departure times to minimise the potential for delay propagation. Weide et al. [102], Dunbar et al. [35] and Duck et al. [33] took the approach of performing crew pairing and aircraft routing in an integrated fashion, which helped to reduce disruptions resulting from dependencies between the crew and the aircraft schedules. Ageeva's [2] approach was to increase the availability of swap recovery actions by encouraging ground time overlaps. This is a similar approach to that used by Smith [96] (see Section 2.2.1) for aircraft swap opportunities (station purity) and Shebalov and Klabjan [94] (see Section 2.2.2) for crew swap opportunities (move-up crews). In general increasing schedule robustness increases the planned cost of a schedule, the point is that in the event of disruptions the robust schedule has a lower recovery cost. This thesis takes the approach of augmenting the robustness of an airline's schedule through the scheduling of reserve crew.

2.3 Reserve crew scheduling

<table>
<tr><th rowspan="2">Phase of planning/ operations</th><th colspan="3">Reserve demand type</th></tr>
<tr><th>Absenteeism</th><th>Delayed crew</th><th>Open time</th></tr>
<tr><td>Manpower planning</td><td>Boissy, 2006</td><td>Gaballa, 1979</td><td></td></tr>
<tr><td>Reserve pairing generation</td><td></td><td></td><td>J.E. Dillon and S Kontogiorgis, 1999</td></tr>
<tr><td rowspan="3">Reserve crew scheduling</td><td colspan="3">Sohoni et al., 1999</td></tr>
<tr><td colspan="2">Paelinck,2001</td><td></td></tr>
<tr><td colspan="2" rowspan="2">Bayliss, 2012, 2013, 2014</td><td></td></tr>
<tr><td>Online reserve policies</td><td>Rosenberger et al., 2002</td></tr>
</table>

Table 2.2: Classification of the existing literature on reserve crew scheduling

Most papers on crew scheduling are aimed directly at crew scheduling and only a few concern reserve crew scheduling. There is a clear overlap between both types of paper and each often contains information relevant to the other. The existing work on reserve crew is summarised in Table 2.3, which defines the phase of planning or operations and the type of reserve demand considered in the specified works.

Sohoni et al. [98] minimise the requirement for reserve cockpit crew (the most expensive crew type) by better predicting the requirement for reserve crew. They report that for the schedule instances they consider, reserve crew utilisation is about 40%. They claim that if reoccurring training is taken into account during crew scheduling the estimated requirement for reserve crew will be more accurate because conflicts with reoccurring training is a leading cause for the requirement of reserve crew. Consequently Sohoni et al. focus on predicting reserve crew demand due to reoccurring training conflicts. They allocate reserve duties using information about training conflicts which give rise to open time (duties which have no crew assigned to them) and also using information derived from an airline operations simulator (SimAir, see Section 2.5.3). Whereas Sohoni et al. focus on reserve crew demand due to schedule conflicts, this thesis focusses on reserve crew demands that occur on the day of operation as a result of unexpected crew absence and delayed crew. In addition, as described in Section 1.1, KLM leave open flights in their schedule which are covered by scheduled reserve crew if they are not used to cover unexpected crew related disruptions or if enough time remains in their reserve block after covering a disruption.

In [74], Paelinck describes a practical approach which was implemented at KLM to improve cabin crew reserve duties. The approach calculates daily demands for reserve crew and the expected number of reserve crew remaining each day, and uses a reserve block stacking approach which aims to have reserve crew available on standby at all times. The work of Paelinck [74] highlights some of the difficulties associated with the planning and scheduling of reserve crew, including how many should be scheduled, what reserve duty start times should be used, when and what is the best way to use reserve crew in response to disruptions. The work of Paelinck [74] provided the starting point for the research presented in this thesis.

Boissy [24] describes a forecast model for absenteeism and a model for minimising the cost of reserve crew and missing crew. Boissy defines tension as the number of disruptions divided by the number of reserves. Using more reserve crew decreases tension but increases the planned crew cost. Boissy's model is used to find the optimal tension, which corresponds to the minimum cost of missing crew plus reserve crew costs. The work of Boissy tackles reserve crew scheduling from a manpower planning perspective. in contrast, the work in this thesis tackles reserve crew scheduling from the point of view of allocating reserve duties to a set of available (already planned manpower) reserve crew.

Gaballa [40] uses the probabilities of callouts as a guide to reserve sizing. Gaballa assumes that reserve crew are used when flights are delayed such that the scheduled crew would exceed their maximum duty length if they start the delayed flight. The main focus is on overnight delays where crew have to be replaced so that their minimum overnight rest constraints are not violated. Gaballa defines a reserve policy as the number of reserve crew scheduled and calculates for each policy the probability of overnight delays and the expected cost of the policy. The author observed that the reserve policy used by Quantas at the time meant that overnight delays due to reserve crew unavailability had a 1 in every 166 years chance of occurring. The approach was implemented and was estimated to save $600,000 (1979) a year. The reserve crew scheduling problem tackled by Gaballa concerned the determination of the number of reserve crew required to minimise the probabilities of overnight delays and reserve crew costs. The allocation of reserve crew duties was based on fixed duty start times, namely early and late duty start times. In this thesis duty start times are the variables and are discretised according to scheduled departure times. Additionally, the work of Gaballa concerns the scheduling of call out reserve crew whereas this thesis is concerned with the scheduling of reserve crew stationed at the hub station, although, callout crew can be modelled as reserve crew stationed at the hub station who have a delayed response time.

Dillon and Kontogiorgis [32] present an approach for pilot reserve crew scheduling that generates reserve pairings which are then allocated using preferential bidding. They focus on quality of life considerations such as regularity. This work helped in negotiations with pilot unions. The work of Dillon and Kontogiorgis refers to the specific case of US airlines, who have permanent reserve crew to fill open or disrupted pairings. Open pairings are crew pairings that do not have crew assigned. Dillon and Kontogiorgis generate call out day pairings for reserve crew and focus on generating sets of reserve pairings of varying lengths which also exhibit regularity. Longer call out pairings allow reserve crew to be used for more different types of pairings including long haul pairings. Generating varying length pairings allows for reserve crew who have different amounts of time off in a given month. Dillon and Kontogiorgis also found that certain times of the year require more reserve crew to be scheduled, for instance in December. In this thesis reserve pairings are of fixed length and the reserve crew being scheduled are not permanent reserve crew, which is because the problem being tackled is based on KLM practices, see Section 3.4 for more details.

The work of Rosenberger et al. [87] on the SimAir simulation tool (see Section 2.5.3) contains one section which concerns the use of reserve crew in response to disruptions, in which they introduce a weighted sum reserve policy for selecting reserve crew in the event of a crew related disruptions. In this thesis, a similar weighted sum reserve policy is developed in Section 6.4, which is incorporated into a (non-simulation) probabilistic model of crew related disruptions and subsequent reserve crew use. The reserve policy of Rosenberger et al. takes reserve-induced delay into account and deadhead times. Whereas, in this thesis, the reserve policy accounts for: reserve-induced delay; remaining reserve duty days; expected future demand and the level of reluctance to using reserve crew in roles below their assigned rank, but does not account for deadhead times. The reason for this is that the types of disruptions considered in this thesis require fast reactionary recovery actions and for these deadheading will not deliver reserve crew to the source of the disruption in time to absorb the disruption. Which is due to the typically low flight volume at spoke stations. Section 12.1.4 discusses the possibility of deadhead modelling in more detail.

2.4 Airline operations

Airline scheduling and planning begins up to 6 years before the day of operation and continues up to two weeks before the day of operation (KLM [57]). So two weeks before the day of operation the schedule is passed on to operations control. Operations control can make minor changes to schedules in response to disruptions leading up to the day of operations. Operations control also oversee the implementation of schedules on the day of operation where recovery decisions such as cancellations, using reserve crew, re-routing passengers and swapping crew and aircraft are made at short notice.

Rabbani [82] develops an airline recovery module for the MEANS (discussed in Section 2.5.3) airline simulator. MEANS is a simulation of the entire American national air traffic system. Rabbani describes operations control as consisting of a range of operators (referred to as desks) each with their own responsibilities, including: dispatchers; traffic management; operations; meteorological. Dispatchers form the main conduit of information flow between all operations control desks and the pilots air-side, they also oversee the refuelling of aircraft, the enforcement of maximum take off weight rules and general safety issues. The traffic management desk negotiates with air traffic control over departure slots and ground delay programs (ground delay programs restrict departures to airports with reduced capacity, which is for safety and congestion reasons). The operations desk has crew schedulers, fleet routers and passenger coordinators who implement recovery actions. The meteorological desk is concerned with the prediction of en route weather and can order cancellations if the safety of a flight may be compromised.

Reserve crew are required to replace delayed or absent crew, operations control may also consider implementing crew swaps, by swapping delayed crew with those currently available. In this thesis, models of reserve crew scheduling are developed which take the availability of other recovery actions such as crew and/or aircraft swaps into account. The models of Chapters 7 and 9 use simulation to estimate the availability of swap recovery actions before reserve crew are scheduled. Chapter 8 represents a theoretical approach to the same problem which does not rely on simulation.

2.4.1 The crew recovery problem

Medard and Sawney [69], Chang [26], Lettovsky et al. [61] and Abdelghany et al. [1] all address the crew recovery problem. In general, models for the crew recovery problem reschedule crew to restore crew schedule feasibility subject to cost minimisation and minimal changes from the original schedule. These objectives are often conflicting because a better crew recovery can be attained if more changes to the original schedule are allowed. Of these papers the work of Lettovsky et al. [61] and Abdelghany et al. [1] explicitly model the use of reserve crew. Medard and Sawney [69] give an account of the airline crew scheduling problem from crew pairing through to the recovery problem. Medard and Sawney [69] introduce an integer programming formulation for the crew recovery problem which is described as a crew pairing model with crew rostering constraints. The reason for this is that disrupted crew pairings need to be repaired and crew have to be assigned to the repaired pairings in a single model. The approach integrates crew pairing and crew rostering for crew recovery problems of limited size. Such an integration is intractable for the original crew pairing and crew assignment scheduling problems.

Chang [26] focusses on the pilot recovery problem. A genetic algorithm approach is presented which takes the original infeasible schedule as input. Crew feasibility constraints in [26] require a maximum of 10 flying hours per day and 32 flying hours per week. The author introduces an object oriented matrix chromosome structure. Each row of which corresponds to a pilot and the flights assigned to that pilot. Rows consist of what are termed CHROMOHEADS which correspond to specific pilots and CHROMOCELLS which correspond to the flights assigned to the pilot. The author selects a mutation rate for the genetic algorithm which is a sum of the violated hard constraints divided by the number of hard constraints multiplied by the number of cells in a chromosome. The algorithm terminates after 30000 generations.

Lettovsky et al. [61] also address the airline crew recovery problem and use an integer programming model to reassign crew and assign reserve crew to minimise the costs of cancellations and deadheading. A number of undisrupted crew pairings are selected for regeneration and reserve crew are treated as crew with empty pairings. Feasible continuations of partially flown pairings are then found using an integer programming formulation with the objective of minimising the total costs of the newly generated pairings, cancellations, deadhead legs and returning crew to their domiciles if pairings end at different stations. Lettovsky et al. [61] introduce problem specific column generation procedures and branching strategies for a branch and bound algorithm. The authors consider 3 disruption scenarios and shows that the solution quality as measured in terms of the number of covered flights increases with the number of undisrupted pairings selected for reassignment. So a trade-off exists between schedule recovery and the deviation from the initial schedule. The model also assumes that crew teams are unsplittable, meaning that new crew teams cannot be created by combining individual crew from different already existing crew pairings. The same assumption is made in this thesis (see assumption **C10** of Section 4.2).

Abdelghany et al. [1] introduce an approach for solving the hub-and-spoke network crew recovery problem which considers crew swaps, reserve crew and deadheading as possible recovery actions. Their model takes the current crew schedule and disruptions as input. Their model is aimed at a hub-and-spoke network. They reason that crew disruptions that occur at spokes are often difficult to deal with as there are few connecting flights to spokes as reserve crew tend to be stationed at the hub. They take the approach of solving crew disruptions at spokes by solving them at the hub before they occur. Their objective is to recover as many disrupted crew pairings as possible, with the least incurred recovery cost. Their objective function has cost contributions for crew swaps, reserve crew, deadheading and cancellations. Their solution approach solves crew disruptions sequentially in chronological order or earliest disruptions first. This means that the recovery decisions for disruptions are determined as disruptions occur. Such an approach is convenient for airlines who operate a rule-based approach to recovery as opposed to an optimisation based approach. This thesis makes a similar assumption (Section 4.2) where airline disruptions are solved in earliest scheduled departure time order first. Their model is also able to anticipate future disruptions, due to minimum rest rules, up to a day in advance and prevents these from occurring.

The crew recovery models of Medard and Sawney [69], Chang [26], Lettovsky et al. [61] and Abdelghany et al. [1] are capable of repairing crew schedules in the event of disrupted crew pairings with cost minimisation as the objective. In this thesis, disrupted crew pairings are modelled as being recovered on a case by case basis as they occur, using swap recovery actions, reserve crew, and cancellations. Whereas the work on the crew recovery problem considers the longer term feasibility of the crew schedule, in this thesis these disruptions are resolved using a rule based approach rather than an optimisation based approach. The work on reserve policies in this thesis could be used as an extension to the crew recovery models considered in this section.

2.4.2 Aircraft re-routing

The crew recovery problem is usually solved using the assumption that aircraft are not disrupted, however, if they are, aircraft re-routing may be required. Furthermore, re-routing aircraft can make the crew schedule infeasible, hence invoking a crew recovery problem. Rosenberger et al. [85] present a model for minimising the costs of re-routing and cancelling flights. Aircraft re-routing is decomposed according to fleet type to avoid complicating the knock-on effects for crew and passengers. In order to make their formulation solvable in real time, they introduce a heuristic to select a subset of aircraft that can be re-routed, the goal is to find a new aircraft routing in which all flights are covered with maintenance feasible routings or are cancelled. Their objective is to minimise the cost of re-routing aircraft and cancelling flight legs. Their model allows for ferrying (transporting aircraft to origin of the next flight), diverting (landing at a destination not initially intended) and overflying (skipping an intermediate destination). Flights are cancelled if they cannot be feasibly included in an aircraft routing or the cancellation threshold of 180 minutes is exceeded. They also use a delay threshold to identify delays for which aircraft re-routing may be desirable. A revised version of their model is considered which accounts for the effect the proposed re-routing has on delays and passenger connections.

The delay and cancellation threshold assumptions used by Rosenberger et al. are also used in this thesis. The delay threshold parameter specifies the minimum delay for which recovery actions (crew swaps and/or aircraft swaps and/or reserve crew use) are considered. Whilst after the application of delay recovery actions, flights are cancelled if their delay still exceeds the cancellation threshold.

2.4.3 Integrated airline recovery

When airline recovery is tackled sequentially, aircraft are re-routed first, then the crew schedule is recovered and then passengers are re-routed. This order is determined by relative cost of disruptions to the different layers of an airlines schedule.

Peterson et al. [76] address the full airline recovery problem from schedule recovery to passenger re-routing. They argue that sequential approaches to the airline recovery problem naturally lead to sub optimal solutions because of the conflicting objectives that exist between each problem. On the other hand a fully integrated approach is intractable, intractability is dealt with in [76] by considering a subset of the full airline recovery problem at a time, i.e. a selection of the disrupted flights and the affected aircraft, crew and passengers are rescheduled at a time. The schedule recovery part of the model has to adhere to time window constraints due to airport gate demand, the aircraft recovery part of the model has to minimise the number of changes to the original schedule (to protect maintenance routings). The crew recovery part of the model reassigns crew at a minimum cost whilst ensuring that non-reassigned crew are deadheaded back to their crew base. The passenger re-routing part of the model ensures that overbooking does not occur with respect to non-disrupted passengers. Their results indicate that an integrated recovery approach costs less than a corresponding sequential approach.

2.4.4 Disruption management

This section of the literature review focusses on decision support tools. The following papers are approaches that aim to aid airline controllers rather than replace them with an automated decision maker, as was the case in Sections 2.4.1 to 2.4.3. Teodorovic and Stojkovic [99] describe an approach to the operational daily airline scheduling problem. The purpose of the approach is to provide airline dispatchers with a decision support tool to aid their rules of thumb and enable them to make globally informed recovery decisions. Their approach is to generate new airline schedules in response to disruptions to provide the airline dispatcher with automatically generated recovery decisions which can be manually overridden. The objective they use is to minimise the number of cancelled flights. When two possible solutions correspond to the same number of cancelled flights the one with the smallest number of passenger delay minutes is chosen.

Kohl et al. [59] provides an overview of airline scheduling and operations. It also provides a good overview of the literature on airline disruption management and reports on the performance of a decision support tool for multiple resource recovery called the Descartes project. The Descartes project is an integrated decision support tool. It contains dedicated crew, aircraft and passenger solvers which are integrated using the "umbrella" system. The system was tested using disruption scenarios developed with the cooperation of operations controllers. They propose two architectures for integrated disruption management: integrated sequential recovery (ISR) and tailored integrated recovery (TIRS). The difference between the two approaches is as follows. ISR treats the dedicated solvers as black boxes, so solutions from one solver have to be checked for feasibility with respect to the solutions of the other solvers. TIRS on the other hand uses a common database for each dedicated solver, so changes made to one layer of the solution are immediately visible within the other solvers and the overall feasibility of a possible plan can be ascertained. The drawback of the TIRS approach is the increase in complexity which drastically increases the required solution times. Experimental results are given for the dedicated solvers but not the proposed integrated approaches.

Mathaisel [68] promotes the idea of a common graphical user interface for different departments of airline control (aircraft, crew and passenger departments). The idea being to enable easier communication between the different departments as well as a system which provides alternative airline schedules in the event of disruptions. The system allows the user to conduct what-if scenarios to check the feasibility of potential recovery actions, the system notifies the user if live events interfere with a what-if scenario. Mathaisel argues that many contributions to airline scheduling and operations in the literature are not utilised because each comes with its own set of input and output format requirements which make them cumbersome to implement.

2.4.5 Airline operations control

Clarke [34] provides an overview of the structure of airline control centres and the information flow within them. A review of decision support tools for recovery decision making in real time is also given. Clarke [34] gives the results of a survey on the causes of delays over 15 minutes in duration. The leading causes of disruptions are weather and maintenance, ground congestion is another important factor.

Kohl et al. [59] note that the roles of dispatchers tend to be slightly different in North America and Europe. In North America dispatchers follow the preparations of flights for take off and report possible problems, whereas in Europe aircraft control perform this task. The crew controllers of airline operations control are responsible for implementing reserve crew recovery actions on the day of operations. Currently no automated systems or advanced decision support tools or used to make these decisions [57]. In this thesis reserve policies that could be used by crew controllers on the day of operations are investigated. See Chapters 8 and 9.

2.5 Modelling uncertainty

The main reason why reserve crew scheduling has the potential to benefit from in depth research is the high level of uncertainty about the level of reserve crew demand on any given day. This section starts with a basic introduction to probability theory, then considers the specific case of how airline operation uncertainty can be modelled.

2.5.1 Introduction to probability theory

Probabilities are the natural tool for modelling events characterised by uncertain outcomes. Higgins [48] provides a good introduction to probability theory and stochastic modelling. The probability of an event is represented by a number between 0 and 1, where 0 corresponds to an event never occurring and 1 corresponds to an event always occurring. When exact theoretical models of uncertain events are not available the probabilities of events are usually estimated from samples of events. The larger the sample the greater the confidence is in the derived probability. $P(x)$ denotes the probability x is the outcome of a given event, therefore $P(X)$ corresponds to the probability distribution of possible outcomes (X) for a given event. Probability theory requires the definition $\sum_{x \in X} P(x) = 1$ which means that one of the set of possible outcomes occurs every time the event is repeated. Events come in two main varieties, those of dependent events and independent events. Independent events are events whose probabilities are influenced by previous events, whereas the probabilities of independent events are not influenced by previous events. The classic example is that of a bag containing marbles of two different colours, in this example an event corresponds to removing a marble from the bag at random and observing the colour. The probability of choosing a marble of each colour is equal to the relative proportions of the colours of the marbles remaining in the bag. If marbles are replaced after each observation the probabilities of observing different coloured marbles remains constant, therefore this provides an example of an independent event. If the marbles are not replaced the probabilities of observing marbles of different colours depends on previous events, which provides an example of a dependent event. Scenario trees [48] are a useful tool for modelling dependent events.

In all but the simplest of processes the outcomes of events depend on prior events (this is particularly the case when considering flight delays), in such cases a simulation model of a process may be used to estimate probabilities of events occurring at different times, see Section 2.5.3 for a review of approaches involving the use of simulation to model uncertainty in transportation networks.

2.5.2 Delay propagation

Reserve crew can be used to replace delayed crew, however before taking such an action it may be useful to evaluate the knock-on effects of recovery actions in terms of downstream delays. The following reviews existing approaches to the modelling of delay propagation.

The concept of delay propagation trees is introduced by AhmadBeygi et al. [3, 4], in [3] they illustrate the basic concept of a delay propagation tree in which each flight in a schedule is considered a root delay, and by performing slack calculations they trace how far a root delay of given length propagates. Delay propagation is then evaluated in terms of the number of subsequent flights delayed and the cumulative sum total of delay minutes. AhmadBeygi et al. [4] build on the ground work of delay propagation trees by using them as part of the objective function for a model that reallocates slack by re-timing departures such that the potential for delay propagation is minimised. They compare single layer and multilayer models (the difference being that in the single layer model delays are only ever considered as propagating to the next flight only and multilayer models place no such constraint on delay propagation). They found that the single layer model provided results as good as those obtained for the multilayer model and this was probably because of the slack that already existed in their real data flight schedules, which tended to absorb delay before the delay spread past the first flight.

Berger et al. [21] introduce a stochastic delay propagation model for calculating train departure and arrival time distributions. The calculation of event time distributions also allows for: delay absorption through slack in the schedule; the potential for trains to catch up by travelling faster; and waiting time rules for passenger transfers. Berger applies the approach in an online setting which receives messages about realised event times and recalculates the effect that they have on the uncertainty of future events. They show in experiments how prediction accuracy increases when more events become known as time progresses. Keyhani et al. [53] present a study of the reliability of connections in a train timetable using a similar underlying model to that in [21]. They derive connection reliability measures to advise passengers on the robustness of their itinerary, in terms of the probabilities of successful connections. The work of Berger et al. [21] and Keyhani [53] is conceptually similar to the statistical delay propagation model of Chapter 8 in the use of discrete journey time distributions to compute distributions for events depending on previous journeys.

Wong and Tsai [105] present a study of delay propagation in an airline network, in which statistical models for arrival and departure delays are developed. They develop hazard functions for arrivals and departures which give the probabilities that delays are recovered at any given time after a delay has started. They analysed real data to determine the effect that different types of disruptions have on how long the resultant disruptions last. In general they found that departure delays caused by crewing disruptions, aircraft maintenance and baggage handling had the greatest reduction in the probability of recovery for each unit increase in these types of disruption. They also found that the arrival delays caused by weather or insufficient block buffer time had the greatest reduction in the probability of recovery for each unit increase in these types of disruption. They used their results to advise airlines on what improvements to arrival and departure procedures would yield the greatest reduction of delays. They also considered the possibility of a recursive model, where their departure and arrival delay functions are used to track delay hazard along aircraft lines of flight, with the output of one model forming the input of the next. This is a similar concept to that used in the statistical delay propagation model of Chapter 8. However Wong and Tsai do not explicitly consider crew absence disruptions or the availability of recovery actions for delay and crew absence disruptions.

2.5.3 Simulation

Domestic aviation network simulators

Rosenberger et al. [87] describe a stochastic model of airline operations, it is basically a description and demonstration of a simulation tool called SimAir. SimAir is used for testing the quality of recovery policies and schedules in a stochastic environment that includes distributions derived from real data concerning the likelihoods of disruptions such as unscheduled maintenance, ground times and block times. The model can assess schedules and recovery policies in terms of operational crew cost, on time performance, and passenger misconnections. They demonstrate how taking operational costs into account in scheduling has the potential to produce schedules with lower operational costs, compared to state of the art approaches based on planned costs ([91] expands on this concept).

Clarke et al. [30] describe the aviation based simulation MEANS, which is a model of the American air traffic system and deals with both air and ground operations. The model has a modular structure (see Rabbani [82]) which means that any of the independent modules can be programmed to any degree accuracy without having to worry about the other modules. The modules include airlines, airport, weather and air traffic control, as a result MEANS can be used by almost any researchers involved in civil aviation. Clarke et al. [30] also describe the different versions of the modules available including historical playback of past data and "human in the loop" versions. The paper describes an investigation that demonstrates a high correlation between simulated events and actual events for two days of operations of the entire American air traffic system.

Simulations in general

The textbook of Rubino and Tuffin [88] addresses the study of rare events in complex systems for which direct methods (analytical) and numerical approximations result in intractable models. Rare events which may be of concern include the catastrophic failure of an aircraft. They show that Monte Carlo Simulation can be an inefficient approach to estimating the probabilities of such events. Especially given that the system being simulated may be have a high complexity, which means that it may not be possible to derive probabilities with a small enough confidence interval in a reasonable amount of time. They introduce methods for overcoming such a problem. One approach is known as Importance Sampling, this approach is based on finding an alternative distribution for the input random variable such that the target event becomes more probable. The task becomes that of finding the alternative random input distribution and how the actual probability of the target event relates to the biased probability estimate of the target event. Another approach for estimating rare event probabilities using Monte Carlo Simulation is known as the splitting technique, the idea is to make copies of simulations that get closer to the target event and use these as the start point of future simulation runs. Simulations that get further away from the target event are deleted. The probability of the target event is then estimated from the estimated number of paths that lead to the target event relative to the total number of paths. In relation to reserve crew scheduling, cancellations due to delays might be considered rare events, but their cost is low in comparison to catastrophic failures, which are the primary focus of rare event simulation methods. However, a splitting technique could be useful when estimating the consequences of crew absence, this is because there is a large number of alternative paths from any start point in an airline network in terms of future disruptions. A splitting technique could be used to identify the worst case outcome associated with different recovery decisions.

The technical report of Frick [39] describes a data driven approach to the study of transportation networks consisting of local area distribution centres and hubs. The main point of interest for mentioning this piece of literature is the idea of data driven modelling, this refers to automating the building of the model of the underlying transportation network. This process saves a lot of time manually changing a hard wired model of a transportation network.

When developing a simulation, such as that presented in Chapter 4, simulation flow charts are a useful tool for representing the logic of the the flow of decisions during any given situation. Waters' [100] textbook on management science defines the conventional elements of flow charts.

A closely related subject to simulation is that of system dynamics modelling [31]. System dynamics modelling is concerned with the optimisation of complex systems by manipulating the "pressure points" or parameters which control the process. In system dynamics "influence diagrams" (similar to a flow diagram) are used for the qualitative analysis of the behaviour of a complex system and simulations are used for quantitative analysis.

See Chapter 4 for the single hub airline simulation tool developed during this research.

2.5.4 Statistical distributions

This section reviews the various ways in which distributional uncertainty can be modelled. In this thesis numerical distributions for crew absence and journey time uncertainty are derived from real data.

Continuous distributions

It is often found that real world data (such as activity durations and the sizes of manufactured components) are well described by continuous mathematical distributions such as the normal distribution. Fitting real data to traditional theoretical statistical distributions often has the reward that the resultant models have properties (such as convexity) that can be exploited in the solution phase. The conference slides of Clarke [29] describe the use of arrival time distributions in a model for minimising the amount of delay propagation from the point of view of the maintenance routing problem. It is suggested that arrival times have a good statistical fit with a log-normal distribution.

Sohoni et al. [97] model block times with a log-laplace distribution, which provide a good fit with block times. This allowed them to simplify complicating chance constraints (functions of the decision variables) that appear in their model for improving the passenger service level of an existing schedule by retiming scheduled departure times within allowable time windows.

Intervals

Interval programming is mathematical programming where real valued coefficients are replaced with intervals. Intervals are used to model the uncertainty of parameters where the only information available are the lower and upper bound values of those parameters. The assumption is that the value of a parameter defined by an interval is uniformly distributed between the lower and upper bound values. The goal of interval programming is to find solutions to problems which have an optimal value with a narrow interval. The narrow interval indicates that the solution is stable with respect to the inherent uncertainties of the problem. Hossny et al. [50] tackle a machine scheduling problem in an algorithm where all arithmetic operations are replaced with interval arithmetic operations. The paper provides tables defining interval arithmetic operations and logical relations for intervals. They perform experiments to show that using interval arithmetic within the scheduling algorithm leads to a better trade-off solution (quality/stability) compared to the alternative approaches where solutions are calculated from crisp parameter values. When the crisp parameter values are the lower bounds of the parameter intervals, the expected completion time is minimised but also results in a high probability of delay. When the crisp values are the upper bounds the expected completion time is high and the probability of delay is low.

Fuzzy sets

Fuzzy sets are a method of expressing the imprecision of linguistic variables, in this way they are truth distributions. Fuzzy logic was first formulated by Zadeh in 1965, the goal was to make computers think more like humans (vague arithmetic). The book of Sakawa [90] covers fuzzy stochastic multi-objective programming, which shows that many real world problems have fuzzy objectives (and soft constraints). Such problems can be tackled using the objective of maximising the minimum membership over a set of fuzzy set constraints.

In this thesis fuzzy logic is not applied because there are quantitative variables for most aspects of the problem under consideration. However, there is one instance in which a function is defined for the subjective/perceived equivalence between delays of different lengths and a flight cancellation (Section 3.5.1). Such a function defines the membership of a delay of a given length to a cancellation.

2.6 Solution methodologies for deterministic problems

2.6.1 Exact methods: Linear and Integer programming

When people discuss exact methods, the first thing that springs to mind is linear programming. Linear programming is concerned with the modelling and solution of problems which can be formulated as mathematical problems with objectives and constraints which have linear terms only. Such problems can be solved with the simplex algorithm (described in [65]). The simplex algorithm exploits the linear structure of problems to traverse the vertices of the solution space, on which an optimal solution must lie. The book of Kallrath and Wilson [52] provides a detailed guide to mathematical programming, which includes linear programming, integer programming and mixed integer linear programming. The authors also explain an alternative to the simplex algorithm, that of interior point methods. Interior point methods, unlike the simplex algorithm, stay away from the boundaries of the feasible region (using penalties), and approach the optimal solution using a more direct route.

There is a vast amount of literature on integer programming applied to the airline scheduling problem [44, 55, 89, 12, 58, 13, 64, 29, 96, 94], in most of this work integer programming models are solved using methods such as branch and price (or column generation). Many refer to a process referred to as Lagrangian relaxation in which the integer requirement is dropped and the remaining problem is solved as a linear program. The solution times are very fast when the integer requirement is dropped, the final solutions are rounded up or down to the nearest integer values, the downside is that sometimes the solutions become either infeasible or non-optimal, however it is usually possible to derive a feasible solution from a non-feasible solution.

The branch and bound algorithm is often used to solve integer programs, in this algorithm a tree search takes place where nodes correspond to partially integer solutions and branches correspond to forcing an additional variable to an integer value above and below their decimal values in the root node's solution. Each branch is resolved to obtain a new node. To speed up the search, branch nodes are bounded if it is deduced that such a root node cannot correspond to an optimal solution. For minimisation (maximisation) objectives the objective value of each node solution is a lower (upper) bound on the objective value of the best solution that can emerge from that node, the upper bound (lower) is the objective value of the best full integer solution found. So if the lower (upper) bound of a partially integer solution exceeds (is less than) the upper (lower) bound, that node can be bounded and no longer branched on. The optimal solution is found when only one unbounded node remains in the tree.

For certain types of problems, specialised integer programming algorithms exist. A prime example is that of crew pairing, which can be formulated as a set partitioning problem. The crew pairing problem when formulated as a set partitioning problem requires that all possible crew pairings are columns in the input coefficient matrix, which for realistic sized problem instances results in an intractable problem size. The branch and price [14] algorithm circumvents this problem by only considering a subset of all possible crew pairings at a time. The algorithm iteratively alternates between a master problem and a pricing problem. The master problem is the set partitioning formulation and the pricing problem determines which columns/crew pairings should be included (column generation) the next time the master problem is solved. For an account of the application of column generation used to solve very large integer programs see the work of Barnhart et al. [14].

In Chapter 9 mixed integer linear programs are formulated and solved, CPLEX is used as the solver.

2.6.2 Meta-heuristics

Many problems can be formulated as integer linear programming problems, however the resultant models can be intractable. In which case one option is to try to develop a more efficient solution algorithm. If on the other hand the problem formulation does not satisfy the linearity or convexity requirements of linear programming solvers, meta-heuristics can still be used to search for good solutions. Meta-heuristics are solution paradigms often based on naturally occurring phenomena (such as evolution or ant colonies) which are emulated as algorithms to find good solutions to a given problem. Meta-heuristics include: local search; tabu search; variable neighbourhood search; genetic algorithms; particle swarm optimisation and ant colony optimisation. [83] contains a series of articles outlining the basics of many meta heuristics and guidelines for parameter selection for their implementation. The advantage of meta-heuristics is that there is no restriction on the complexity of the model of the problem being solved, all that is required is a means of evaluating the objective value of a given solution to the problem. In this thesis, the probabilistic models of Chapters 5 to 8 provide such a means of solution evaluation for the reserve crew scheduling problem.

Meta-heuristics can be divided according to local search and population based approaches. Local search approaches usually involve a single incumbent solution whose neighbourhood is explored to determine the most promising trajectory across the solution space to the best solution that is within reach. Population based approaches are used when a diverse set of possible solutions can be exploited in some way. In multi-objective problems population based approaches can be used to find a Pareto optimal set of trade-off solutions.

Population based approaches

The textbook of Burke and Kendall [25] describes (among other approaches) swarm intelligence methods such as ant colony algorithms, these approaches work on the basis that the collective behaviour of large numbers of simple individuals can appear to act in an intelligent fashion. The analogy for ant colony methods applied to shortest path problems is the most intuitive. Ants explore for food and once they find food return to the nest, ants being blind lay pheromone trails that they can use to retrace the steps back to the nest. The logic is that over time the shortest path to and from the food will end up with a stronger pheromone trail, because ants will make the highest number of trips backwards and forwards over this path compared to any other path over the same amount of time leading to more ants following the shortest path to the food.

In this thesis an ant colony optimisation approach is applied in Chapter 5 to schedule reserve crew, in which reserve crew are scheduled to begin standby duties at departure times visited by ants. Pheromone distributions are used to stochastically generate the ant paths, in each iteration the pheromone distributions are updated according an evaporation rate and the quality of reserve crew schedules generated by the ants.

Genetic algorithms (described by Burke and Kendall [25], Michalewics and Fogel [70] and Goldberg [43]) are based on the analogy of evolution and DNA, solutions are coded as strings (genetic code). Genes with desirable characteristics are selected and crossed over (analogy of sexual reproduction), the result is that the offspring (solutions) will have the desirable characteristics of its parents. Continuing in this fashion will lead to convergence towards a good solution. Premature convergence can be prevented with mutation and parameters that control the rate at which operators such as cross-over and mutation are applied to the population of solutions. Hart et al. [46] describe memetic algorithms and contains in depth information on practical applications. Memetic algorithms can be viewed as a hybrid technique of genetic algorithms and local search, in each generation of a genetic algorithm, local search is performed starting from each incumbent solution. Hart et al. [46] describe two different approaches to the implementation of memetic algorithms, one is Lamarkianism in which the result of the local search is directly used in the next generation of the genetic algorithm. The other approach is based on the Baldwinian effect in which the objective value achieved from the local search is used to determine which of the original solutions are placed in the next generation, but the local search solutions do not replace the original members of the population. The name memetic comes from meme which can be viewed as the intellectual equivalent of a gene or an idea, ideas can change in a life time, whereas genes are relatively fixed. The two approaches reflect possible types of interactions between memes and genes.

In this thesis a genetic algorithm is applied in Chapter 5, one of the issues faced was that cross-over often leads to infeasible solutions, this was circumvented by using a greedy correction heuristic to maintain a feasible population of solutions. In Chapter 10 a genetic algorithm is used in which the mutation operation is replaced with single iterations of a simulated annealing algorithm, making it similar to a memetic algorithm.

Local search based approaches

Other popular meta-heuristics with many reported successes include tabu search and simulated annealing. Tabu search, introduced by Glover [42] is a local search based algorithm which uses a short term memory of the solution space already searched to guide the search into new regions. The memory feature of the tabu search algorithm is used to encourage search diversification, the memory can also be used to store promising areas of the solution space which could be searched in greater detail, this corresponds to an intensification mechanism.

Simulated annealing, developed by Kirkpatrick et al. [54] is based on considering the relationship between combinatorial optimisation problems and statistical mechanics. Simulated annealing is often explained using the analogy of cooling molten substances (random initial solutions) slowly to create substances (solutions) with desirable properties such as large orderly crystals (optimised objective values). The tabu search and simulated annealing techniques can be applied to scheduling problems due to their versatility. The draw back of the wide ranging applicability of tabu search and simulated annealing (and genetic algorithms) is that representations of the solution space and parameter choices have to be worked out before the methods become effective, even then the solution obtained are not guaranteed to be optimal or even close to optimal.

Tabu search and simulated annealing are both applied in Chapter 5 and simulated annealing in Chapter 10.

2.6.3 Hybrid approaches

Grunert [45] tackles a direct flight network design problem based on that faced by a postal service. The problem consists of assigning flights between different depots to ensure that all letters reach their destination's local airport overnight. They present an integer programming formulation of the problem with capacity and flow constraints. They propose a hybrid tabu search and branch and bound solution approach. They start by finding an initial feasible solution using a greedy algorithm. In each iteration of the hybrid algorithm, a subset of integer variables are fixed and the remainder are solved in a branch and bound phase. Then a tabu search is performed on the variables that were fixed in the branch and bound phase in order to explore possible feasible changes to the fixed variables ready for the next iteration. The authors used a stopping criteria of a maximum of a thousand iterations. In computational experiments they investigate the effect of changing the parameters of the tabu search part of the algorithm in terms of solution time and the deviation from the best available (known) solution. In particular they found it beneficial to use a tabu tenure greater than 5.

2.7 Solution methodologies for stochastic problems

This section considers the existing research on solution methodologies for problems with uncertain input parameters. The main methods are stochastic programming and robust optimisation. Some variant models for robust optimisation are discussed as these, in part, inspired the work on the scenario-based approach to reserve crew scheduling in Chapter 9.

2.7.1 Stochastic programming

In Shapiro et. al [93] stochastic programming is described as being a very broad subject area that lacks a characteristic standard problem formulation, but instead is a framework of conditions and techniques under which the more tradition approaches to mathematical programming can be applied to problems with uncertain parameters. In particular they state that stochastic programming requires that the feasible solution space is convex (just as in linear, quadratic and convex optimisation). However they state that convexity requires the assumption that the probabilities of future events (in multi-stage problems) do not depend on previous (stage) decisions. This is an assumption which is not justifiable for many real world problems. An example where this independence assumption is invalid occurs in reserve crew scheduling. The probability of reserve crew demand depends on the recovery decisions made previously, which may directly or indirectly influence the current demand for reserve crew. For example, a crew absence may have been covered previously, meaning such a disruption does not still need to covered. Or a crew-related delay covered previously by reserve crew will not propagate and require reserve crew at a later time. In general the probabilities of disruptions in the future do depend on the recovery decisions made in the past.

Kall [51] explains that solving problems with uncertainty using the expected values of uncertain parameters can lead to solutions that are not suitable in any of the possible outcomes of that problem. They give the example of a development project with two possible outcomes in terms of profit, either very high or very low. Basing the optimisation problem on the average profit will not reflect either possible outcome. So stochastic programming differs from normal mathematical programming problems in that it fully acknowledges the effect that decisions have in each of a set of different possible outcomes. Where typically, the different possible outcomes have associated probabilities of occurring. In general a stochastic programming solution will give a solution which performs well over a range of possible outcomes, as opposed to a solution that is optimal in the expected outcome. In this thesis the probabilistic models are based on expected outcomes. In Section 6.1.7 this leads to a problem which is solved by extending the model to account for a range of expected outcomes. In stochastic programming problems, uncertainty can be modelled in several ways. Chance constraints (Shapiro et al. [93], Kall [51] and Birge and Louveaux [23]) can be used when a constraint which depends on uncertain parameters has to be satisfied with some specified probability (this approach is used by Sohoni et al. [97]).

Uncertainty in stochastic programming problems can also be captured with error terms for the uncertain parameters of the model. The error terms usually take the form of statistical distributions. Solution techniques applicable to stochastic programming problems often derive discrete approximations of the error term distributions so that a range of possible outcomes can be explicitly included in the constraints of the stochastic program.

Stochastic programs can be defined by the number of stages involved in solving the model. Single stage models (Birge and Louveaux [23]) solve the entire the problem in a single stage so that the decisions and recourse (recovery) actions are determined at the same time. In two-stage stochastic programming, the decision variables are solved first and the second stage solves the recourse action variables after some information on the outcomes becomes available.

The mixed integer programming simulation scenario model of Chapter 9 is influenced by approaches to stochastic programming (as well as others), in its explicit modelling of the effect that a reserve crew schedule has on the (post recovery) expected level of disruption in each of a set of disruption scenarios.

2.7.2 Robust optimisation

Bertsimas et al. [22] define robust optimisation as a general approach to formulating and solving optimisation problems with uncertain parameters, in which uncertainty is captured in the form of independent sets of (simultaneous) realisations of those parameters. These sets of uncertain parameters are referred to as the uncertainty set. A single element of the uncertainty set corresponds to a deterministic equivalent of the problem being solved. The solutions to such formulations must be feasible with respect to the uncertainty set.

Robust optimisation differs from stochastic programming in several ways.

- Stochastic programming models parameter uncertainty using probability distributions, whereas robust optimisation uses varied sets of well defined parameters, where each set corresponds to a deterministic equivalent of the problem being solved.

- Stochastic programming is motivated by theoretical developments, whereas robust optimisation is motivated by tractability/solvability. Robust optimisation is a pragmatic approach to solving optimisation problems with uncertain parameters.

- Stochastic programming allows some constraints to be satisfied with a threshold minimum probability, whilst robust optimisation requires feasibility over the entire uncertainty set.

Robust optimisation is often criticised for its inflexibility with respect to the requirement that solutions have to be feasible over an entire uncertainty set. Section 2.7.3 reviews work on several variants of robust optimisation where the strict feasibility requirements have been relaxed in some way, the examples given apply to the research area of train scheduling, a domain with many parallels to airline scheduling.

2.7.3 Variants of robust optimisation

Light robustness

Fischetti and Monaci [38] introduce the concept of light robustness as an alternative to robust scheduling and stochastic programming approaches. The advantage of their approach is that the strict requirement that solutions are feasible over the entire uncertainty set is relaxed, meaning the resultant solutions are not over conservative. The example they use is a linear program with a coefficient matrix whose values take on uncertain values. In light robustness a maximum deterioration limit of the objective value is set, slack variables capture the violation of constraints and hence the robustness of the solution. The auxiliary objective function is to minimise the slack such that the resultant solution is as robust as possible with an objective function value that is acceptable. The light robustness approach is similar to stochastic programming in that the slack variables play a similar role to the secondary recourse variables in stochastic programming. As in recoverable robustness (see Section 2.7.3) there is a set maximum number of parameters that can simultaneously take on their worse case values.

Recoverable robustness

Liebchen et al. [63] Introduce the concept of recoverable robustness, they describe the concept as the integration of scheduling and delay management. They state that recoverable robustness differs from strict robustness because the solutions from recovery robust optimisation only have to be within a recoverable distance from a fully feasible solution given a recovery algorithm. Whereas, strictly robust solutions have to be feasible in all scenarios before the consideration of recovery, and as a result leads very expensive solutions. The specific example used by Liebchen et al. is based on the deterministic timetabling problem, which can be represented as a graph with nodes corresponding to events (trains arriving and departing from stations) and vertices representing activities (driving between stations, picking up and dropping off passengers at stations). The variables in such a model are the times corresponding to event nodes. The objective is to minimise the sum of the product of activity lengths (arc lengths) and the number of passengers expected on the corresponding arcs, as to minimise overall passenger waiting time. The recovery algorithm for their train timetabling example permits modifications of the scheduled times of departures within certain allowable limits. They define a recovery robust optimisation problem mathematically as consisting of three sets, the original optimisation problem, a set of (disruption) scenarios and a set of recovery algorithms. A feasible solution is defined as a solution that is feasible in all scenarios in the scenario set given the set of recovery algorithms. The disruption scenarios considered each contain a maximum of a single disruption. Cicerone et al. [28] generalise the concept of recoverable robustness to the case where disruption scenarios contain any number of disruptions, as they find the solution given by Liebchen et al. [63] unsatisfying. In other work Cicerone et al. [27] apply the approach developed by Liebchen et al. [63] to the railway shunting problem in which railway cars are assembled according to their final destination, the problem is that the arrival times of trains at the shunting yard is uncertain, which can lead to inefficient shunting operations that lead to more delays. Cicerone et al. [27] highlight the fact that for each application of recoverable robustness in different domains the challenge is that of formulating the applicable recovery algorithms.

Recoverable robustness is not a directly applicable approach for solving the reserve crew scheduling problem, because in recoverable robustness the recovery actions are fixed like a policy, whereas in reserve crew scheduling the recovery action (reserve crew) is what is being scheduled. However, recoverable robustness could be applied to the crew scheduling problem, with reserve crew forming the fixed recovery algorithm. It follows that it could be beneficial to integrate crew scheduling and reserve crew scheduling, perhaps using an iterative approach as used by Weide (Section 2.2.3) to gradually increase the overall schedule robustness.

2.7.4 Methods for multi-stage decision making

As stated in Chapter 1 one of the goals of this project is to investigate online reserve use policies in terms of their effect on the expected levels of day of operation disruptions. The online reserve crew use problem can be cast as a multi-stage decision problem, where stages correspond to possibly disrupted flights. When a flight is disrupted a recovery decision is required which should take into account the remaining available reserve crew and the potential for future disruptions for which those reserve crew may be better utilised. Should reserve crew be used to absorb an immediate disruption or alternatively should the reserve crew be held in anticipation of possibly larger disruptions that may occur later? In this section a review of the family of techniques applicable to multi-stage decision making problems is given.

Bellman's dynamic programming [20] provides the basis for most approaches for multi-stage decision making problems. Dynamic programming works well for problems with the correct properties, provided that the number of states and available actions at each state are small enough to avoid computational intractability. For dynamic programming to be a suitable solution technique, a problem is typically required to have the property that the optimal decision at each stage depends only on the current state of the system and not on the history of how that state was reached. Dynamic programming is viewed as an efficient enumeration algorithm. Dynamic programming can be implemented in forwards and backwards varieties. Backwards dynamic programming starts from the final stage and works back to the first stage. Backwards dynamic programming can be used for problems for which the desired state in the final stage is known. In each stage of backwards dynamic programming, the values of states are updated according to the best state they can reach in the next stage (considered previously) plus the (transition) cost of reaching that state. Proceeding in this fashion all the way back to the initial stage will reveal the optimal sequence of decisions required to reach the desired final state. Depending on the context of the problem that is being solved, forwards dynamic programming is also possible. Dijkstra's algorithm is a special case of forwards dynamic programming.

The textbook of Powell [79] gives the Bellman equation as follows.

$$V_t(S_t) = \max_{x_t}(C_t(S_t, x_t) + V_{t+1}(S_{t+1})) \tag{2.4}$$

Equation 2.4 states that the value (V_t) of being in state (S_t) is equal to the value of the decision (x_t) that leads to the greatest sum of immediate profit (C_t) (cost of the decision) and the value of the state you end up in (V_{t+1}) having enacted that decision. For some problems finding the value function (V) is the biggest problem, especially when external uncertainty exists in the transition from the current state to the new state, given an implemented decision. When external uncertainties are present, the $V_{t+1}(S_{t+1})$ term is replaced with the expectation term $E\{V_{t+1}(S_{t+1})|S_t\}$, which makes finding the optimal value function even more difficult.

Markov decision processes (described by Powell [79]) are multi-stage decision making problems that are characterised by: states which obey the Markov property; states which have discrete sets of possible actions; known models of uncertainty for state transitions given a current state and decision; and a known reward for each outcome. The Markov property states that the value of a state in a system is the same regardless of the history that led to that state. The aim when solving Markov decision problems is to find a policy which returns an action for each state such that the objective contribution accumulated over the time horizon of the problem is optimised. If the time horizon is infinite or the rewards for decisions made now have a delay, a discount factor is used to assess the current value of rewards with respect to an interest rate. This approach ensures that the yielded policy finds the optimal decision based on the current value of a decision. Several different approaches to solving Markov decision making problems exist. Powell [79] describes policy iteration which starts with a policy, then determines the value of the policy, then uses this information to update the policy. This process continues until some convergence criteria on the value of the policy is reached. Alternatively, value iteration uses a value function which stores the value of all possible decisions in all states. Value iteration iteratively determines the value function until it converges, the optimal policy can then be extracted from the value function by selecting the highest value decision for each state.

When multi-stage decision problems do not have known models for the transition function, or have too many states and actions, the techniques of approximate dynamic programming (Nascimento and Powell [72], Powell et al. [81], Powell [80] and Balakrishnan [6]) can provide a solution. Approximate dynamic programming is an umbrella term for a wide range of techniques. In general approximate dynamic programming can be used whenever any of the above described approaches cannot be applied because one or more of the aspects of the given problem leads to computational intractability. For example, the expectation term in Bellman's equation for solving dynamic programs, with environmental uncertainties, can become intractable and as a result requires a Monte-Carlo simulation based approximate dynamic programming approach to approximate it.

Nascimento and Powell [72] apply approximate dynamic programming to the energy dispatch problem. The problem is to allocate stored energy and energy from external sources to different energy pathways. They show that this problem suffers from the curses of dimensionality and apply a Monte-Carlo based approximate dynamic programming approach. In the simplified model the state is defined as the total amount of stored energy, the contribution function sums the costs of under supply, over supply, discarding, buying and selling energy. The value function is piecewise linear and they use linear programming to determine value function slopes.

Powell et al. [81] consider a vehicle routing application of approximate dynamic programming in which orders become known in real time. Drivers have to be kept close to where orders are most likely to occur. Powell [80] provides a manual for practitioners considering using approximate dynamic programming, it warns that there are three main communities that use approximate dynamic programming (control theory, computer science and operational research) and each has there own set of standard notation and terminology. Powell [80] also gives guidance for step size schemes used to update the value function. $1/n$ is described as theoretically sound, and is equivalent to averaging the observed values obtained from the given state over n runs. Powell [80] also gives guidelines for a successful application of approximate dynamic programming, stating that all aspects of a problem's structure must be exploited.

Balakrishnan [6] describes an application of approximate dynamic programming to the problem of airport taxi time prediction, a diffusion wavelet value function is experimented with and compared to other value function structures. The following variants of approximate dynamic programming represent different ways to learn value functions. Reinforcement learning (Si et al. [95] book chapter 2) is a strand of approximate dynamic programming and can be used for multi-stage decision problems for which explicit transition functions and reward functions are not available, but a simulator of the process of concern is available. Reinforcement learning learns how to make decisions based on trial and error. Q-learning [79, 95] is a technique of approximate dynamic programming which uses a state/decision value function. This approach is most applicable when the number of decisions possible in each state is small. The entries in the state/decision pair value function are referred to as Q-values and store the value of making the given decision in the given state.

In this thesis, probabilistic models are developed (Chapters 5 to 8) which evaluate potential reserve crew schedules in terms of their effect on the expected level of delay and cancellation disruptions that occur on the day of operation. When these models are applied in an online context (Chapters 8 and 10) to evaluate alternative reserve use decisions, they perform the same role as that of a value function (see above). In Chapter 4 a look-up table value function is derived from simulation which is applied in Chapter 10 as a reserve policy.

2.8 Aspects of problem solving

This section considers aspects of problem solving in general.

2.8.1 Modelling

Mathematical modelling is the process of taking a real world problem and converting it into a mathematical problem. The book of Edwards and Hamson [37] gives a guide to mathematical modelling. Through numerous example problems they outline the stages that are often followed in a mathematical modelling cycle. The process will start with some sort of problem description. The problem is then usually abstracted by making a number of assumptions to identify the important variables. A solution is formulated and tested to identify any problems. This process continues iteratively until the model gives the required solution or the model matches the reality. This process is mirrored in this thesis.

2.8.2 Solution space

Michalewics and Fogel [70] discuss the importance of knowing the structure of the solution space, this allows for the development of enumeration algorithms and is also a prerequisite for developing solution representations that are required to apply modern heuristic methods such as simulated annealing, tabu search and genetic algorithms. For a typical scheduling problem the solution space consists of all possible allocations of resources to tasks which do not violate any of the constraints. When resources are identical the solution space consists of all combinations of tasks assigned to a given number of identical resources, when resources are distinct, for example all personnel have different skill sets, the allocation of resources to tasks has a solution space consisting of all permutations of resources assigned to the given tasks. Pinedo [77] states that machine scheduling problems can also have the property that the order in which tasks are completed effects the time taken to complete each tasks. This means that the scheduling problem has two interacting layers (task allocation and task ordering) which makes the solution space even larger than a typical permutation sized problem.

2.8.3 Analogies with other problem domains

There are a host of other problem domains which have analogous problem structures to those of the problem of reserve crew scheduling under uncertainty. In scheduling tools and algorithms (by Pinedo [77]) an analogy is drawn between machine scheduling and airport operations, namely that gates can be treated as machines and scheduled aircraft departure times treated as job completion times.

The topic of inventory control provides another possible source of analogies, see the book called factory physics [49] for more information on inventory control and related subjects. A reorder policy approach would be most applicable to the problem of callout reserve crew, who are stationed at home and can be called to the airport when the number of reserve crew available at the airport drops below a certain threshold (the reorder point). Winston [104] also covers inventory control, presenting the traditional economic order quantity (EOQ) models. Winston points out that these models are often based on the unrealistic assumption of constant demand. $var(demand)/\bar{demand}^2 \leq 0.2$ is given as the condition under which such an assumption is reasonable. An EOQ approach would therefore be appropriate for manpower planning, i.e. determining numbers of reserve crew required each day. But when considering the allocation of standby duty start times at a per flight level the per flight demand may be too irregular for a direct application of an EOQ model.

2.9 Chapter summary

This chapter has provided the necessary background information required to undertake research on the problem of reserve crew scheduling under uncertainty: A wide knowledge of airline scheduling and airline operations helps to understand the wider ecosystem of which airline reserve crew are part; A detailed knowledge of the current literature on reserve crew scheduling helps to identify the niche that this research is trying to fill, and; A survey of the current approaches to modelling uncertainty and solution methodologies for such problems provide food for thought on how to proceed.

Chapter 3

Problem description and definitions

In this chapter Section 3.1 gives a general formulation for the problem of airline reserve crew scheduling under uncertainty. The problem has two main elements, those of offline reserve crew scheduling and online reserve policies, these aspects of the overall problem are described in more detail in Sections 3.2 and 3.3 respectively.

Chapter structure

Section 3.1 gives a general problem formulation for airline reserve crew scheduling under uncertainty. Section 3.2 lists the factors that should be considered when trying to schedule reserve crew in a good or optimal way. Section 3.3 lists the research questions relating to the secondary objective of this research, that of online reserve policies. Section 3.4 discusses the details of the real world case study that is considered in this thesis, which is based on the operations of KLM. Section 3.5 defines a number of concepts and conventions that are common to the remaining chapters of this thesis. Section 3.6 summarises the main points from this chapter.

3.1 General problem formulation

The combined problem of the offline scheduling and the online utilisation of reserve crew with the objective of minimising schedule disruptions under operation uncertainties can be formulated as in Equations 3.1 and 3.2. The explanation of these equations and the definitions of the variables (x and y) and the inputs (S, U and P) are given in Table 3.1.

$$\min_{x \in X,\, y \in Y} E\left(f\left(x, y, S, U, P\right)\right) \tag{3.1}$$

$$x \in X | \sum_{j=1}^{n} x_{i,j} = R_i,\, \forall i \in T \tag{3.2}$$

The main overall objective (Equation 3.1) of this thesis is to find a combination of a reserve crew schedule (x) and a corresponding reserve policy (y)

		Decision variables
x	:	Reserve crew schedule, $x_{i,j}$ specifies the number of reserve crew of type i assigned to standby reserve duties that start daily at time index j.
y	:	A reserve policy y specifies if reserve crew are to be used to replace delayed or absent crew in any given situation, and if so, which reserve crew of those available are to be used. The reserve policy space is denoted y and includes rule of thumb policies as well as optimisation based policies.
		Inputs
S	:	The airline's schedule S takes the form of a list of scheduled flights, listed in earliest scheduled departure time order. For each scheduled departure the schedule specifies the scheduled departure time, the destination, and the crew and aircraft who are assigned to that flight.
U	:	Uncertainty U in this case takes the form of unexpected crew absence and unknown journey times. For each of these statistical distributions are derived from real data.
P	:	The airline's recovery policy P specifies how the airline will respond to any given disruption. For example, if a departure is delayed due to the delayed arrival of a previous flight then the airline will consider swapping the resources assigned to the delayed flight in order to mitigate the delay. The reserve policy y defined above is an element of the airline's overall recovery policy, because their are alternative recovery actions that the airline may like to consider before using reserve crew.
n	:	The number of hub departures in the airline's schedule during the time horizon over which the available reserve crew are being scheduled.
T	:	The set of reserve crew rank-qualification combinations

Table 3.1: Problem description notation

that minimises the expected (E) level of delay and cancellation disruptions that occur when an airline schedule (S) is implemented. The schedule is implemented in an environment which is subject to uncertainty (U), namely crew absence uncertainty and journey time uncertainty, both of which perturb the scheduled events with respect to those planned.

The airline implements a recovery policy (P) to recover from schedule disruptions. Besides using reserve crew, the airline may have alternative recovery actions such as swapping crew and aircraft or cancelling flights. The reserve policy is an element of the the airline's overall recovery policy.

In Objective 3.1, X is the set of feasible reserve crew schedules. A reserve crew schedule is feasible if no more than the available set of reserve crew are scheduled. This can be expressed by Constraint 3.2, in which T is the set of types of reserve crew and R_i is the number of reserve crew of type i which are available for scheduling. A reserve crew type is defined by a combination of a rank (which defines the roles they can undertake), and a qualification (which defines the fleets they can operate on).

A reserve policy y determines whether or not reserve crew should be used to cover a given crew related disruption, and if so, which reserve crew of those available should be used. A reserve policy y can be a simple rule of thumb such as using reserve crew as demand occurs, in earliest start time order (the default policy, see Section 3.5.2). It could also be a function of the given reserve crew schedule x and the current state of the airline's operations as well as the expected future demand for reserve crew at the given time (see Section 8.2.5).

The main difficulty in the problem defined by Objective 3.1 and Constraint 3.2 is in how to accurately and efficiently compute the expected level of disruption associated with a given combination of a reserve crew schedule x and a reserve policy y. The difficulty is caused by the presence of operational uncertainty (U) when implementing a schedule S. In particular, crew absence and journey time uncertainty means that there are a multitude of ways in which a day's operations can unfold, which makes computing the expectation term of Objective 3.1 challenging. This problem is made more difficult by the non-trivial interaction that exists between the airline's recovery policy and the outcomes of uncertain events on the day of operation. This is because small changes in the outcome of an event can change which recovery action is applied, which in turn changes the recovery actions which will be available later on. Furthermore, there is the potential for an interaction between the reserve schedule x and the reserve policy y. For example, changing the reserve policy can change what the optimal reserve crew schedule is, and changing the reserve crew schedule can change what the optimal reserve policy is. This last point is the main justification for making online reserve policies a secondary research objective in this thesis.

The reserve crew scheduling problem can be formulated as a two-stage stochastic integer programming problem, where the first stage variables define the reserve crew schedule. The second stage variables determine how those scheduled reserve crew are used given any realisation of a disruption scenario. The variables in both stages are integer decision variables. Dyer and Stougie [36] provide a proof for the #P-hard complexity of two-stage stochastic integer programming problems with discretely distributed parameters.

3.2 Offline reserve crew scheduling

When trying to schedule reserve crew in a good or optimal way the following issues need to be considered. For each issue a forwards reference is provided to the section of the thesis that address that issue.

- **Reserve crew demand is influenced by journey time uncertainty**. The time between an aircraft leaving the gate at the origin station and arriving at the gate of the destination station is known as the block time. Block time (see Figure 3.1) includes the time spent

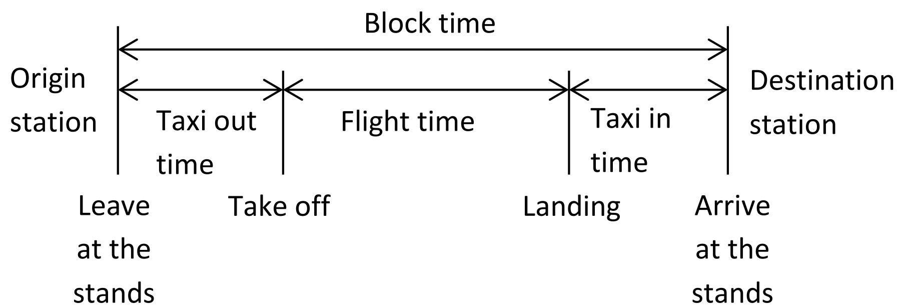

Figure 3.1: Block times and their constituent parts

travelling between the gate and the runway at the origin and destination stations (known as taxi time), and also the time spent in the air in transit to the destination. In this thesis taxi times times are included as part of the journey time and so journey times correspond to block times. Journey times are influenced by congestion and weather conditions, giving rise to journey time uncertainty. Reserve crew can be used to replace delayed crew, and therefore reserve crew demand is influenced by journey time uncertainty.

In this thesis, Chapter 7 introduces a probabilistic model of reserve crew demand as a result of journey time uncertainty. Chapter 8 introduces a statistical delay propagation model. Chapter 9 introduces a scenario-based approach to modelling reserve crew demand due to crew-related delays as well as crew absence.

- **Reserve crew demand is influenced by crew absence uncertainty**. Crew can, for various reasons, be absent at short notice, this may be because of illness or other extenuating circumstances. When crew are absent a flight may be prohibited from flying, unless reserve crew are used to replace them.

  In this thesis Chapter 5 introduces a probabilistic model for the allocation of reserve crew used to cover for absent crew. Chapter 6 improves upon this initial model, by addressing the simplifying assumptions made in Chapter 5. Chapter 9 introduces a scenario-based approach for scheduling reserve crew in anticipation of both crew absence and delay disruptions.

- **Other recovery actions may reduce reserve crew demand**. In the event of delays which can be mitigated by replacing the delayed crew with reserve crew, other recovery actions may also be available. Swap recovery actions can be used to absorb delays, which may reduce the need for reserve crew at such times. When scheduling reserve crew an appreciation of the airline's recovery policy could help to avoid scheduling reserve crew when they are not needed.

  In this thesis, the probabilistic model of Chapter 7 uses a simulation, which implements swap recovery actions, to learn how crew-related delays propagate through an airline's schedule. This information is then used to schedule reserve crew in a way that minimises the total expected crew related delay whilst also taking the availability of swap recovery action into account. The scenario-based approach of Chapter 9 also uses simulation to provide a means of allowing for the availability of swap recovery actions when scheduling reserve crew. Chapter 8 introduces a statistical delay propagation model, which does not rely on simulation, that is able to calculate departure time distributions for all flights in a schedule as a function of an airlines recovery policy, including swap recovery actions. Section 8.1.4 shows how this is possible.

- **The times at which reserve crew are scheduled influences the potential for reserve crew induced delay**. When reserve crew are used to cover for absent crew it may be the case that the affected flight has to be delayed until the reserve crew begin their standby duty. Scheduling reserve crew at different times will on average induce different levels of reserve-induced delay.

In this thesis, Section 6.1.3 shows how the improved probabilistic crew absence model allows for reserve-induced delay. Only Chapters 5 and 7 do not explicitly allow for reserve-induced delay. Instead, these models use the simplifying assumption that reserve crew can only be used for disruptions which occur whilst they are on standby duty and not before.

- **The airline's reserve policy influences the effectiveness of any given reserve crew schedule**. In the event of disruptions which can be mitigated by using reserve crew, the airline may have a policy for selecting which of the available reserve crew to use or whether to use reserve crew at all. Knowledge of an airline's recovery policy can be used to find an associated reserve crew schedule that works best with that policy.

  In this thesis reserve policies are a reoccurring theme, Figure 1.1 and Section 3.3 give more details.

- **Reserve crew have ranks and qualifications that limit which roles and fleet types they are feasible for**. Airlines often have a range of fleets (aircraft types). Reserve crew must be qualified to operate on the fleet type associated with a given disruption. Furthermore, reserve crew have ranks which specify which roles on flight legs they can fulfil. So an approach to optimising a reserve crew schedule should also take into account the ranks and qualifications of the reserve crew being scheduled.

  Sections 6.3 and 9.8 are concerned with modifying the probabilistic and scenario-based models, for the case of multiple fleet types, crew ranks and qualifications.

- **The structure of the airline's schedule dictates how disruptions may propagate through the schedule**. The structure of an airline's schedule in terms of the scheduled departure and arrival times influence how likely it is that a delay from one flight will spread to the next flight. If a short turnaround time is scheduled between the arrival of one flight and the departure of the next, the probability of delay propagation is increased. Knowledge of the airline's schedule can be used to identify flights which have a high risk of delay propagation. Reserve crew can then be scheduled to minimise the potential for delays affecting those flights.

  Exploiting the structure of an airline's schedule is precisely the aim of the probabilistic crew delay model of Chapter 7. Chapter 8 improves upon this initial probabilistic crew delay model by modelling delays in general using a fully theoretical model which does not rely on a simulation learning phase.

- **The structure of crew pairings determine the maximum number of cancellations in the event of uncovered crew absence**. When crew absence occurs which cannot be covered by using reserve crew, flights may need to be cancelled. In a hub and spoke network, a cancellation at the hub also requires the cancellation of the subsequent return journey from the spoke back to the hub. Cancellations continue until absent crew can be replaced with reserve crew. So the maximum number of cancellations for any given crew pairing is the number of flights in that crew pairing. Ideally reserve crew would be scheduled to prevent all cancellations. However, this is not always possible, instead reserve crew can be scheduled in a way that minimises the total potential for flight cancellations due to crew absence.

Section 6.1.1 shows how the improved probabilistic crew absence model takes the structure of crew pairings into account when scheduling reserve crew. The scenario-based approach of Chapter 9 also allows for the effects of the structure of crew pairings, see Section 9.2.2.

3.3 Online reserve policy

A reserve policy determines whether a crew related disruption should be covered by using reserve crew, and if so, which of the feasible reserve crew should be used. The investigation of online reserve policies is closely related to the primary objective of investigating approaches for offline reserve crew scheduling. When scheduling reserve crew, the knowledge of an assumed reserve policy can help to improve the quality of the reserve crew schedule. So the question is, what is the best reserve policy?
In this thesis the investigation of online reserve policies considers the following questions.

- **Given a crew related disruption and a limited availability of reserve crew, is reserve crew use or reserve crew holding the most appropriate action?** On a given day the demand for reserve crew may be low or high, the best reserve policy should adapt to the conditions on the day of operation as events unfold. For example, whether or not reserve crew are used to cover a small delay should depend on whether doing so significantly increases the risk of being unable to recover from larger disruptions later on, such as crew absence or large delays. It may also be the case that a swap recovery action is available to mitigate the same disruption, which may be a cheaper recovery action and therefore a more favourable option to the airline. A reserve holding policy needs to be able to make globally informed decisions that may not necessarily be immediately beneficial. Sections 3.5.2, 4.7.2, 4.7.3 and 8.2.5 consider a range of possible reserve holding policies, which are all compared with one another in Section 10.5.

- **Given that reserve crew come in a range of rank and qualification combinations and there exist numerous combinations of reserve crew which would be feasible to cover for a given crew related disruption, which combination should be utilised?** The demand for reserve crew with particular ranks and qualifications can vary on a day to day basis. On some days some reserve divisions (groups of reserve crew with the same rank and qualification) may be in high demand. On such days an adaptive reserve policy should consider using the combination of reserve crew with the lowest expected future demand. In other situations it may be beneficial to consider the possibility of using reserve crew in roles below their assigned rank. What is required is a policy which can provide a recommended decision based on the merits of using different combinations of crew in terms of the immediate benefit, long term benefit and the expected future demand for reserve crew. Section 6.4 introduces a parameterised policy for selecting reserve crew combinations taking the above described considerations into account.

- **How can reserve policies be modelled during offline reserve crew scheduling?** Once a reserve policy is found that performs well in operations, the problem then becomes that of how this reserve policy can be taken into account when scheduling reserve crew. Section 5.2.1 introduces the basic modelling principle which enables the modelling of a reserve policy in a probabilistic reserve crew scheduling model.

3.4 KLM specific problem

As described in Section 1.1, this research project arose from a collaboration between the University of Nottingham and KLM. As a result the problem tackled in this thesis is based on KLM practices.

KLM operate a single hub and spoke network, with Schipol as the hub station. Almost all flights involve Schipol as the origin or destination (a minority involve other intermediate stops). As a result of this network structure, smooth operations at the hub are of vital importance. This thesis is focussed on the scheduling of reserve crew who will be on standby duty at the hub station and will be used to cover for disruptions that occur there.

At KLM reserve crew are regular crew who are contractually obliged to undertake a number of reserve blocks per year. Reserve blocks are two weeks in length. The purpose of reserve blocks is to give the crew schedule a level of recoverability from crew-related disruptions. In the first five days of a reserve block, reserve crew can be used to adopt any crew disrupted pairing. If they are not used in this period they get two days off, in the second week they can only be assigned to open pairings (short sequences of flights left unassigned exactly for this purpose). If reserve crew are used to cover crew disrupted pairings in the first week, the reserve crew can only be used to cover open pairings in the remainder of their reserve block provided that their minimum rest requirements are satisfied. The two week length of a reserve block allows reserve crew to be used for any disrupted pairings (long or short haul) in the first week of the reserve block whilst ensuring that the pairings assigned to the reserve crew after their reserve blocks (follow on pairings) suffer no knock-on effects.

In this thesis, reserve crew are scheduled in ways that reflect the expected demand for reserve crew due to day of operation crew absence and delay disruptions. The use of reserve crew to cover for open pairings is not explicitly considered in this thesis. This is because the primary role of reserve crew is for use covering unexpected crew-related disruptions. The use of reserve crew to cover open pairings is only considered after they have been used to cover for crew-related disruptions. The open pairings are used to fill the remainder of reserve blocks provided that the minimum rest requirements of reserve are satisfied.

Due to KLM's reserve block structure, reserve crew are on standby for 5 days unless they are used to cover a disrupted crew pairing. As a result, the only constraints for feasible reserve crew use are: 1) a reserve must be on a standby duty day when a crew related disruption occurs; and 2) the expected finish time of the initial disrupted crew duty must be within the standby duty length of the reserve crew member(s).

Crew at KLM come in a variety of rank and qualification combinations known as divisions. Qualifications determine the set of fleets crew can operate on, typically crew are qualified for between 1 and 3 different fleet types. Ranks, determine which roles crew can undertake on a flight. At KLM there are 3 main ranks: pursers; 2 band; and 1 band. In general all flights require at least one purser. At KLM "flying above rank" is permitted on the day of operations (but cannot be planned in scheduling), but only for the cases of pursers flying as senior pursers and 1 band (lowest rank) flying as 2 band (second rank). The cost is that they have to be paid at the higher rank rate. On the other hand, "flying below rank", is allowed provided that separate qualified crew are allocated to each role, the cost is that the crew still have to be paid according to their designated rank. In this thesis reserve "flying above rank" is not considered, because a more detailed model would be required to model the exceptional circumstances in which this is permitted. "Flying below rank" is explicitly modelled in this thesis (see Sections 6.3 and 9.8).

Each fleet type has a minimum required number of qualified crew of each rank. In general, the crew requirements per fleet are proportional to the passenger capacity of the aircraft of that fleet type. If the number of passengers is below a certain threshold, the minimum crew requirements can sometimes be relaxed on the day of operation and "flying minus one" is allowed. In Chapter 10 the test instances are based on the case where there are 3 fleet types and 2 crew ranks. Crew are each qualified for a separate combination of 2 fleet types. Each fleet type has distinct crewing requirements, in terms of the required number of crew of each rank.

At KLM, a fixed number of new reserve blocks are available for scheduling per day. This thesis considers the case where a fixed number of reserve crew are available for scheduling within a specified time horizon. Additionally, KLM have fixed salaried staff this means that reserve crew costs are not a variable, thus costs do not form any part of the objectives considered in this thesis. So, in this thesis the main objectives are always related to disruption minimisation with respect to fixed reserve crew availability.

3.5 Definitions

This section introduces a number of definitions that recur across the remaining chapters of this thesis.

3.5.1 Cancellation measure of a delay

Crew related disruptions include absent crew and crew delayed on connecting flights. When crew are absent and no reserve crew are available it may be necessary to cancel a flight. In the event of a crew-related delay the consequences of not absorbing the delay may not be as extreme, as it may be feasible to simply operate a delayed schedule. In terms of minimising disruptions there are two separate objectives to consider, cancellation minimisation and delay minimisation. In order to keep the simplicity of a single objective optimisation problem, Equation 3.3 is used to map delays to a measure of cancellation.

$$cancellation\ measure = \left(\frac{delay}{cancellation\ threshold}\right)^{delay\ exponent} \tag{3.3}$$

The overall objective is then to minimise the sum of the expected cancellations and the cancellation measures of the expected delays. Equation 3.3 is designed to capture the subjective equality between a cancellation and delays of different sizes. The *delay exponent* is a parameter controlled by the decision maker. The value of the *delay exponent* will typically be greater than 1 to capture the perception that a cancellation is worse than any number of delays whose delays sum to the *cancellation threshold*, and that small delays are much less important than long delays. The *cancellation threshold* is the assumed maximum delay before a flight is cancelled, a value of 3 hours is assumed in this thesis (see Section 2.4.2). Equation 3.3 is used when the probabilistic crew absence model takes reserve-induced delay into account (Chapter 6), in the statistical delay propagation model (Chapter 8) and in the mixed integer programming simulation scenario model (Chapter 9) to penalise delays in general.

Figure 3.2 shows the effect of varying the *delay exponent* on the cancellation and delay performance of reserve crew schedules derived using the delay cancellation measure function as part of the objective function (the other part being the expected cancellation due to crew absence). Small values of the *delay exponent* lead to low delays at the expense of increased cancellations (due to uncovered crew absence) and high values of the *delay exponent* lead to higher delays and reduced cancellations (due to uncovered crew absence). The value of the *delay exponent* is therefore a decision maker parameter for selecting one reserve crew schedule from a set of trade-off solutions. Hereafter, a *delay exponent* of 2 is assumed. A *delay exponent* of 2 corresponds to giving sub-linear weight to delays below the *cancellation threshold*, but giving them increasing importance as they approach the *cancellation threshold*. The proposed methods in this thesis still work for any value of the *delay exponent*.

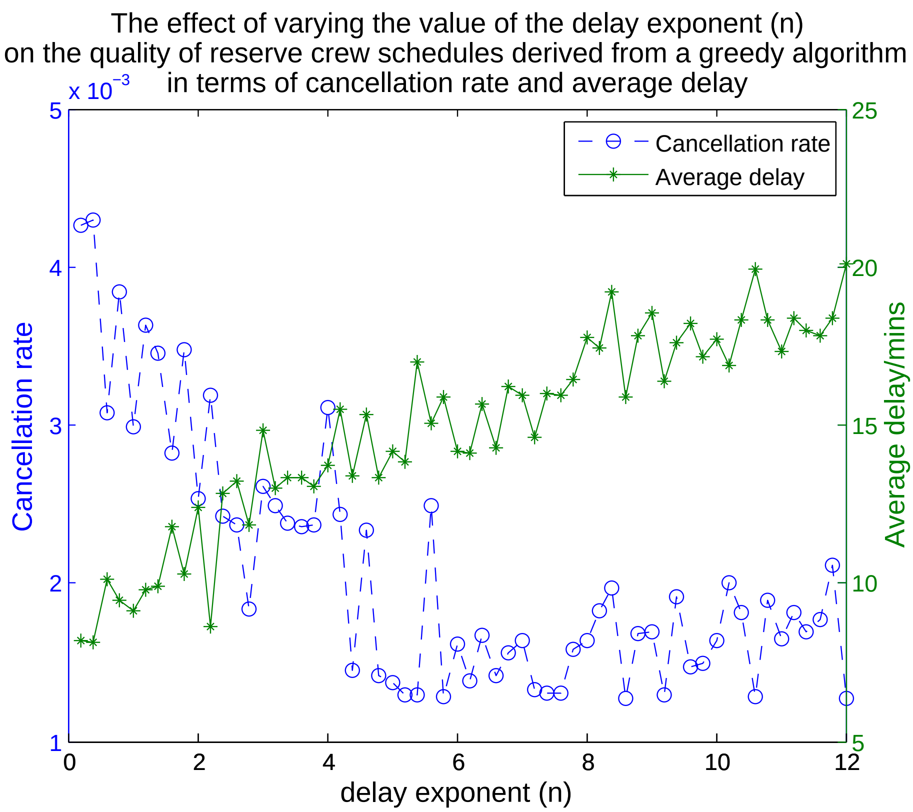

Figure 3.2: The effect of the delay exponent (used to penalise delay) on average delay and cancellation rate

3.5.2 Rule of thumb reserve policies

Default reserve policy

As described in Section 3.3 reserve policies are treated as the secondary objective in this thesis. This section defines the default reserve policy. The default reserve policy is the assumed policy for reserve use when no other reserve policy is defined.

The default reserve policy is a rule of thumb reserve policy, in which reserve crew are used as demand occurs, using reserve crew in earliest start time order first. The default policy never holds reserve crew in the event that their use is immediately beneficial. The default reserve policy is used in Chapters 5 and 7 when the probabilistic models are first introduced, where the main focus is on reserve crew scheduling. Chapters 4, 6, 8 and 9 consider improved reserve policies.

Absence only policy

Another rule of thumb policy used in subsequent chapters is called the *abs only* policy. In the *abs only* policy the default policy is used whenever crew are absent, but when crew related delays occur, reserve crew are never used to replaced the delayed crew.

3.5.3 Heuristic reserve crew scheduling approaches

Heuristic or "rule of thumb" reserve crew scheduling approaches provide initial benchmark solutions when developing new more advanced approaches, the more advanced approaches must at the very least result in reserve crew schedules of higher quality schedules than the rule of thumb approaches. The only exception to this is when a rule of thumb approach is known to result in an optimal solution.

Zeros

The *Zeros* rule of thumb approach to reserve crew scheduling consists of scheduling all available reserve crew to begin their standby duties at the start of the first day of the considered scheduling time horizon.

Uniform start rate

The *uniform start rate* (or USR) rule of thumb approach to reserve crew scheduling schedules the reserve crew of each type to begin standby duties at equal time intervals over the considered time horizon.

No reserves

The *no reserves* (or No Res) rule of thumb approach to reserve crew scheduling schedules no reserve crew at all. This approach is useful when assessing the total impact that scheduling any reserve crew at all has on the expected level of disruption in an airline schedule.

3.5.4 Frequently used solution methodologies

This section describes the search methodologies that are used frequently in this thesis. Subsequent chapters will refer back to these explanations. The probabilistic models of Chapters 5 to 8 provide a means of evaluating possible reserve crew schedules in terms of their associated delay and cancellation minimisation effects. The probabilistic models can be used as the evaluation function in a variety of heuristic solution methodologies. These include the following.

Greedy algorithm

A greedy algorithm, when applied to the problem of reserve crew scheduling, adds one reserve crew to the schedule at a time, at the time that results in the largest decrease in the objective value.

Local search

Local search (steepest ascent) proceeds by evaluating all solutions neighbouring the incumbent solution and moving to the solution that decreases the objective value the most. This process continues until no improving moves are available in the local neighbourhood, i.e. a local optimum has been found.

Tabu search

Tabu search [42] is a local search based approach which avoids becoming trapped in local optima by allowing moves to non-improving neighbouring solutions. To avoid cycling between local optima, an adaptive memory is used to store solutions which have been visited recently. Once a move to a neighbouring solution is made, the new solution is added to a list of "tabu" moves which cannot be visited for a specified number of iterations. The number of iterations a solution remains in the tabu list is sometimes known as "tabu tenure". After solutions have served their tenure, they are removed from the tabu list and can be visited again. However, the intention is that the search will have moved to a different region possibly containing the global optimum. See Section 2.6.2 for more details.

Simulated annealing

Simulated annealing [54] (also see Section 2.6.2) is an approach to searching for solutions to optimisation problems which is based on the analogy of cooling molten substances to achieve orderly final states. The simulated annealing algorithm for optimisation problems requires the specification of an initial temperature (T_0), a cooling scheme ($T(n)$) and the total number of iterations ($maxN$) before the algorithm terminates, where n is the current iteration number. The algorithm starts from a given initial solution, each iteration randomly selects a neighbouring solution, which is accepted if it is an improving move. A non-improving move is accepted with probability $e^{-\delta/T(n)}$. δ is the increase in the objective value associated with the non-improving move, $T(n)$ is the current temperature. In this thesis, a cut and insert neighbourhood is used, this corresponds to taking a reserve crew member and changing their start time. Also, the cooling scheme (value of T at any given iteration) is based on an exponential decay, starting from a specified initial temperature (T_0) and reaching a final temperature (T_{maxN}) after $maxN$ iterations or a maximum time limit.

For the problem of reserve scheduling, the maximum change in the objective value ($max\delta$) that can occur by moving to a neighbouring solution, is approximately equal to the maximum number of hub departures in a crew pairing. I.e. the maximum number of cancellations that may result from making reserve crew unavailable for that pairing. When $max\delta$ is used as the initial temperature, it can be referred to as the boiling point. This approach has the effect of only accepting the worst possible move to a neighbouring solution at the beginning of the algorithm.

Genetic algorithm

Section 2.6.2 and [43] explain the basic principles behind genetic algorithms. In this thesis genetic algorithms are applied In Chapters 5, 6 and 10. In Chapter 5 when a genetic algorithm is used to search for cancellation minimising reserve crew schedule the crossover operation which is used leads to infeasible solutions which are corrected using heuristics. In Chapters 6 and 10 a different solution representation is used with which this problem does not occur.

In Chapters 6 and 10 the mutation operator is replaced with a single iteration of a simulated annealing algorithm, the cooling scheme of which therefore controls the mutation rate over the course of the algorithm. The use of simulated annealing as the mutation operation makes the algorithm similar to a memetic algorithm (described in Section 2.6.2).

3.6 Chapter summary

This chapter has given a general formulation for the problem of reserve crew scheduling and reserve policy selection under operational uncertainty. The main elements to consider when addressing this problem, from both offline scheduling and online policy perspectives, were discussed. Specific details were also given for the structure of the real world case study which is to be considered in this thesis. The chapter finished with a section defining recurring concepts that do no belong to any one of the subsequent chapters.

Chapter 4

Simulation model and methodologies

This chapter introduces a simulation framework for a single hub airline. The simulation allows the study of possible reserve crew scheduling solutions in the full range of possible disruption scenarios. In a single hub and spoke network there are many sources of uncertainty influencing the way in which an airline schedule unfolds on the day of operation. Crew can be absent, flights can arrive late at the hub station (which may also delay subsequent flights), airport congestion can cause delays and aircraft may require unscheduled maintenance. Additionally, there are many possible ways in which a disrupted schedule can be recovered, for instance, flights may be cancelled, airline resources may be swapped to mitigate delays and reserve crew may be used to cover for absent or delayed crew. In short, a combinatorial explosion arises due to the many sources of uncertainty in an aviation network and the way that recovery decisions influence the possible future outcomes of scheduled events.

The continuous interaction between operational uncertainty and airline recovery makes it difficult to theoretically derive statistical distributions for the outcomes of scheduled events (although Chapter 8 attempts this). This task is simplified by considering a single realisation of a sequence of events at a time and repeating this process to derive the required distributions. A simulation is the ideal tool for this purpose. Section 2.5.3 described several existing airline simulators which were used to address airline scheduling problems under uncertainty.

The development of this simulation tool was initially motivated by the desire to develop a scenario-based approach (see Chapter 9) to reserve crew scheduling, as an alternative to the probabilistic model approaches (Chapters 5 to 8), where the simulation's purpose was to generate the required input scenarios. The simulation tool introduced in this chapter has also proven to be an invaluable research tool in its own right and is used in many of the subsequent chapters of this thesis.

Chapter structure

In this chapter, Section 4.1 describes the goals of the simulation. Section 4.2 introduces the assumptions and assertions which underpin the simulation. Section 4.3 illustrates a high level view of a single run of the simulation. Section 4.4 describes the modelling of airline recovery. Sections 4.5 and 4.6 describe the simulation inputs and outputs. Section 4.7 describes the reserve crew scheduling and reserve policy methodologies that are based on this simulation.

4.1 Goals

The goals of the simulation include:

- **To help to develop intuition of the problem.** It can be quite difficult to gain intuition about a problem characterised by the continuous interaction of uncertainty and decision making. A simulation tool can be used to form a research cycle of testing, feedback and refinement. Simulation provides immediate feedback on the performance of new approaches. Simulation visualisation also provides an additional method of testing (debugging) the implementation of experimental approaches.

- **To derive input parameters for solution methodologies.** Simulation can be used to derive input data for various solution methodologies. In Chapter 7 simulation is used to derive probabilities of delay propagation. In Section 9.2 simulation is used to generate disruption scenarios for scenario-based approaches to reserve crew scheduling. In Section 4.7 a simulation based reserve crew scheduling approach is described which learns when demand for reserve crew occurs and then schedules reserve crew to meet those demands.

- **To compare and validate alternative approaches to reserve scheduling and online decision making.** The simulation is also a valuable tool for comparing alternative approaches to reserve crew scheduling and online decision making and for determining what the strengths and weaknesses of the various approaches are.

4.2 Simulation assumptions and assertions

The simulation consists of a number of distinct components including: the airline's resources (crew and aircraft); the assignments of those resources (lines of flight); the airline's recovery policy; the possible recovery actions; journey time uncertainty; crew absence uncertainty; the flow of information; and the airline's schedule. For each of these elements of the simulation a number of assertions and assumptions are made. The assumptions fall into the categories of: Simplifying (S); True (T) (or intrinsic to the problem); or KLM specific (K). Intuitive names are also given for the simplifying assumptions (S) in italics.

Crew

The assumptions made regarding airline flight crew are as follows:

- **C1:** (T) Each flight leg requires a team of crew.
- **C2:** (T) A team of crew consists of a number of crew with a range of ranks. The required number of crew of each rank is determined by the fleet type (see assumption **A2**).
- **C3:** (T) Crew have qualifications that determine the fleet types they can operate on.
- **C4:** (S) *Crew homogeneity assumption.* The crew in a team with the same rank and qualifications are treated as homogeneous.
- **C5:** (S) *Binary crew absence assumption.* Absent crew are unavailable for an entire crew pairing. In reality, if the crew become available after the beginning of their assigned pairing they are used for other open assignments (see Section 2.3).
- **C6:** (T) Crew have specified duty start times for each day of their assigned crew pairing.
- **C7:** (S) *Fixed duty length assumption.* Crew are limited to a fixed maximum daily duty length (typically 12 hours).
- **C8:** (T) Crew have minimum rest periods between consecutive flights. If a crew stays on the same aircraft for consecutive flights, the minimum rest time, takes on the value of the scheduled ground time, that is, if the scheduled ground time is less than the minimum rest time.
- **C9:** Crew are swappable if:
  - **C9a:** (S) the crew teams are assigned to duties involving the same fleet type. *Fleet purity of crew pairings assumption.*
  - **C9b:** (T) the crew teams can complete each others duties within their respective maximum duty lengths.
  - **C9c:** (T) the crew teams are assigned to crew pairings (crew lines of flight) that are swappable (see lines of flight below).
- **C10:** (S) *Indivisible crew teams assumption.* Crew teams stay together throughout the crew pairings that they are assigned to.

Aircraft

The assumptions made regarding the airlines airframes are as follows.

- **A1:** (T) Aircraft come in a range of fleets.
- **A2:** (T) The fleet type determines the number of qualified crew of each rank required for the legal operation of a flight involving the given fleet.

- **A3:** (T) Aircraft have minimum turn times between consecutive flights.

- **A4:** (S) *Fleet homogeneity assumption.* Aircraft are swappable if the two aircraft are of the same fleet type.

Lines of flight (crew pairings or aircraft routings)

The assumptions made regarding (crew or aircraft) lines of flight are as follows.

- **L1:** (T) Flight duties are sequences of flight legs that can legally be performed in a single day or shift by an airline resource (crew or aircraft). Lines of flight consist of a sequence of flight duties which can be assigned to a single airline resource (crew or aircraft).

- **L2:** (T) Lines of flight have an associated fleet, which in the case of aircraft routings, determines the fleet type required to operate the line of flight, and in the case of a crew pairing determines the required number of qualified crew of each rank.

- **L3:** (K) Lines of flight are fleet pure. I.e. all flights on a line of flight involve the same fleet type.

- **L4:** Lines of flight are swappable if:

  - **L4a:** (S) they are associated with the same fleet type. *Same fleet swaps only assumption.*
  - **L4b:** (S) they share the same overnight station where the swap can be reversed. *Same overnight station swap assumption.*
  - **L4c:** (S) the next scheduled flight on the replacement resource's line of flight is later than the scheduled departure time of the delayed line of flight for which recovery is sought. Otherwise the swap can only increase overall delay, because a delayed resource will delay an earlier flight even more (This assumption covers the vast majority of beneficial resource swaps, See Section 8.1.4 and theorem 1 in Appendix C). *Later flights are delayed less by delayed resources assumption.*
  - **L4d:** (S) the replacement resource for the delayed line of flight is not delayed for its own next scheduled flight. Otherwise the swap just redistributes the total delay without reducing it. *The replacement resources must not be delayed for their own next scheduled flight assumption.*

Airline recovery policy

The assumptions made regarding the airline recovery policy are as follows.

- **RP1:** (S) *Sequential recovery assumption.* Airline recovery actions are made at the scheduled departure time of a flight, at that time it is assumed that sufficiently accurate arrival time estimates will be available for the flights which may provide swap opportunities (see assumption **I1** below).

- **RP2:** (S) *Delay threshold assumption.* Departures delayed by more than 15 minutes are considered delayed. Recovery actions are considered for such delayed departures.

- **RP3:** (S) *Cancellation threshold assumption.* Departures delayed by more than 180 minutes (the cancellation threshold) after the application of recovery actions, are cancelled.

- **RP4:** (S) *Crew absence cancellation assumption.* Flights are cancelled if absent crew from the assigned crew team cannot be replaced with reserve crew. The exception to this is when "flying minus one" is allowed, but this depends on passenger numbers being below a certain threshold level.

- **RP5:** (S) *Reserve policy assumption.* The reserve policy determines which reserve crew should be used to cover a given absence or crew-related delay. The reserve policy also has the final say on whether or not reserve crew should be used then or held for later use. A reserve policy is a useful concept for thinking about how reserve crew are used.

- **RP6:** (S) *Deadheading not viable for solving short notice delays and unexpected crew absence disruptions assumption.* Disruptions that occur at spoke stations are solved at those spoke stations. Namely that crew absence that effects crew stationed at spoke stations are covered using reserve crew stationed at that spoke station, which makes for a trivial reserve crew scheduling problem due to low flight volume. The exception to this assumption is when reserve crew are deadheaded to a spoke, but this is rarely useful for the types of disruptions which are the focus of this work.

- **RP7:** (S) *Low spoke station flight volume assumption.* Swap recovery actions are not available at spoke stations due to low flight volume, and therefore any delays simply propagate back to the hub station via the return leg.

Recovery actions

The assumptions made regarding airline recovery actions are as follows.

- **RA1:** (T) Delays can be absorbed by swapping crew and/or aircraft.

- **RA2:** (S) *Delay reducing swaps only assumption.* Delay reducing swap recovery actions must not cause an increased delay of another flight. See Theorem 1, Appendix C

- **RA3:** (T) If no delay reducing recovery actions are available but the flight is still feasible, operate a delayed schedule (i.e. do nothing, and accept the delay).
- **RA4:** (T) Reserve crew can be used to cover for absent or delayed crew.
- **RA5:** (T) Flights that are still infeasible after the after the consideration of recovery actions are cancelled. For a single hub network it is usually only a single out and back cycle that is cancelled (a cancellation cycle, see Section 2.2.1), the line of flight resumes (if then feasible) at the next scheduled hub departure.
- **RA6:** (S) *Swap recovery selection assumption.* After the consideration of delay recovery actions, select the recovery action which minimises the delay of the disrupted flight. Ties are broken by selecting the recovery action involving the least number of changes from the original schedule. Where the degree of schedule changes is quantified according to the product of the number of individual crew or aircraft swaps and the number of other lines of flight directly affected by the swap recovery action. After this, remaining ties are then broken by selecting the swap with the largest amount of common ground time in which the swap recovery action can be undone.

Reserve crew

The assumptions made regarding reserve crew are as follows.

- **RC1:** (T) Reserve crew come in a variety of rank and qualification combinations, which must be respected when using reserve crew.
- **RC2:** (S) *Fly below rank assumption.* Reserve crew can fly below rank, but not above (flying above rank, although possible, is not considered because it is much less common).
- **RC3:** (S) *Reserve block regularity assumption.* Reserve crew standby duties are regular, in the sense that they begin at the same time each day, within a reserve block.
- **RC4:** (T) The earliest start time of a grouping of reserve crew is the maximum of the start times of those reserve crew. The maximum duty finish time of a grouping of reserve crew is the minimum of the duty finish times of those reserve crew.
- **RC5:** (S) *Reserve crew use assumption.* Reserve crew can be used to replace delayed or absent crew provided that this does not result in a departure delay exceeding the cancellation threshold.
- **RC6:** (S) *Reserve duty time feasibility assumption.* Reserve crew can be used if the adopted line of flight is scheduled to finish on or before the reserve crew's maximum duty finish time.

- **RC7:** (S) *Reserve crew domicile assumption.* Reserve crew are stationed at the hub station.

- **RC8:** (S) *Reserve crew response time assumption.* Reserve crew have a zero response time.

- **RC9:** (S) *Once only reserve crew use assumption.* Once reserve crew have been used once to replace absent or delayed crew they adopt the disrupted crew pairing, and do not return to being standby reserve crew.

Journey time uncertainty

- **J1:** (S) *Journey time uncertainty assumption.* For each origin-destination pair there is a statistical distribution which captures journey time uncertainty. Although not considered in this thesis, the natural extension of this is that different distributions apply at different times of the day (due to congestion) and also in different weather conditions.

Crew absence uncertainty

- **CA1:** (S) *Crew absence independence assumption.* Each individual crew member has an independent probability of being absent. For each crew team the simulation requires a cumulative probability distribution of different numbers of crew of each rank being simultaneously absent. Note that the probabilities of crew absence may be affected by factors such as seasonal trends in contagious diseases. Although this is not explicitly considered in this thesis, different distributions could be used to reflect such trends.

Information flow

- **I1:** (S) *Perfect ETA knowledge assumption.* The uncertain arrival time of a flight is known with perfect accuracy at all times after the departure of the flight. The justification for this assumption is that the knowledge of arrival times will typically only be required at times close to the scheduled arrival time of those flights, so it is reasonable to assume that very accurate estimates of arrival times will exist at those times. Although not considered in this thesis, the alternative is to model arrival time distributions shrinking in width towards randomly generated journey times whilst journeys are in progress.

Schedule

- **S1:** (T) All flight legs are planned with specific origins, destinations, departure times, arrival times, assigned crew and aircraft.

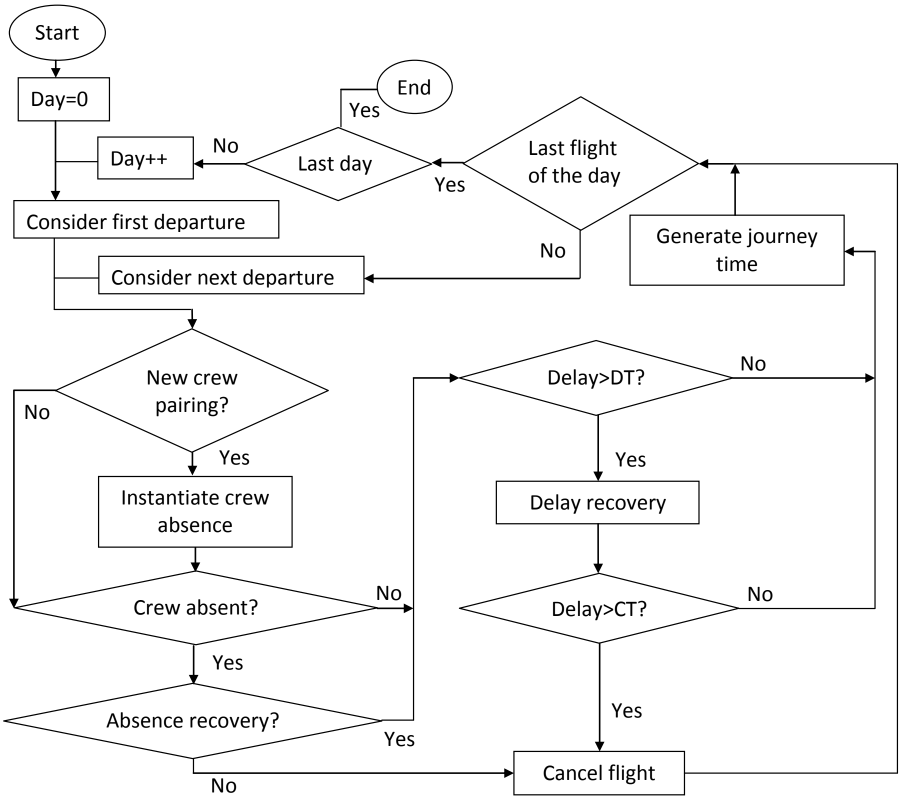

Figure 4.1: Simulation flow chart

4.3 Simulation flow chart

Figure 4.1 shows a flow chart corresponding to a single run of the simulation. The simulation considers each departure in departure time order. If the given departure corresponds to the start of a new crew pairing, the number of absent crew affecting that crew pairing is stochastically generated from the corresponding statistical distribution. If crew are absent, and reserve crew are available, they are used to replace the absent crew, otherwise the flight has to be cancelled. The simulation then computes the earliest departure time based on the previous arrival times of the scheduled resources. If the departure is not delayed beyond the delay threshold the flight goes ahead and a journey time is generated from the relevant journey time distribution. If the flight is delayed beyond the delay threshold, recovery actions are considered. If, after the implementation of recovery actions, the delay exceeds the cancellation threshold, the flight is cancelled. Otherwise the flight goes ahead with a journey time generated from the relevant distribution. Swaps are undone at the end of each day (not shown in the flowchart). The simulation continues in this fashion until the whole schedule has been implemented.

4.4 Airline recovery

This section concerns delay recovery and absence recovery (see Figure 4.1). In the event of a delay exceeding the delay threshold, the simulation considers all combinations of single crew and aircraft swaps. Additionally, teams of reserve crew are considered in conjunction with aircraft swaps, the delay recovery action that minimises the delay without increasing the delay of another flight is preferred (assumption **RA6**, the *swap recovery selection assumption*). The assumed airline recovery policy is defined by the assumptions of Section 4.2. Assumptions **C9** (crew swap assumptions) and **A4** (*fleet homogeneity assumption*) define which crew and aircraft are swappable. Assumption **L4** (swappable line of flight assumptions) defines the condition under which lines of flight are swappable. Assumptions **RP1-4** specify when recovery actions are required and the order in which disrupted flights are recovered. Assumptions **RA1-6** define what the available recovery actions are. Assumptions **RC1-6** define the feasibility of reserve crew in the event of crew-related disruptions.

4.4.1 Delay recovery

As stated by Assumption **RP2** (*delay threshold assumption*), delay recovery is considered if a departure h is delayed beyond the delay threshold, in which case all combinations of single crew and single aircraft swaps are considered. As in Assumption **RA2** (*delay reducing swaps only assumption*), a swap recovery action must reduce the delay of departure h, without invoking additional delay on the departures directly affected by the swap. In Appendix C this is shown to be a necessary condition for a beneficial resource swap for the case of a single resource swap and a sufficient condition for simultaneous crew and aircraft swaps.

The earliest departure time of a flight is a function of the earliest crew and aircraft ready times, therefore the effect of a crew swap cannot be appreciated in isolation from the effects of aircraft swaps. As a result, all combinations of single crew and single aircraft swaps have to be enumerated and evaluated in terms of their associated delays. To avoid enumerating all combinations of single crew and aircraft swaps, assumptions **C9** (crew swap assumptions), **A4** (*fleet homogeneity assumption*) and **L4** (swappable lines of flight assumptions) reduce this to considering only the swappable resources if they are assigned to swappable lines of flight. This leaves a set of feasible crew assigned to swappable crew pairings and feasible aircraft assigned to swappable aircraft routings. Furthermore, individual resources can be ruled out if their ETAs—from previous flights—are greater than the ETA of the delayed resource assigned to departure h, or greater than the cancellation threshold (CT) of departure h. If a single resource satisfies all of these requirements, a feasible swap time window is calculated for the resource (denoted CTW and ATW for crew and aircraft respectively). If the time window has non-negative width, the resource is added to the list of feasible resources (FC and FA for crew and aircraft respectively) to be considered as possible combinations of single crew and single aircraft swaps.

The feasible time window for crew k is as follows.

$$CTW_{lb}^{k} = \max\{D_h,\, ceta_k + MS\} \tag{4.1}$$
$$CTW_{ub}^{k} = \min\{ceta_{C_h} + MS,\, D_h + CT\} \tag{4.2}$$

MS denotes the minimum (sit) rest time for crew between consecutive flights, TT the minimum turn time for aircraft between consecutive flights, D_h the departure time of flight h, C_h the crew assigned to flight h, and $ceta_k$ the ETA of the k^{th} swap feasible crew. Equivalent expressions exist for feasible aircraft swap time windows. After this, all combinations of single crew and single aircraft swaps are considered. Feasible combinations have crew and aircraft feasible swap time windows that overlap. They will also have the property that the replaced resources will not increase the delay of the next departures on the lines of flight they adopt. Assumption **RA6** (*swap recovery selection assumption*) defines the preference order for selecting a recovery action.

Algorithm 1 Pseudocode finding a delay minimising swap recovery action

for $\forall l \in FC$ **do**
  for $\forall m \in FA$ **do**
    if an overlap exists between the feasible swap time windows CTW^{FC_l} and ATW^{FA_m} and the swap does not increase the delay of other affected flights **then**
      if $\max\{CTW_{lb}^{FC_l},\, ATW_{lb}^{FA_m}\}$ < best alternative departure time **then**
        Update best swap recovery action
        Update best alternative departure time
      else
        if $\max\{CTW_{lb}^{FC_l},\, ATW_{lb}^{FA_m}\} \equiv$ best alternative departure time **then**
          if this swap is easier to implement and undo later than the current best swap recovery action **then**
            Update best swap recovery action
            Update best alternative departure time
          end if
        end if
      end if
    end if
  end for
end for

Algorithm 1 shows how the delay minimising swap recovery action is calculated from the list of feasible crew and aircraft and their associated feasible swap time windows.

4.4.2 Reserve teams constructed and used to replace delayed connecting crew

After the consideration of all feasible combinations of single crew and single aircraft swaps, departure h may still be delayed. If the flight is delayed due to a delayed crew, delay can be absorbed by constructing a replacement team of crew out of individual reserve crew, provided that the generated reserve team is feasible and reduces the delay (assumptions **RC1-6**).

If the reserve policy permits the use of a reserve team for this purpose, the replaced delayed crew can potentially be used for swaps later on. A swap with a reserve team is modelled as the delayed crew adopting an empty line of flight.

4.4.3 Absence recovery

In the event that some crew are absent from a team of crew, flights on the affected crew pairing are delayed or even cancelled until reserve crew become available to replace the absent crew. Reserve crew use is the only recovery action that can avoid flight cancellation. In this thesis a number of reserve policies allow for the possibility of holding back reserve crew in the event of crew absence. The potential benefits of reserve holding include: the possibility that a more damaging instance of crew absence could be covered instead; the possibility that a given instance of crew absence uses too many reserve crew which could have been used to cover other disruptions for greater overall reward; and also that using reserve crew may induce a large amount of delay which may have a high probability of propagating.

4.5 Simulation input

The simulation requires an input schedule and statistical distributions corresponding to the uncertain elements of the simulation.

4.5.1 Input schedules

The required schedule inputs for the simulation are the set of flight legs complete with scheduled departure times, arrival times, origins, destinations and the crew and aircraft assigned to each flight. The schedules used with this simulation are mostly based on real data (Chapter 5 is the only exception). The real schedule data is based on KLM's operations in August 2012. To avoid creating methods that are specialised for one particular example schedule, a random schedule generator has been created. The schedule generator has parameters that control different aspects of schedules to make them more or less challenging in terms of maintaining crew feasibility. The tightness of generated schedules can be controlled using a parameter which controls the probability that journeys can be completed within their allocated time windows. The occurrence of crew-related delays can be controlled by varying the rate at which crew change aircraft during duties. Such schedules are used in Chapters 6 to 9 during the development of several of the approaches to reserve crew scheduling.

When considering the case of multiple fleets, crew ranks and qualifications (Sections 6.3, 9.8 and Chapter 10) real aircraft lines of flight were derived from the August 2012 data, and crew pairings were generated using a set partitioning model with CPLEX as the solver. Chapter 10 uses three real schedules of varying sizes to compare all approaches to reserve crew scheduling and online reserve decision making considered in this thesis. In addition, versions of each of these schedules are generated which have a heightened delay risk and minimum scheduled ground times. The purpose of these schedules is to highlight the potential of the models which aim to minimise delay propagation in instances where this is a significant issue. The schedules based on real data have minimal risk of delay propagation, since they include scheduled slack.

4.5.2 Stochastic elements of the simulation

The simulation requires input probability distributions for journey times and crew absence. The schedule data for KLM in August 2012 provides scheduled and actual departure and arrival times. This allowed the derivation of journey time distributions for each origin-destination pair. The simulation does not explicitly model ground delays due to airport congestion, as these are included as part of the journey time. Additionally, KLM also provided block time (gate to gate) distributions that they had used in one of their own simulation studies. This data was used in the schedule generator described above. KLM also provided historical data on crew absence for a 7 year period. From this, probabilities that any given crew member will be absent were approximated (see the crew absence uncertainty assumptions of Section 4.2).

4.6 Simulation output

Simulation is used to provide quick feedback about possible changes to a system without actually implementing the changes in reality. The output from this simulation includes the following.

- Schedule visualisation.
- Average performance measures for delays, cancellations, reserve crew demand and utilisation.
- Graphs regarding the types of recovery actions used for different disrupted flights.
- Animation of the paths of airline resources through the airline network.

Figure 4.2 shows the schedule visualisation output of the simulation. This is useful when checking that the input schedules are consistent and are free of errors. Green lines represent flights from the hub station to spoke stations. Yellow lines indicate flights from spoke stations to the hub station. Whilst line lengths indicate the planned duration of flight legs. Horizontal sequences of flight legs correspond to aircraft routings, whilst blue lines indicate crew pairings. I.e. non-horizontal blue lines show crew changing aircraft between consecutive flights. If an input schedule contains errors these may be apparent in the schedule visualisation as, for example, aircraft routings whose arrival time from one flight overlaps the departure time of the next flight or multiple crew teams connecting to the same flight

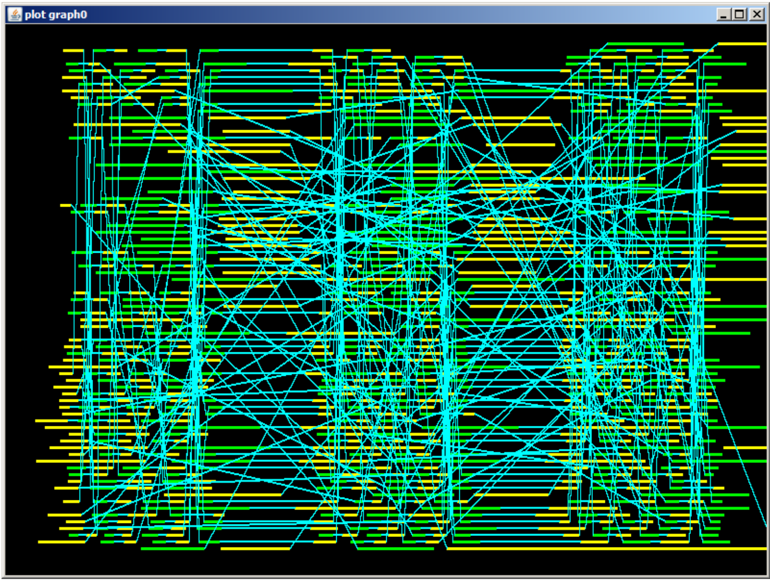

Figure 4.2: Airline schedule visualisation, the flight legs to the hub (yellow) or to a spoke (green) which are assigned to aircraft (rows) and crew (blue lines)

leg. The simulation visualisation gives clues as to what the source of the error may be.

The output performance measures from the simulation include: cancellation rate (overall and per fleet); reserve utilisation (overall, per reserve type and disruption type); delay (average total delay per flight, average crew delay per flight, average delay size, probability of delay (and probability of delay$\geq$30 minutes); crew swap rate; aircraft swap rate; cancellation measure (average and maximum). These performance measures are calculated from the average performance over a large number of repeat simulation runs. As well as average performance measures, an optional output is a list of the total accumulated cancellation measures from each individual run of the simulation, this information can be used to analyse the variance of the performance of a given reserve crew schedule and reserve policy combination (see Figure 9.6 for an example of this).

Figure 4.3 shows an example of the recovery action output from the simulation, for a 2 day schedule example. The graph shows how often various recovery actions were applied at each departure over a number of repeat simulation runs. The bars for each recovery action at each departure are stacked on top of each other. The use of reserve crew to cover for absent crew is uniform, due to the assumption that all crew have an equal chance of being absent. Delays typically become more common over the course of a day's operations, as subsequent flights may experience delays propagated from previous flights. As a result, delay recovery actions such as crew swaps and aircraft swaps occur more frequently at these times. Towards the end of the day large red spikes usually correspond to reserve crew being grouped into teams and used to replace delayed connecting crew. This happens at the end of the first day in the example given.

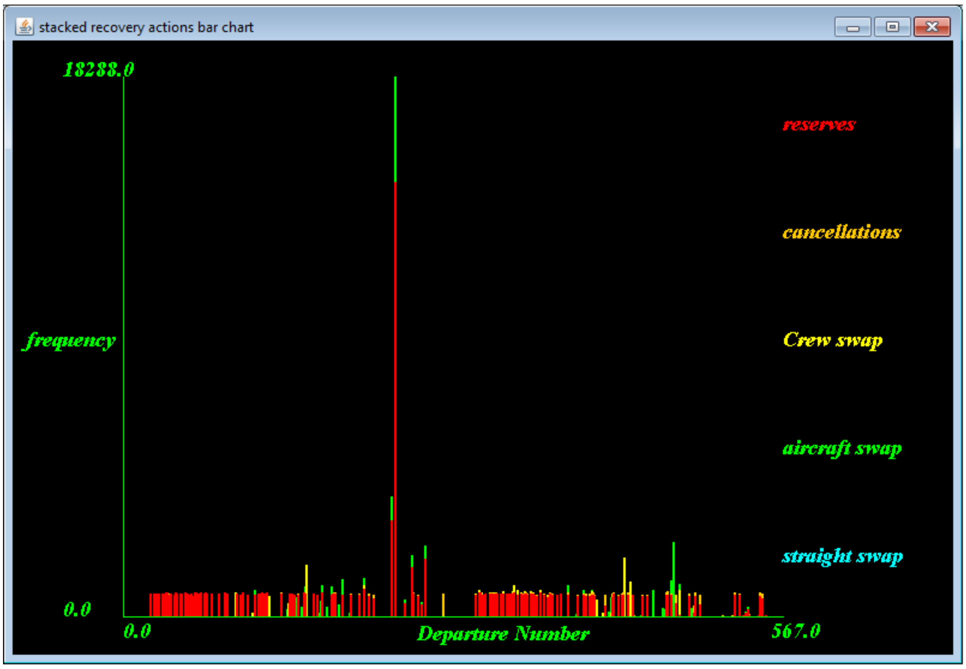

Figure 4.3: Example recovery actions graph

A visual simulation output such as that of Figure 4.3 can be used to provide immediate feedback on whether or not a proposed approach to reserve crew scheduling is effective or if there are any obvious weaknesses in a proposed solution.

4.7 Simulation based methodologies

One of the goals of this simulation is to provide data which can be used for reserve crew scheduling and reserve policy methodologies. In this section a simulation based reserve crew scheduling heuristic is introduced called the *area under the graph* approach. Then two simulation based reserve policies are described.

4.7.1 The *area under the graph* approach to reserve crew scheduling

The area under the graph approach to reserve crew scheduling is based on running a simulation without reserve crew as an available recovery action and keeping a record of the times at which reserve crew were in demand. After a large number of repeat simulations, the accumulated data can be used to schedule reserve crew at equal intervals of accumulated demand. The name *area under the graph* comes from an intuitive visualisation of this approach, in which the demand for reserve crew over time can be represented as a time series graph, and the process of scheduling reserve crew to satisfy demand evenly, can be visualised as scheduling reserve crew at equal intervals of area under the demand graph.

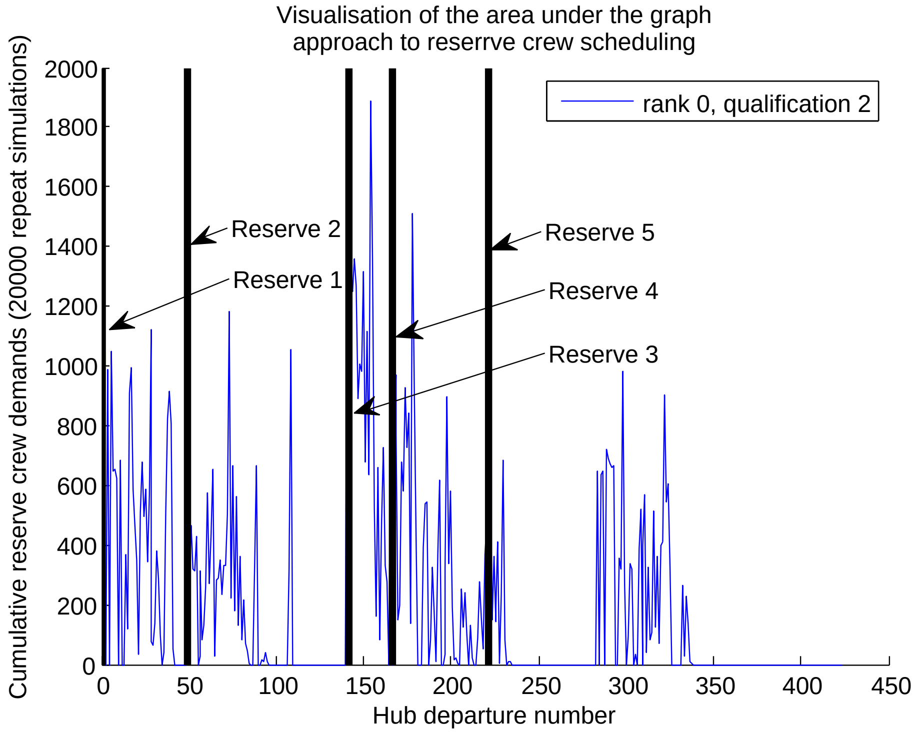

Figure 4.4: Reserve demand graph and equal total demand intervals

Figure 4.4 shows an example graph of the cumulative reserve demand derived from 20000 repeat simulation runs. The particular example shows the demand for reserve crew of low rank with qualification type 2, of which 5 reserve crew are available for scheduling. The black vertical lines mark equal intervals of total demand. These correspond to the times at which reserve crew are scheduled in this approach. More reserve crew will be allocated at times of the day when demand is high. The jump in the cumulative demand for reserve crew at around the 150^{th} hub departure of the day can be attributed to an increased rate of delayed connecting crew after the first round of flights of the day, as reserve crew can be used to replace delayed connecting crew.

In Chapters 5, 7, 9 and 10 the area under the graph approach is compared to various other approaches to reserve crew scheduling which were developed during this research.

4.7.2 Simulation reserve policy: SIM

The first simulation based reserve policy, referred to as SIM hereafter, runs a number of repeat simulations to evaluate and compare alternative reserve use decisions, such as holding reserve crew or using them to cover a given disruption. When this policy is implemented during a simulation which is being used to evaluate a given reserve crew schedule, a branch-off simulation is used to evaluate alternative reserve crew decisions in terms of the associated overall expected cancellation measure. The branch-off simulations that are used to evaluate reserve decisions require a rule based reserve policy, otherwise extra branch-off simulations will be recursively created, leading to computational intractability. The SIM policy is a realistic proposition in a real time setting, however when using the SIM policy in the validation simulation itself, the number of repeat simulations used to evaluate each alternative reserve decision has to be limited to control the amount of time taken to obtain results from the validation simulation. The cost of limiting the number of repeat simulations used to evaluate each alternative reserve decision is an increased risk that the limited sample of simulations is not representative of the true value of a given reserve decision. This is because confidence intervals are wider when they are derived from smaller samples. The SIM policy is used as a method of comparison in Chapter 10. In Appendix Section D.3 several variants of the SIM policy are also compared with a policy based on a probabilistic model and several rule of thumb heuristic policies.

4.7.3 Look up table policy: LUT

This section describes a look up table policy, where the table stores the expected values of states in terms of the total expected future cancellation measure (see Section 3.5.1) that will be accumulated thereafter if the given state is reached. The simulation is used to populate the look up table (LUT) with the values of states. A state is defined by the total number of reserve crew of each rank and qualification combination that remain available and the time/departure number at which those reserve crew are available. Such a look up table can be used to determine whether or not using reserve crew to cover a given disruption represents a globally optimal decision and if so which combination of reserve crew should be used. When a crew-related disruption occurs for which reserve crew use is at least an immediately beneficial action, the LUT can be used to determine the value of each of the available decisions in terms of the overall benefit. To evaluate a given decision (reserve use or holding reserve crew); firstly, find the state that decision leads to (i.e. the number of reserve crew of each type that will be remaining come the next scheduled departure), and secondly, obtain the value of that state from the LUT and then finally, add a value contribution corresponding to the immediate cost of implementing that decision (delay cancellation measure or cancellations directly caused by the decision). After evaluating all available actions, choose and implement that with the best value (i.e. that which minimises the total expected future cancellation measure).

Let Z denote the set of available reserve crew decisions, including reserve holding, let $S(t)$ denote the state at departure t. Let $T(z)$ denote the vector containing the number of reserve crew of each type which are used if decision z is implemented, and let $CM(z)$ denote the cancellation measure experienced by departure t if decision z is implemented. Let $V(S,t)$ denote the value of being in state S at departure t (the values stored in the LUT). Then the optimal reserve crew decision can be expressed by Objective 4.3:

$$\max_{z \in Z} V(S(t) - T(z), t+1) + CM(z) \tag{4.3}$$

Objective 4.3 determines the solution with the lowest total expected future cancellation measure, i.e the optimal decision, given no knowledge of future outcomes.

The underlying assumption of a LUT policy is that the values of states ($V(s,t)$) correspond to an optimal policy being followed thereafter. This means that the LUT has to be populated with values for all states corresponding to an optimal reserve policy. So, to determine an optimal policy, knowledge of the optimal policy is required in the first place (see Section 2.7.4), which makes this a challenging problem.

Instead, the LUT policy proposed in this section learns the values of states during a simulation learning phase in which the absence only policy (see Section 3.5.2) is used as the reserve policy. It is claimed here that the absence only policy approximates the behaviour of an optimal reserve policy. This claim is justified by the results of Section 10.5 and appendix Section D.3, which show that the best reserve policies nearly always cover crew absence disruptions but rarely cover crew-related delay disruptions. This approach to learning the value of states make this LUT policy an opportunistic reserve policy which can be used to find the exceptional situations in which the absence only policy should be overruled. In future work (see Chapter 12) the aim is to use approximate dynamic programming to learn a truly optimal policy from a simulation such as the simulation which is presented in this chapter. In Chapter 10 the LUT policy is compared to a variety of other reserve policies developed during this project in a number of realistic problem instances.

4.8 Chapter summary

This chapter has introduced the simulation tool that will be used for a variety of different purposes in this thesis, from parameter generation to the validation and comparison of the approaches proposed herein. The assumptions and assertions of the simulation were introduced. Details of the simulation implementation were given, especially regarding the modelling of airline swap recovery actions, which is a recovery action that has the potential for reducing reserve crew demand. The simulation's input requirements and it's outputs were discussed. Finally, several simulation based methodologies for reserve crew scheduling and reserve policies were introduced.

Chapter 5

Probabilistic crew absence model

The aim of this chapter is to model, in the simplest way possible, the problem of crew related disruptions and reserve crew used to cover for them, including only those features which characterise the underlying problem, whilst ignoring the features that are considered to be non-complicating details. This chapter introduces a model for crew absence disruptions and reserve crew used to cover for them. Crew absence disruptions are modelled as having probabilities of occurring and reserve crew are modelled as having probabilities of remaining available. The model is able to calculate the effect a given reserve crew schedule has on the probabilities that crew absence disruptions go uncovered, given the initial probabilities that they occur. The purpose of the model is for use as an evaluator of reserve crew schedules, which can be used in a range of methodologies to search for reserve crew schedules which offer the greatest level of protection against operational infeasibility. The search methodologies in which the proposed simplified probabilistic crew absence model ($SPCAM$) is used as an evaluator include: constructive/greedy algorithms; a variety of meta-heuristics; and a number of dynamic programming based heuristics. These search methodologies are compared to the optimal solutions for a number of small test instances.

The model presented in this chapter was inspired by the work of Paelinck, which was described in Section 2.3. In this chapter, reserve crew demand is modelled (in more detail) at a per flight level and the demand per flight need not be constant. The model also features a reserve use order policy that is modelled by the fundamental equations which are used for evaluating the effect that any given reserve crew schedule has on the probabilities that crew-related disruptions still go uncovered.

The $SPCAM$ introduced in this chapter underpins all of the subsequent probabilistic models presented in this thesis (see Chapters 6 to 8).

This chapter presents the first probabilistic approach to reserve crew scheduling considered in this project and corresponds to the first full conference paper [17] published during this research.

Chapter structure

The structure of this chapter is as follows. Section 5.1 describes the simplified problem. Section 5.2 introduces the $SPCAM$. Section 5.3 considers possible objective functions for the model. Section 5.4 investigates alternative solution methodologies to full enumeration. Section 5.5 outlines the areas for improvement for the $SPCAM$. Section 5.6 summarises the main findings from this chapter.

5.1 The simplified problem

The $SPCAM$ is aimed at early planning, just after the flight legs have been determined, when very little is known about crew disruptions on the day of operations. In its simplest form the crew unavailability problem can be modelled as a stream of departures from a single station (the hub). Each departure has a number of scheduled crew and each scheduled crew has some probability of absence. The end goal is to use the predictive model ($SPCAM$) that is developed in this chapter to schedule a fixed number of reserve crew, before any information is available about crew absence, to cover as much crew absence as possible. Whenever a crew absence occurs, it is assumed (in this simplified problem) that, reserves are used in earliest start time order (assumed to be the same as earliest finishing time order). Reserve crew have fixed standby duty start times and duty lengths, which determine which departures they can feasibly cover in the event of crew absence. The simplified problem requires the following assumptions. Some of these assumptions are themselves simplifications of the assumptions of Section 4.2, and are referred to below.

5.1.1 Assumptions

1. The reserve crew scheduling problem consists of a set of departures from a crew base where each departure has probabilities of crew absence and therefore the possible need for reserve crew. This assumption represents a simplification of assumptions **L1** (definition of flight duty) and **S1** (definition of an airline's schedule).

2. The maximum demand for crew per departure is 1, or this can be interpreted as a single team of reserve crew. This assumption represents a simplification of assumption **C2** (crew teams consist of a number of crew of a variety of ranks).

3. The chance of crew absence is captured accurately by unique probabilities for each departure. That is, the probability that a flight is affected by crew absence does not depend on the probabilities that other flights are affected by crew absence. This approach reflects the case where different crew are assigned to all flights. This assumption corresponds to assumption **CA1** (*crew absence independence assumption*), but has been reworded to fit the simplified problem that is considered in this chapter.

4. Reserve crew cover for the first crew absence that occurs within their fixed length duty period, if more than one reserve crew member is available then the crew member who started their duty period first is used. See assumption **RP5**, the *reserve policy assumption*.

5. Reserve crew can undertake a maximum of a single duty within any one duty period. This can be justified by the fact that, when reserve crew are used, they typically adopt the remainder of the pairing (string of flight duties) of the absent crew member they are covering for. See assumptions **RC1-6** (assumptions regarding reserve crew).

6. Reserve crew duties begin at times corresponding to scheduled departures. See assumption **RC3** (*reserve block regularity assumption*).

Assumptions 1 and 3 mean that the base problem can be represented as a vector of probabilities (Q) where q_i represents the probability that the crew scheduled for departure i will be absent. Assumption 2 means that the problem can be represented as a vector as opposed to a matrix for specifying the probabilities of different numbers of crew being absent at each departure (see Chapter 6 for the relaxation of assumption 2). Assumptions 4 and 5 state the way that reserve crew are to be used: reserve crew cover for the first disruption that occurs within their duty period, they can only cover one flight and therefore once they undertake a cover duty they cannot cover for any of the remaining departures that occur within their duty period. Assumptions 4 and 5 also define the precedence ordering for the use of reserve crew in the event that more than one reserve crew is available for a given crew absence disruption, which is to use the feasible reserve crew member who has been on duty the longest. Assumption 6 tries to minimise wasting reserve crew duty time by not scheduling them before the first time at which they may be required to cover crew absence. A consequence of assumptions 2 and 6 is that an optimal solution to the problem will involve no more than one reserve crew scheduled to each possible start time. This can be used to reduce the number of feasible solutions searched when trying to find a good or optimal solution to the problem.
Simplifying assumptions 1, 2 and 4 are removed in subsequent chapters. Table 5.1 defines the notation used in the model.

		Inputs
Q	:	Vector of probabilities. The probabilities that the originally assigned (regular) crew for each flight are absent. q_d is the probability that the crew assigned to flight d are absent.
L	:	Length of a reserve duty period.
R	:	Number of reserve crew available for scheduling.
n	:	Total number of departures in the airline's schedule during the reserve crew scheduling horizon.
		Decision variable
X	:	Vector of start time indices for standby duty reserve crew duties. x_k states the standby duty start time index of reserve crew member k.
		Output
P	:	Vector of probabilities. The probabilities that neither the originally assigned crew or reserve crew are unavailable for each flight. p_d is the probability that the originally assigned crew are absent and reserve crew are not available to replace those absent crew.

Table 5.1: Notation

5.1.2 Fundamental equations

The $SPCAM$ is based on the idea that the probability that a crew related disruption goes uncovered depends on the probability that the disruption occurs in the first place and the probability that reserve crew are available to cover the disruption. The $SPCAM$ (namely Algorithm 2) that is presented in this chapter outlines how to calculate the probabilities that all flights in a schedule are cancelled due to crew unavailability (i.e. P) after taking into account the effect that a reserve crew schedule (X) has on reducing the probabilities of cancellations due to crew absence. In short, the $SPCAM$ is a function of the initial probabilities of crew absence Q and a reserve crew schedule X, which returns P. I.e. $P = f(Q, X)$. The end goal is to use this function as the basis for a surrogate objective function (Section 5.3) that will be used to find an optimal (disruption minimising) reserve crew schedule. The advantage of this approach is that the proposed model is much faster than using repeat simulations to evaluate the quality of different potential reserve crew schedules. An iterative scheme is employed to calculate P for any reserve crew schedule X starting from any Q. The iterative scheme is based on the following assignment equations which are applied for each reserve crew (k) in turn and for each crew absence disruption (d) each reserve crew could feasibly be used for.

$$r_k^{d+1} := r_k^d (1 - p_d) \tag{5.1}$$

$$p_d := p_d \left(1 - r_k^d\right) \tag{5.2}$$

In Equations 5.1 and 5.2 d denotes the departure number and r_k^d the probability that reserve crew member k is available at departure d. p_d gives the probability that crew are unavailable for departure d, after taking reserve crew availability into account. Equation 5.1 shows how to update the probability that the reserve crew member remains available for subsequent disruptions, given that they might be used to cover the given disruption. Equation 5.2 shows how to calculate the probability that a crew related disruption still occurs given that a reserve crew member is available with some probability. These equations provide the basic modelling principle on which Chapters 5 to 8 are based. The following example is used to illustrate the basic idea for the example of a single flight, with a 0.1 chance of crew absence, and a single reserve crew is scheduled to begin their standby duty at the departure time of that flight. This example corresponds to the first iteration of Algorithm 2.

$$
\begin{aligned}
& \text{If, } p_1 = 0.1\text{, i.e. there is a 0.1 chance of crew absence} \\
& \text{and, } r_1^1 = 1\text{, i.e. reserve 1's duty has just begun} \\
\implies \quad & r_1^2 = 1 \times (1 - 0.1) = 0.9 \\
& \text{because the reserve has a 0.1 chance of being used} \\
\implies \quad & p_1 = 0.1 \times (1 - 1) = 0 \\
& \text{because the reserve will definitely be available}
\end{aligned}
$$

These equations are applied iteratively to each departure, considering the effect of each scheduled reserve crew member, the procedure for doing this is outlined in Algorithm 2, in Section 5.2.1.

5.2 The model

The problem inputs can be represented as a vector of probabilities Q, where each element denotes the probability of crew absence for a particular departure. A reserve schedule X can be represented as a list of reserve duty start times, or equivalently departure numbers (because possible reserve duty start times are discretised according to the scheduled departure times). For example, $x_k = 5$ means that reserve k starts their standby duty at the scheduled departure time of flight 5, i.e. lower case is used to refer to elements of vectors P, Q and X.

$$Q = \{q_1 \; q_2 \; q_3 \; q_4 \; ...\} \tag{5.3}$$
$$X = \{x_1 \; x_2 \; x_3 \; x_4 \; ...\} \tag{5.4}$$

For a given set of departures with associated probabilities of crew absence and a reserve schedule, we can determine the effect the given reserve schedule has on the probabilities of crew unavailability. The vector of probabilities of crew unavailability is denoted P and is a function of the reserve schedule and the probabilities of crew absence of the originally scheduled crew, $P = f(Q, X)$. The procedure for finding P for a given reserve schedule in Algorithm 2 reflects the assumptions given in Section 5.1.1.

5.2.1 Calculating the effects of a reserve crew schedule on the probabilities of crew unavailability

Algorithm 2 Scheme for calculating crew unavailability probabilities given a reserve crew schedule

1: $P = Q$
2: **for** $k = 1$ to R **do**
3: $\quad r_k^{x_k} = 1$
4: $\quad$ **for** $j = x_k$ to $min(n, x_k + L - 1)$ **do**
5: $\qquad r_k^{j+1} = r_k^j(1 - p_j)$
6: $\qquad p_j = p_j(1 - r_k^j)$
7: $\quad$ **end for**
8: **end for**

The first line of Algorithm 2 sets P equal to Q. Q represents the initial problem and P will finally represent the probabilities of crew unavailability after the reserve crew schedule has been taken into account. The role of Algorithm 2 is to calculate P. Algorithm 2 states that, for each reserve crew that begins a duty (lines 2 to 7), initialise the probability of that reserve crew's availability to 1 (line 3). For each departure that occurs (lines 4 to 7) in a reserve crew's duty, update the probabilities that they remain available for subsequent disruptions (line 5) and update the probability that no crew are available for that departure (line 6).

Here it is assumed that departures take place at equal time intervals, as a result, reserve crew duty lengths are defined in terms of a constant number of departures (L) that occur within their duties. In subsequent chapters departure times are based on real schedule data and the feasibility of reserve crew also depends on the expected durations of the disrupted crew duties.

In Algorithm 2 reserve crew are considered in earliest start time order, because X is sorted in an ascending order. In this way Algorithm 2 reflects the policy that reserve crew are used in earliest start time order (also referred to as the default policy, see Section 3.5.2). Changing the order in which reserve crew are considered allows other (order based) reserve policies to be modelled.

5.3 Surrogate objective function investigation

The main motivation for developing this probabilistic model is to provide a surrogate objective function for evaluating potential reserve crew schedules, where the probabilistic model is to be used to replace evaluation by simulation. Evaluation by simulation is computationally expensive and is not a viable option for the evaluation intensive search methodologies that are to be used to search for a good reserve crew schedule.

In order to use this model to search for a good quality reserve schedule, the reserve schedule that minimises the vector P must be determined. A single valued measure of P is required, this will enable simple comparison of the relative quality of alternative reserve crew schedules. A number of single valued measure of P are possible, Table 5.2 lists the possibilities.

Symbol	Objective function	Equation
A	Sum of P	$\sum_{j=1}^{n} P_j$
B	Max P	$Max_{j=1..n}(P_j)$
C	Standard deviation	s
D	Coefficient of variation	$\frac{s}{\bar{P}}$
E	Product of mean and standard deviation	$s\bar{P}$
F	Weighted sum of mean and max P	$a\bar{P} + b\max_{j=1..n}(P_j)$
G	Mean absolute deviation	m
H	D with mean absolute deviation	$\frac{m}{\bar{P}}$
I	E with mean absolute deviation	$m\bar{P}$

Table 5.2: Objective functions

In order to find the best form of the surrogate objective function based on the $SPCAM$, a simulation framework is used to derive performance measures corresponding to the optimal reserve schedules according to each of the surrogate objective functions given in Table 5.2. Reserve utilisation and flight cancellation will be the performance measures used to determine the most effective form for the surrogate objective function. The simulation

Objective function	Reserve utilisation (rank)	Cancellation rate (rank)
A	0.9455 (1)	0.1848 (1)
B	0.9303 (7)	0.1904 (7)
C	0.9439 (4)	0.1864 (5)
D	0.8747 (9)	0.2475 (9)
E	0.9442 (3)	0.1859 (3)
F	0.9410 (6)	0.1871 (6)
G	0.9435 (5)	0.1860 (4)
H	0.8916 (8)	0.2378 (8)
I	0.9444 (2)	0.1856 (2)

Table 5.3: Reserve utilisation and cancellation rates in 2000 simulations (50000 flights)

experiments are based on 20 randomly generated problem instances consisting of 25 departures, each with uniformly random generated probabilities of crew absence for each departure and 9 reserve crew available for scheduling. The number of possible reserve schedules—or combinations of start times for reserve crew—for the simplified crew absence problem is as follows:

$$Reserve\ schedules = \begin{pmatrix} n \\ R \end{pmatrix} = \frac{n!}{R!(n-R)!} \tag{5.5}$$

The structure of the solution space is such that the natural definition of neighbouring solutions is that of moving a reserve crew's start time to an alternative start time which currently has no reserve crew assigned. This neighbourhood structure is used in the local search based approaches described in Section 5.4.1. For the case when $n = 25$ and $R = 9$ the number of possible reserve crew schedules is approximately 2 million. This excludes reserve schedules involving more than one reserve beginning duties at the same time, which for this particular formulation of the problem are guaranteed to be suboptimal. This problem size is small enough to allow the enumeration of the optimal reserve crew schedules corresponding to each surrogate objective function. For each of the 20 problem instances, 100 repeat simulations were performed, making 2000 simulations in total. Table 5.3 contains a summary of the results obtained. It shows that the 'sum of P' objective function (A) ranked first on both criteria. It was found that the 'coefficient of variation' objective function (D) gave the lowest reserve utilisation and highest cancellation rates. There also appears to be a correlation between the rankings on both criteria, the only difference in the results is that objective functions C and G swaps ranks.

5.4 Solution methodology investigation

Based on the performance of the sum of P surrogate objective function in terms of simulation derived reserve utilisation and cancellation rates, this objective function now forms the basis of an investigation of search/solution methodologies. Enumeration becomes less viable as the problem size grows so more intelligent solution methods are required. A variety of solution methods are tested using the reserve utilisation and cancellation rate criteria. The same experimental design as in Section 5.3 is used. The additional feature of the proposed *SPCAM* is that expected reserve utilisation and cancellation rates can be derived from the sum of P objective values, using the following equations:

$$Expected\ reserve\ utilisation\ rate = \frac{(\sum_{i=1}^{n} q_i) - (\sum_{i=1}^{n} p_i)}{R} \tag{5.6}$$

Intuitively, Equation 5.6 means that the average probability of each reserve crew member being used is the expected reduction in the number of uncovered crew absences as a result of the given reserve crew schedule divided by the number of reserve crew. In words, the expected reserve utilisation is the expected number of cancellations when no reserve crew are available minus the expected number of cancellation when reserve crew are available divided by the number of reserve crew. That is, the average number of cancellation prevented per reserve crew.

$$Expected\ cancellation\ rate = \frac{\sum_{i=1}^{n} p_i}{n} \tag{5.7}$$

The expected cancellation rate is simply the expected number of flights without crew divided by the number of flights.

5.4.1 Description of solution methods

Having established a fundamental modelling principle and an appropriate objective function for the problem of reserve crew scheduling, the *SPCAM* is now used as the method of evaluating potential solutions in a variety of search methodologies. It will be obvious from Equation 5.5 that enumeration (ENUM) becomes intractable for realistic sized problem instances. So alternative methods are considered and described below.

Pruned dynamic programming algorithm

The method referred to as a pruned dynamic programming algorithm (DP) is based on dynamic programming. It uses the idea of states (number of reserve crew assigned) and stages (departure number) to implicitly enumerate the solution space. It is also a branch and bound algorithm because entire branches (partial solutions) can be eliminated early during the search using lower and upper bound estimates of the objective values of partial solutions. The algorithm constructs and searches a binary tree in a breadth first manner where each level of branching represents a departure time and each path from the root of the tree to a leaf represents a partial solution. Each iteration of the algorithm considers the next departure and adds a layer of depth to the tree. The algorithm branches on each leaf remaining from the end of the previous iteration by either scheduling a reserve crew member or not scheduling a reserve crew member (hence a binary tree), until all of the departures have been considered. The lower bounds and upper bounds that are used to eliminate partial solutions as early as possible are heuristically estimated. The heuristics complete the partial solutions and the objective values of these completed solutions are used as the bound values for pruning. The lower bound heuristic is such that it always gives a solution of better quality (where better quality here means lower objective value) than the upper bound heuristic with both starting from the same partial solution. Partial solutions are then eliminated if their lower bound is greater than the minimum upper bound of partial solutions in an equal or higher state. The use of upper and lower bound heuristics for pruning partial solutions makes this method a heuristic, but a heuristic with a high probability of obtaining optimal solutions. The choice of upper and lower bound heuristics influence how ruthless the pruning strategy is, and therefore the probability that the partial solution corresponding to the optimal solution is pruned. The algorithm can be made faster or more cautious by using (respectively) lower or higher quality upper bound heuristics. The lower bound heuristic should always outperform the lower bound heuristic (by definition) and should be of the highest quality available, whilst considering that it should ideally be fast.

Variant dynamic programming heuristic

An alternative approach to the application of dynamic programming to this reserve crew scheduling problem is to alternate the definitions of states and stages which were used for the DP method described in Section 5.4.1. In the variant dynamic programming (VDP) approach, stages correspond to the total number of reserve crew scheduled and states to the latest departure number at which a reserve crew is scheduled, in a given reserve crew schedule. In this heuristic approach to dynamic programming, only the best partial reserve crew schedule is retained in each state in each state, and no partial solution completion heuristics are used. The objective values used, are those of the partial solutions. In each stage, each remaining partial solution is branched on by adding one reserve crew at later times than the latest scheduled reserve crew in that schedule. At each stage, the number of reserve crew in each partial solution is equal to the stage number. After the final stage, the reserve crew schedule with the lowest sum of P is taken as the solution.

Population based heuristics

The implementation of a Genetic Algorithm (GA) uses a binary vector representation of candidate reserve crew schedules, 2 competitor tournament selection, single point crossover and a mutation rate of 0.001 applied to every chromosome. The population size is 100 and 100 generations are performed before the algorithm terminates. The constraint on using a fixed number of reserve crew means that crossover can lead to infeasible solutions with either more or less than the required number of reserve crew. This issue was dealt with by applying greedy heuristics (backwards and forwards heuristics, see 5.4.1) to obtain feasible candidate reserve crew schedules with the required number of reserve crew. This is similar to the approach used in memetic algorithms (see Section 2.6.2 for more details), however in this application

the heuristics are only used to correct infeasible members of the population in each generation.
In the Ant Colony (AC) approach each of the 100 ants visits R of N positions. For each move made by an ant, a cumulative distribution corresponding to the ant's next possible moves is created from the pheromone vector. A random input is compared with the cumulative distribution to determine where the ant will move to. The sum of P is computed for each ant's tour, and an evaporation factor of 0.95 is applied to the pheromone vector in each of 200 iterations. The ant with the smallest objective value is used to lay pheromone in such a way that replenishes the amount evaporated (so that the cumulative probability of each ant move distribution remains 1). For more details about ant colony optimisation, see Section 2.6.2.

Constructive heuristics

The Backwards Heuristic (BH) starts with the (optimal but) infeasible solution of reserve crew assigned to each period. Reserve crew are removed one at a time choosing the one that increases the objective value the least.
The Forwards Heuristic (FH) starts with no reserve crew assigned and adds one at a time choosing the one that decreases the objective value the most.
The Basic Greedy (BG) approach positions reserve crew corresponding to the R highest probabilities in the original probability of crew absence vector Q.
The Even Distribution (ED) heuristic allocates reserve crews evenly across departures so that the reserve crew are spread evenly over the set of departures.

Local search based methods

Hill Climbing (HC) searches the local neighbourhood and takes the best move only if its better than the current best, using the same neighbourhood structure as the tabu search and simulated annealing implementations.
The Simulated Annealing implementation (SA) [54] is based on: a temperature reduction applied every 4 iterations (epoch=4); an initial temperature of 3; a final temperature of 0.001; a geometric temperature reduction factor of 0.999; the same neighbourhood structure as used in tabu Search.
The Tabu Search (TS) [42] implementation uses the neighbourhood structure described after equation 5.5. A recency tabu list is retained in which the swapping of reserve crew between two positions in the schedule is prevented for a tenure of 50 iterations after a swap is made, this means the tabu list refers to elements of the search neighbourhood rather than the solutions themselves. This approach prevents moves from immediately being undone (i.e. cycling) whilst encouraging the exploration of new regions of the solution space. The method uses 200 iterations always accepting the best non-tabu move.
The Variable Neighbourhood Search method (VNS) [71] uses 5 neighbourhoods (in order): single swap; cut and swap; single point crossover; sideways shift and randomly generated neighbourhoods. If a neighbourhood contains a better solution the solution is accepted as the current solution and a new

iteration begins starting from neighbourhood one. If a better solution is not found the next neighbourhood is tested, if no improving solution is found after cycling through all neighbourhoods the procedure is terminated.
The VDP+HC method is the HC method starting from the VDP solution.

5.4.2 Solution method results

Each of the solution methodologies described in Section 5.4.1 were applied to the 20 problem instances introduced in Section 5.3, and the solutions from each were tested in 10000 repeat simulations[1] of the corresponding problem instances (200000 simulations in total for each approach). Table 5.4 compares the simulation results with the expected results derived theoretically from the surrogate objective values for the criteria of reserve crew utilisation and cancellation rates. The objective function in each case is the sum of P as this was found to be the most effective surrogate objective function in Section 5.3. Table 5.4 also gives the solution times[2] for each method. The first row contains the results found from enumeration in Section 5.3 and corresponds to the optimal solution, this gives a benchmark for judging the effectiveness of the various search and optimisation techniques. Table 5.4

Solution method	Objective value	Reserve crew utilisation	Expected reserve crew utilisation	Cancellation rate	Expected cancellation rate	Solution time (s)
ENUM	**4.5315**	0.9439	0.9433	**0.1818**	0.1813	1296
ED	5.1054	0.8784	0.8795	0.2035	0.2042	0.016
BG	4.8258	0.9236	0.9106	0.1930	0.1930	0.047
FH	4.6125	0.9383	0.9343	0.1853	0.1845	0.125
BH	4.5655	0.9425	0.9395	0.1830	0.1826	0.141
HC	4.5322	0.9445	0.9432	**0.1818**	0.1813	1.654
SA	4.5320	0.9438	0.9432	**0.1818**	0.1813	35.072
TS	**4.5315**	0.9439	0.9433	**0.1818**	0.1813	32.639
VNS	4.5320	0.9436	0.9432	0.1819	0.1813	34.804
GA	**4.5315**	0.9439	0.9433	**0.1818**	0.1813	**9.278**
AC	4.7416	0.9388	0.9199	0.1903	0.1897	24.546
DP	**4.5315**	0.9445	0.9433	**0.1818**	0.1813	38.957
VDP	4.5450	0.9427	0.9418	0.1823	0.1818	0.422
VDP+HC	4.5316	0.9438	0.9433	**0.1818**	0.1813	**0.437**

Table 5.4: Objective values, simulation coverage levels and solution times for a variety of solution methods

shows that, apart from full enumeration, the TS, GA and DP approaches were able to find the theoretical optimal solutions to all problem instances. The GA approach required the least time to do this. The cancellation rate achieved in repeat simulations was minimised by the ENUM, HC, SA, TS,

[1]The average cancellation rate RMSE for 10-fold cross validation drops to 0.0005 for a sample size of 10000 simulations

[2]Matlab, dual core 1.86ghz, 2gb, windows vista

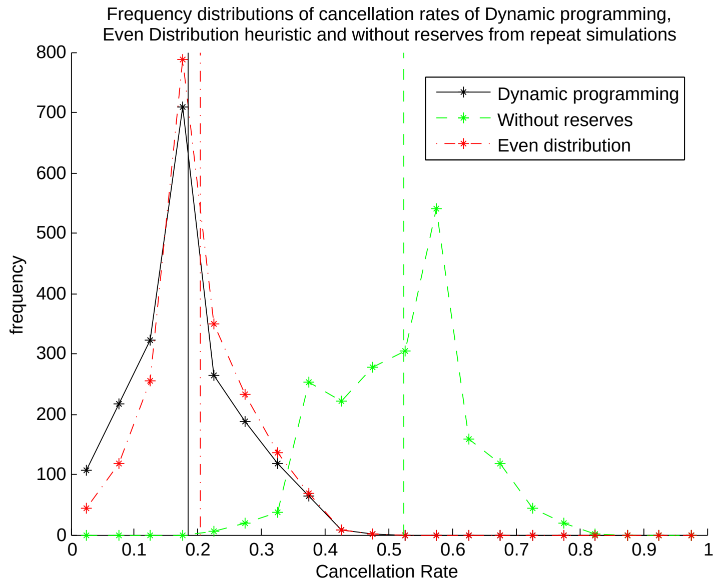

Figure 5.1: Range of cancellation rates provided by the dynamic programming solution, the even distribution heuristic and with no reserves allocated

GA, DP and the VDP+HC approaches. The VDP+HC approach also has a very low solution time. The main conclusion here is that meta-heuristics provide a suitable alternative to enumeration.

The results of Table 5.4 also show that the expected reserve utilisations and the expected cancellation rates, which were derived theoretically using Equations 5.6 and 5.7, are accurate predictions for the reserve utilisation rates and the cancellation rates which were attained in simulation testing. In conclusion, average reserve crew schedule performance can be predicted using equations 5.6 and 5.7, but simulation is still required to determine the variability of the performance of reserve schedules in various crew absence scenarios. The performance variability of the optimal solution compared to the even distribution heuristic and when no reserves are assigned is demonstrated in Figure 5.1. Figure 5.1 gives an idea of the worst case performance, best case performance and expected performance in terms of cancellation rate for the dynamic programming solution, the solution derived from the even distribution heuristic and when no reserves are allocated. The reason for focussing on these methods is that the even distribution heuristic represents how reserves are scheduled when probabilities of crew absence are not considered in decision making (i.e. the reserve crew are allocated at equal time intervals), the results therefore illustrate the effectiveness of the probabilistic approach, when solved using the dynamic programming approach. The results for the dynamic programming approach are skewed towards a lower cancellation rate compared to the even distribution heuristic. The result is a lower average cancellation rate (see vertical lines). When no reserves are allocated, the average cancellation rate is approximately 0.5, as expected, due to the problem instances being generated using uniform random numbers. The asymmetric shape of the cancellation frequency distribution when no reserve crew are scheduled corresponds to the fact that

20 different problem instances were used to test each method. The reason why the symmetry is restored when reserve crew are scheduled is because the variance of the cancellation rate is reduced by the presence of reserve crew, this means that reserve crew have a stabilising effect on the expected level of cancellations.

5.5 Possible model improvements

The assumptions of Section 5.1.1 can be modified to accommodate real world airline operations. The assumption that each departure requires one member of crew (which was justified as representing a team of crew) can be replaced with each departure requiring any specified number of crew. The multiple crew absence version of this model requires that for each departure there is a probability distribution for the number of crew that will be absent. The following chapter extends the model in this direction, in which the vector P becomes a matrix. Evaluating the effect of reserve schedules on the probabilities of cancellations due to crew absence requires significant modifications compared to Algorithm 2.

An underlying assumption made in this simplified model is that all crew can be treated as identically functional units, the reality is that crew come in a variety of ranks and qualifications. The next chapter also presents an extended formulation of the model where these considerations are included. The extended formulation takes the structure of crew pairings into account, which was ignored in this simplified model.

The current model deals with crew absence disruptions only. In Chapter 7 an analogous model is developed where crew-related delay disruptions are considered whilst crew absence disruptions are ignored.

Another assumption made in this model was that reserve crew can only be used to cover for disruptions whose departure time falls within their duty period. A relaxation of this is considered in the following chapter where reserve crew can be used for departures occurring before the start of their duties. This means that departures might be delayed whilst waiting for reserve crew to begin their duties.

Reserve crew and flight cancellation are the only recovery actions used in this model, however preprocessing of the input probabilities can be performed to reflect the probable availabilities of other more preferable recovery actions such as crew swaps, this is especially the case for crew-related delay disruptions. Chapters 7 and 8 represent two different probabilistic approaches (based on this chapter) that account for crew-related delays and also the availability of swap recovery actions.

5.6 Chapter summary

This chapter has introduced a probabilistic model of crew absence uncertainty and the process of replacing absent crew with reserve crew. The model outlines the fundamental modelling principles to follow when tackling the problem of reserve crew scheduling under uncertainty. An investigation of the possible surrogate objective functions based on the $SPCAM$ was carried out. Then, a variety of heuristic solution methodologies were investigated. It was found that optimal solution could be found, besides a full enumeration, using dynamic programming, genetic algorithm, tabu search. The results of this chapter underpin the models developed in subsequent chapters.

Chapter 6

Improved probabilistic crew absence model

The probabilistic crew absence model of Chapter 5 provides a fundamental modelling principle for modelling the uncertainty of crew related disruptions and the use of reserve crew to recover from such disruptions. However, Chapter 5 made several simplifying assumptions. In this chapter the Simplified Probabilistic Crew Absence Model ($SPCAM$) of Chapter 5 is extended and improved in several ways, resulting in an improved probabilistic crew absence model (CAM). The weaknesses of the $SPCAM$ of Chapter 5 that are addressed in this chapter are as follows.

1) **Crew absence was modelled with a single probability per departure**. This simplification did not allow for the possibility that more than one of the crew assigned to a flight can be absent simultaneously. In this improved model the single probability of crew absence for each flight is replaced with a probability distribution containing the probabilities that different numbers of crew are simultaneously absent. As a result of considering the possibility of multiple crew from a crew team being simultaneously absent, reserve crew feasibility depends on the combined feasibility of a group of individual reserve crew. Furthermore, flight cancellation due to crew unavailability can only be avoided if and only if all absent crew are replaced, one for one, with reserve crew.

2) **The probabilities of crew absence affecting flights in a schedule were independent**. This simplification ignored the structure of an airline's schedule, that is, crew teams are often assigned to multiple flights, and therefore the probabilities of crew absence affecting those flights are dependent upon one another. In this chapter, the *binary crew absence assumption* (see assumption **C5** of Section 4.2) is made that crew, if absent, are unavailable for all flights in their assigned crew pairing. The CAM takes the structure of the crew schedule into account.

3) **Reserve crew were feasible to cover for crew absence disruptions affecting a fixed number of departures after the beginning of their standby duty**. The $SPCAM$ ignored the details of the airline's schedule, such as departure times and arrival times. In this chapter, such details are not ignored, so that now, reserve crew feasibility is based on assumptions **RC1-6** of Section 4.2.

4) **Reserve crew were not feasible to cover crew absence disruptions affecting flights whose scheduled departure time was before the start of their standby duty**. This simplification did not allow for the possibility that, a flight which was affected by crew absence, can wait for reserve crew to begin their standby duties. The relaxation of this simplification introduces the possibility that using reserve crew to avoid flight cancellation can introduce a delay. In this chapter, these reserve use induced delays are incorporated into the expected cancellations objective function using a function which maps delays to measures of cancellation. The delay cancellation measure function was introduced in Section 3.5.1.

An additional weakness of the investigations of the previous chapter were that they considered very small problem instances. The CAM is applied to a larger and more realistic sized problem instance than the $SPCAM$ was. Section 6.1.7 shows that as a result of considering a larger problem instance and the possibility of multiple absent crew per crew pairing, the CAM is found to underestimate cancellation rates. An explanation is found, which relates to the variance that exists in the total number of crew that can be absent, which is not modelled in the CAM. A model refinement is given which ensures this variance is accounted for.

Chapter structure

In this chapter, Table 6.1 defines the notation for the CAM, Section 6.1 presents the CAM. The model refinement described above is also presented in this section. Section 6.2 validates the CAM in computational experiments. Section 6.3 shows how the CAM can be modified and applied to the case where aircraft come in a variety of fleet types and crew come in a variety of ranks and qualifications. Section 6.4 introduces a generalised version of the default online reserve policy which can be encoded within the CAM. Section 6.5 concludes with a chapter summary. Section 6.6 discusses how the CAM is used as a building block in subsequent approaches to reserve crew scheduling considered in this thesis.

6.1 New probabilistic crew absence model formulation

The $SPCAM$ of Chapter 5 is now extended to the case where crew teams consist of a number of individual crew each with independent probabilities of being absent. The single probability of crew absence for each departure used in the $SPCAM$ is replaced with a discrete probability distribution that describes the probability that different numbers of crew are simultaneously absent for each departure. Let P now be a matrix containing the set of all such distributions. So that $p_{d,e}$ corresponds to the probability that e crew are simultaneously absent for a departure d.

In the CAM, the fundamental equations (Equations 5.1 and 5.2) of

$a_{d,e}$	:	Probability that e reserve crew are available to cover e absent crew affecting departure d
A_d	:	Scheduled arrival time of departure d
b	:	Delay exponent used in the delay cancellation measure function
$C_{l,m}$	:	departure number of the m^{th} hub departure in crew pairing l
CM_d	:	Cancellation measure of reserve crew use induced delay at departure d corresponding to a given reserve schedule X
CT	:	Cancellation threshold (maximum delay before a flight is cancelled)
D_d	:	Scheduled departure time of departure d
DL	:	Maximum duration of a reserve standby duty period
DT	:	Delay threshold (minimum delay for which delay recovery actions are considered)
EDT_d	:	Expected departure time for departure d
F_d	:	Crew pairing assigned to hub departure d
$Feas_{t,d}$	:	Feasibility of a reserve crew member with start time index t covering crew absence affecting departure d (binary matrix)
$L_{l,d}$	:	departure number of the last flight of the day of crew pairing l on the day of departure d
$leafNodes$	:	The number of leaf nodes currently in the reserve crew combination tree
$Leaves$	:	The set of leaf nodes in the reserve crew combination tree at any given time
M	:	The set of reserve crew feasible for a given departure to cover crew absence
$maxCA$	:	Maximum number of crew that can be absent from a pairing. Equals the number of scheduled crew in each crew team
n	:	Number of hub departures in the airline schedule
N_δ	:	δ^{th} reserve crew node
N_δ^{len}	:	Number of reserve crew in the reserve crew combination corresponding to node δ in the reserve combination tree
N_δ^{par}	:	Parent node of node δ in the reserve crew combination tree
N_δ^{res}	:	Reserve number of the reserve crew member corresponding to node δ in the reserve crew combination tree
$nodeProb$	:	Probability that a given combination of reserve crew are simultaneously available for covering crew absence
$p_{d,e}$	:	Probability that e crew are unavailable for departure d
$q_{d,e}$	:	Initial probability of e absent crew at departure d before the affects of a reserve crew schedule are taken into account
$r_{d,k}$	:	Probability that reserve crew member k is available to cover crew absence affecting departure d
R	:	Number of reserve crew in a reserve crew schedule
$ResCom$	:	Vector containing the combination of reserve crew corresponding to a given node in the reserve crew combination tree
$u_{d,k}$	:	Probability that reserve crew member k is used to cover crew absence affecting departure d
X	:	Reserve crew schedule
X_k	:	start time index of the k^{th} reserve scheduled to begin a reserve pairing
ξ	:	Number of nodes currently in the reserve crew combination tree

Table 6.1: Notation

Chapter 5 are replaced with the analogous Equations 6.1 and 6.2.

$$p_{d,e} := p_{d,e}\,(1 - a_{d,e}) \tag{6.1}$$
$$r_k^{d+1} := r_k^d - u_{d,k} \tag{6.2}$$

Where $a_{d,e}$ is the probability that e reserve crew are simultaneously available at departure d. Equation 6.1 states that the probability that e crew are unavailable for departure d depends on the probability that e crew are absent in the first place and the probability that e reserve crew are not available (simultaneously) to cover the absence affecting departure d. Equation 6.2 gives the probability that reserve k remains available for subsequent use given that they have a probability of $u_{d,k}$ of being used to cover absence at departure d. Equations 6.1 and 6.2 are applied: for each scheduled reserve crew in earliest start time order, to reflect the assumed reserve policy; and for each scheduled departure in earliest departure time order, to reflect the assumed priority order for airline recovery actions (see **RP1** of Section

4.2 the *sequential recovery assumption*). In this case the preference order reserve policy is to use reserve crew in earliest start time order, which is referred to as the default policy (see Section 3.5.2). The calculation of $a_{d,e}$ requires the enumeration of all combinations of e reserve crew which are feasible for departure d. In section 6.1.5, the algorithm for enumerating all feasible combinations of e reserves is introduced.

Before this an explanation is given of how the CAM takes the detailed structure of crew pairings into account. The importance of doing this was highlighted in the last bullet point of Section 3.2 (the structure of crew pairings determines the maximum number of cancellations in the event of uncovered crew absence).

6.1.1 Crew pairings

Crew schedules have a hierarchical structure, single flights are called flight legs, a set of flight legs comprising a days work or shift is called a duty, a string of duties spanning a number of days that begin and end at the crew's home base is called a crew pairing.

The $SPCAM$ assumed that each flight's crew had an independent probability of being absent. This simplifying assumption does not hold when we consider that numerous departures from an airport can correspond to the same team of crew (i.e. the same crew pairing). To allow for the structure of crew pairings, the replacement assumption is as follows.

Assumption 1: The probabilities of crew absence of flights operated by the same team of crew are equal. This assumption is the probabilistic equivalent of the *binary crew absence assumption* (**C5**) of Section 4.2.
As a result of assumption 1, the CAM requires an additional assumption.

Assumption 2: Accurate estimates are available or can be derived for the probabilities that different numbers of crew will be absent from each crew team and as a result are unavailable for their entire assigned crew pairing.

These assumptions imply that crew absence becomes known on or before the beginning of the crew's assigned pairing, and that, if absent, they will not become available for the pairing at a later time. The effect of assumptions 1 and 2 on probabilistic reserve crew schedule evaluation is that the initial probability of crew unavailability for a flight, before the consideration of reserve crew, is initially equal to that of the previous flight on the same crew pairing after taking the effects of reserve crew into account. I.e. the probability that reserve crew are required to cover absence depends on the probability the crew absence was not covered at a previous flight on the same crew pairing. This means the CAM allows for the possibility that crew absence, if not covered, might still be covered at later flights along the same crew pairing.

6.1.2 Reserve crew feasibility

This section defines the feasibility of reserve crew scheduled at different times for covering flights which may be affected by crew absence. In contrast to the $SPCAM$, reserve feasibility in the CAM depends on the structure of crew pairings and the duration of the cancellation threshold (CT).

$$Feas_{t,d} = \begin{cases} 1 & D_t < D_d + CT \text{ and } D_t + DL \geq A_s \\ & (s = L_{F_d,d}) \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

Noting that possible reserve duty start times are discretised according to scheduled departure times, Equation 6.3 states that reserve crew with start time index t are feasible for a disrupted departure d, if: 1) they begin their standby duty before the cancellation threshold of the disrupted departure, and 2) their duty finishes ($D_t + DL$) at or after the final arrival time of the disrupted pairing (A_s). In the event that e crew are absent, a combination of e reserve crew is feasible if and only if each of the individual reserve crew are feasible. The feasibility of reserve crew with start time index t for covering a crew absence disrupted departure d can be pre-calculated and stored in the form of a binary matrix ($Feas$).

6.1.3 Reserve use induced delay

Since reserve crew can be used any time between the scheduled departure time of a crew disrupted flight and the cancellation threshold of the flight, it is possible that some combinations of reserve crew, although preventing a cancellation, may introduce a delay to the crew disrupted flight. Such a delay is caused by waiting for the reserve crew to begin their standby duties before they can be utilised. To penalise reserve-induced delay in the evaluation of a given reserve crew schedule, the delay cancellation measure function is used (Equation 3.3 of Section 3.5.1 which is repeated here).

$$Cancellation\ measure = \left(\frac{delay}{CT}\right)^b \tag{6.4}$$

Equation 6.4 gives the cancellation measure of a delay as the ratio of the $delay$ relative to the cancellation threshold (CT), raised to the power b (delay exponent).

$$delay = \max_{k \in ResCom} (D_k - D_d,\ EDT_d) \tag{6.5}$$

Equation 6.5 gives the size of the reserve-induced delay associated with using a given combination of reserve crew ($ResCom$) to cover crew absence affecting departure d. The delay depends on the reserve crew (k) who has the latest duty start time. EDT_d is the expected delay of departure d before the effects of reserve crew are considered, and can be estimated from a simulation (Chapter 4), in which there are no reserve crew scheduled. EDT_d is the expected delay associated with the aircraft which is assigned to departure d.

$$CM_d = CM_d + g\left(\frac{delay}{CT}\right)^b \tag{6.6}$$

Equation 6.6 gives the objective value contribution (penalty) associated with reserve-induced delay due to a combination of reserve crew who have a probability of g of being used to cover for absent crew affecting departure d. CM_d denotes the total cancellation measure of reserve-induced delays affecting departure d. The calculation of g is addressed in section 6.1.5.

In Section 6.2 the CAM is tested with and without penalties for reserve-induced delay. To distinguish between the two models, the model is called the Static Delay Model (SDM) when it includes the penalty term for reserve-induced delays. Note that the word static hints at the model which is introduced in Chapter 8 which in contrast, is dynamic and responsive to the given reserve crew schedule that is being evaluated, as will be explained in that chapter.

6.1.4 Evaluating expected cancellations associated with a given reserve crew schedule

Algorithm 3 Outline of reserve crew schedule evaluation procedure

Inputs: airline schedule, the assumed reserve policy, crew absence probabilities (Q), expected delays before reserve recovery
Outputs: For all flights: cancellation probabilities, reserve use induced delay cancellation measure contributions
$P = Q$
$r_k^1 = 1, \ \forall k \in \{1...R\}$
for $d = 1$ to n **do**
  Reset a and u
  $M =$\{Feasible reserves for departure d in earliest start time order\}
  For all feasible reserves in M generate all reserve combinations containing between 1 to $|p_d|$ reserve crew
  Determine the probability that each combination is used, given that reserves are used in earliest start time order
  Determine (a) the total probability of different numbers of reserve crew being simultaneously available
  Determine (u) the probability that each individual reserve is used to cover crew absence in departure d
  Update probabilities of cancellation due to different numbers of crew absence and reserve availability for subsequent crew absence
  $p_{d,e} = p_{d,e}\,(1 - a_{d,e})\,, \ \forall e \in \{1...|p_d|\}$
  $p_{w,e} = p_{d,e}, \ \forall w \in$ \{subsequent departures assigned to crew pairing F_d\}, $\forall e \in \{1...|p_d|\}$
  $r_k^{d+1} = r_k^d - u_{d,k}, \ \forall k \in M$
end for

Algorithm 3 outlines the procedure followed by the CAM when evaluating the expected number of cancellations due to crew absence associated with a given reserve crew schedule (X). In general, Algorithm 3 considers each scheduled departure in order. For each, it enumerates feasible combinations of reserve crew and their associated probabilities of being considered for use. The probabilities that different numbers of reserve crew are simultaneously available are used to update the probabilities that flights are cancelled due to crew unavailability. The probabilities that reserve crew remain available for subsequent disruptions depends on the probabilities that they were used for crew absence affecting the given departure.

In more detail, the algorithm firstly initialises (line 3) the probabilities that different numbers of crew are unavailable for each departure (P) to the probabilities that different numbers of crew are absent for each departure (Q). Then (line 4) the reserve crew availability probabilities are initialised to 1. The algorithm then considers each scheduled departure (line 5) in earliest departure time order. For each departure, all combinations of the reserve crew which are feasible to cover absent crew affecting that departure are generated (line 8) and their probabilities of being considered for use, given that more preferable combinations are not available, are calculated (line 9). The probabilities that different numbers of reserve crew are simultaneously available (line 10) are used to calculate the probabilities that crew absence disruptions affecting the given crew pairing still go uncovered (lines 13 and 14). The probabilities that each individual reserve crew is used to cover absence at the given departure (line 11) are used to update the probabilities that each reserve crew remains available for subsequent crew absence disruptions (line 15).

The details of how the feasible combinations of reserve crew are actually generated (line 8) and their corresponding probabilities calculated (line 9) are the subject of Section 6.1.5.

6.1.5 Enumerating feasible combinations of reserve crew and associated probabilities

Lines 8 to 11 of Algorithm 3 involve enumerating feasible combinations of reserve crew of different sizes and calculating their probabilities of actually being utilised. These probabilities depend on the probabilities that different numbers of crew are not available and the probabilities that more preferable combinations of reserve crew are available for the same disruptions. Reserve combination preference is defined by the reserve use order policy, which in this case is assumed to be the earliest start time order, as to minimise reserve-induced delay. Note that any other order based policy can be used instead.

Algorithm 4 enumerates the feasible combinations of reserve crew for each possible crew absence disruption. It turns out that simply enumerating all combinations of reserve crew of different sizes also yields combinations that in reality would never occur, given the reserve use policy. Such combinations include:

1. Combinations that have been generated for a previous flight in the same crew pairing, as crew absence is covered at the earliest opportunity as it makes no sense to hold reserve crew when they can be used to cover crew absence to prevent flight cancellation. Such combinations are filtered out by line 15 of Algorithm 4.

2. Combinations involving non-consecutive reserve crew numbers with identical duty start times or identical flight feasibility. This is because

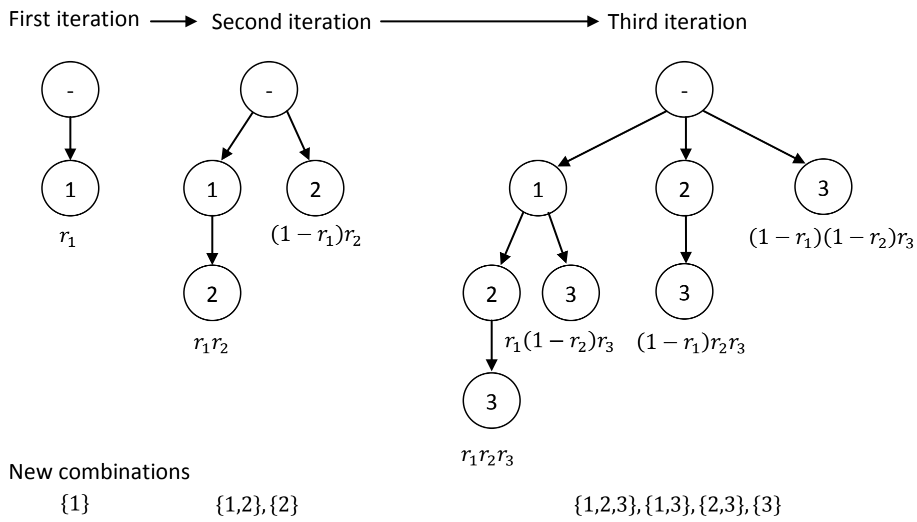

Figure 6.1: A growing reserve crew combination tree

reserve crew are used in order (which is the order specified by the given reserve policy) and an ordering is applied to reserve crew with the same start time, to eliminate ambiguity. Such combinations are not generated by Algorithm 4, lines 26 to 28 ensure this.

Another aspect of reserve combination generation that requires careful consideration is the derivation of the probabilities that given combinations are considered for use, given that they pass reserve combination filters 1 and 2 (above). For example, suppose reserve crew 1, 2, 3 and 4 are feasible recovery actions when 2 crew are absent from a given crew team. Then, a feasible reserve combination such as (1,4) (which implies 2 and 3 have different start times/flight feasibility to 1 and 4, see reserve combination filter 2) has an associated probability of $(r_1\,(1-r_2)\,(1-r_3)\,r_4)$ of being considered for use to cover the disruption. However, the $(1-r_2)$ term can be removed if the reserve crew combination (1,2) was a feasible combination for a previous flight on the same crew pairing, because had this been the case, the reserve crew combination (1,2) would be have been used to cover the two absent crew at that time. In this case, the use of reserve 4 does not depend on reserve 2 not being available, in fact their usage is mutually exclusive to each other. The same reasoning applies to the $(1-r_3)$ term. This concept is referred to as *reserve non-dependency* later on, and is asserted in lines 17 to 21 of Algorithm 4.

Algorithm 4 outlines the procedure for generating feasible combinations of reserve crew and calculating their associated probabilities of being used. The algorithm is based on building a tree of nodes, where nodes correspond to particular feasible reserve crew, and paths from the root to a leaf correspond to combinations of reserve crew.

Figure 6.1 illustrates how starting from a root node (top) the reserve crew combination tree is generated in stages, in each stage the next preferred reserve crew is added to the tree. The node probabilities for the new reserve combinations generated in each iteration are stated below the newly

Algorithm 4 Generation of reserve combinations and associated probabilities at departure d

Inputs: d, r, F, $Feas$, P and EDT
Outputs: a_d, u_d and CM_d
Create root node N_1 corresponding to the empty reserve crew combination, with no parent node of its own, path length 0 and node probability 1, $N_1^{res} = null$, $N_1^{par} = null$, $N_1^{len} = 0$, $N_1^p = 1$
$\xi = 1$ (nodes used)
$Leaves \leftarrow N_1$
$leafNodes = |Leaves|$ (nodes to be branched on)
for each $k \in M$ **do**
  for $\delta = 1$ to $leafNodes$ **do**
    if $N_\delta^{len} < maxCA$ **then**
      Branch on N_δ with reserve k
      increment ξ
      $N_\xi^{res} = k$, $N_\xi^{par} = N_\delta$, $N_\xi^{len} = N_\delta^{len} + 1$, $N_\xi^p = N_\delta^p \times r_k$
      $Leaves \leftarrow N_\xi$
      $ResCom$ =The combination of reserve crew corresponding to the path from N_ξ to the root node
      if $ResCom$ was not feasible for any previous flight on crew pairing F_d **then**
        $nodeProb = N_\xi^p$
        for each feasible reserve crew s which is not in $ResCom$, with start time$\leq D_{X_k}$ **do**
          if replacing reserve k with reserve s in $ResCom$ results in a combination of reserve crew that was feasible for any previous flight on crew pairing F_d **then**
            $ResCom$ probability does not depend on reserve s not being available, therefore $nodeProb = \frac{nodeProb}{(1-r_s)}$
          end if
        end for
        $a_{d,N_\xi^{len}} = a_{d,N_\xi^{len}} + nProb$
        $u_{d,\gamma} = u_{d,\gamma} + g, \forall \{\gamma \in ResCom\}$, where $g = \left(nProb \times p_{d,N_\xi^{len}}\right)$
        $CM_d = CM_d + g \left(\frac{\max_{\gamma \in ResCom}\left(D_{X_\gamma} - D_d,\ EDT_d\right)}{CT}\right)^b$
      end if
      if N_δ^{res} has identical feasibility to reserve k **then**
        Remove N_δ from $Leaves$
        decrement δ
      else
        $N_\delta^p = N_\delta^p \times (1 - r_k)$
      end if
    end if
  end for
  $leafNodes = |Leaves|$
end for

generated nodes. The new reserve crew combinations generated at each stage are listed at the bottom of the diagram.

In Algorithm 4 *Leaves* denotes the set of leaf nodes and *LeafNodes* the number of leaf nodes in the reserve combination tree at any given stage of the algorithm. N_{ξ}^{len} corresponds to the number of reserves in the combination of reserves beginning at the root node and ending at node ξ. N_{ξ}^{par} is the parent node of node ξ in the tree. N_{ξ}^{res} gives the reserve number corresponding to node ξ. Line 3 defines the root node as node 1, corresponding to a reserve combination of 0 reserves, without it's own parent node and a node probability of 1. Lines 5 and 6 add the root node to the set of nodes (*Leaves*) that are to be branched on with nodes corresponding to the first feasible reserve crew in the first iteration of the algorithm.

The reserve combination tree is then grown by branching on each leaf node with nodes for each feasible reserve crew in turn, in earliest start time order (lines 7 and 8). This means that each path from the root node to a leaf node defines a combination of reserve crew listed in earliest start time order, with no repeat reserve crew. Additionally, no leaf nodes are more than 'the maximum number of absent crew' away from the root node (line 9), as such reserve crew combinations are never required. ξ is the number of nodes in the reserve crew combination tree at any given time. So node ξ always corresponds to the newest reserve crew combination (*ResCom*) generated by the tree. The probability that the reserve combination (*ResCom*) corresponding to node N_{ξ} is used depends on the probability that more preferable reserve crew are not available (asserted on line 30) or whether or not the reserve combination is subject to the *reserve non-dependency* described above (asserted on lines 17 to 21). Every time a node is branched on by a new reserve node, the node which was branched on remains a leaf node, but the node probability is updated so that it corresponds to the newest reserve not being available (line 30). The branch node then corresponds to combinations which that reserve is a member of. The branch node is added to the set *Leaves*, to be branched on by subsequent reserve crew. Given the probability (*nodeProb*) that the reserve crew combination *ResCom* is considered for use: line 22 updates the probability that a total of N_{ξ}^{len} reserve crew are available (a) to cover N_{ξ}^{len} absent crew affecting departure d; line 23 updates the probabilities that the individual reserve crew in *ResCom* are used (u) to cover crew absence affecting the given departure (d); and line 24 updates the delay cancellation measure contribution for departure d, corresponding to *ResCom* (see Equations 6.5 and 6.6). Lines 26 to 31 ensure that reserve combinations are not generated that fall into the category of reserve combination filter 2 (see above). Line 34 sets the number of nodes that are to be branched on when nodes corresponding to the reserve with the next highest start time are added to the reserve crew combination tree.
The solution space of this reserve crew scheduling problem is considered in the next section.

6.1.6 Solution space

The number of possible reserve schedules is as follows.

$$\sum_{j=ceil\left(\frac{R}{MaxCA}\right)}^{R} \frac{n!}{j!(n-j)!} y(j,R) \tag{6.7}$$

Equation 6.7 gives the number of ways R reserve crew can be assigned to the n different possible reserve standby duty start time indices (scheduled departure times), where no more the $MaxCA$ (maximum number of crew absent from each crew pairing) reserve crew are assigned to any individual start time index. Where $y(j,R)$ is the number of combinations of j integers $(1 \leq integers \leq MaxCA)$ that sum to R. For the case where no restriction is placed on the number of reserve crew that can begin duties at the same time the $y(j,R)$ values are in the 'Bell number' sequence. In Equation 6.7 the summation accounts for each of the numbers of partitions that R reserve crew can be divided into, where each partition contains no more than $MaxCA$ reserve crew. The factorial term accounts for the number of ways such a number of partitions can be allocated across n possible start times. The $y(j,R)$ term accounts for number of ways R reserve crew can be divided into j partitions. Understanding Equation 6.7 can be useful when considering possible neighbourhood structures for local search heuristics.

The reserve crew schedule X specifies the start time index in start time order of each reserve crew scheduled. Where a start time index X_k corresponds to the beginning of a standby reserve pairing, where standby duties begin at time D_{X_k} daily. A feasible solution must contain the correct total number of reserves (R) and have no more than $maxCA$ scheduled to begin their duty at the departure time d (as no more than this will be required for covering crew absence at departure d).
The objective of the SDM is to minimise cancellations due to crew absence plus the cancellation measure contributions due to reserve-induced delay. For the CAM there is no cancellation measure term.

$$ObjVal = \sum_{d=1}^{n} \sum_{e=1}^{|p_d|} p_{d,e} + \sum_{d=1}^{n} CM_d \tag{6.8}$$

In Chapter 8 the form of the objective function is the same as for the SDM, however the cancellation measure term is calculated using a statistical model of delay propagation, which allows for delays from all causes, their propagation, swap recovery actions as well as reserve crew used to replace delayed crew.

6.1.7 Improved model

In this section a pitfall of the CAM as presented so far is demonstrated and a model improvement is proposed. The example is based on the airline schedule described at the start of Section 6.2, with a reserve crew schedule, consisting of 12 individuals, derived from a greedy heuristic algorithm. Figure 6.2 shows the predicted cancellation probabilities derived from the

CAM before (purple) and after the model refinement (green), compared to predictions derived from repeat simulations (blue). The predictions from the $SPCAM$ are also given (cyan). Figure 6.2 shows the problem with the

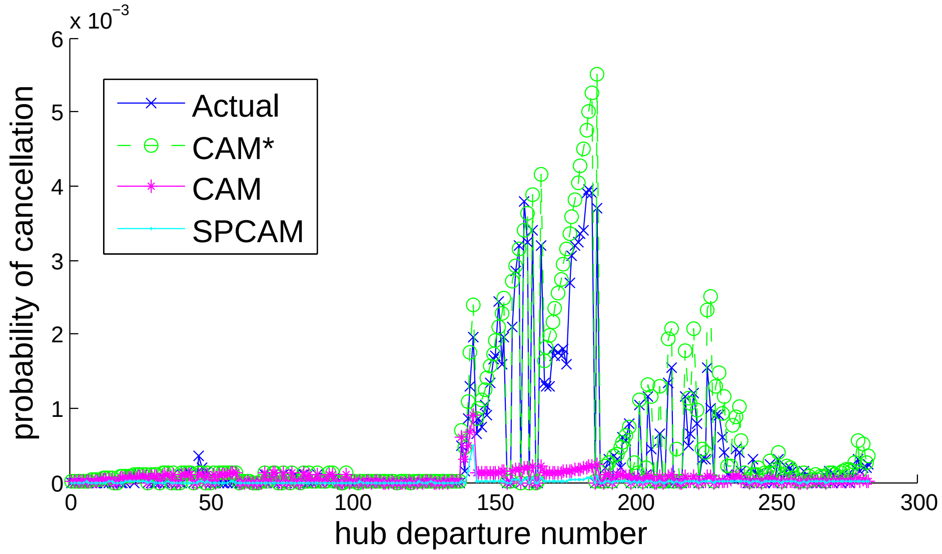

Figure 6.2: Cancellation predictions from evaluations of the $SPCAM$, CAM and refined probabilistic crew absence model ($CAM*$) compared to those derived from repeat simulations

CAM, which is, that it underestimates the probabilities of cancellations compared with predictions derived from simulation. The explanation for this is that the CAM implicitly assumes that the total number of absent crew is always exactly the expected value. However, in the simulation the total number of crew that are absent in any given run of the simulation varies a great deal, and on bad days, once reserve crew have been used, cancellations spike. In fact, when all crew have equal probabilities of absence (as assumed in this chapter) the total number of absent crew follows a binomial distribution. Let $p1$ be the probability that a single crew member is absent and there are $(MaxCA \times n)$ crew in total, the probability that z crew are absent in total is given by $\frac{(MaxCA \times n)!}{z!((MaxCA \times n)-z)!} p1^z (1-p1)^{((MaxCA \times n)-z)}$. The factorial term is the binomial coefficient, i.e. the number of ways that exactly z crew can absent out of $(MaxCA \times n)$ crew. The second part $(p1^z(1-p1)^{((MaxCA \times n)-z)})$ is the probability with which each of those instances (of exactly z crew being absent) occurs.

To remedy the cancellation underestimation problem, a refined probabilistic crew absence model ($CAM*$) uses the distribution of the total number of absent crew (i.e. the binomial distribution) to evaluate reserve crew schedules simultaneously over a distribution of P matrices, denoted O, where matrix O_z corresponds to a P matrix where the total expected number of absent crew is z. The Matrix O_z is therefore the P matrix corresponding to the case where the probability that any crew member is absent is equal to $p1 = \frac{z}{MaxCA \times n}$. The weight of O_z in the objective function of the $CAM*$ is taken from the binomial distribution, i.e. the probability that z absences occur in $n \times MaxCA$ opportunities ($binomP\,(z, n \times MaxCA, p1)$).

In summary, the variance of the total crew that can be absent is captured by modelling the variance of the probability that a single member of crew is absent. So the CAM is evaluated over a set of P matrices, each of which is constructed from a different probability that each single member of crew is absent. The new approach also captures the variance that exists in the probabilities that each reserve crew member remains available to replace disrupted crew affecting each of the scheduled flights in the airline's schedule.

Algorithm 5 Procedure for incorporating the variance of the total number of crew that can be absent into reserve crew schedule evaluations using the CAM

1: $objVal = 0$
2: **for** $z = 1$ to n **do**
3: $p1 = \frac{z}{MaxCA \times n}$ probability a given crew member is absent
4: $objVal = objVal + binomP(z, n \times MaxCA, p1) \times evaluation(O_z, X)$
5: **end for**

Algorithm 5 shows that the $CAM*$ evaluation procedure is identical to the CAM except that the procedure is repeated (line 2) over a distribution of P matrices (O). For high values of z (total number absent) the associated binomial distribution probabilities become very small and can be ignored. One possibility for doing this is to limit evaluation to the total number of absent crew corresponding to a cumulative probability of 0.95 or 0.99, in the following 0.999 is used.
Figure 6.2 demonstrates that the $CAM*$ gives cancellation predictions of greater accuracy compared to the CAM, where the simulation predictions are treated as the target values.

6.2 Experimental results

The $SPCAM$, CAM and SDM are now validated through experimentation. Two more models are also tested: $CAM*$ and $SDM*$, which correspond to CAM and SDM respectively, with the addition of the model refinement of Section 6.1.7. The $SPCAM$ implementation uses the probability of at least one crew absence affecting each departure as the single input probability for each departure. Apart from this the $SPCAM$ is the same as the CAM.

6.2.1 Test instance

The experiments are based on real airline schedule data. The schedule is 2 days in length, with 283 departures from the hub station. There are 209 teams of crew and 74 aircraft covering a total of 566 flights. 140 of the crew teams begin their pairings at the hub station, these crew teams are subject to crew absence uncertainty. Each member of crew has a 1% chance of being absent, this value was approximated from actual crew absence data. The proposed approach will still work even if there is sufficient data to derive individual crew absence probabilities for all crew, the matrix P would then have to be constructed from these. Each team of crew consists of 4 members. There are 12 reserve crew available for scheduling. The aircraft routings are taken directly from real airline schedule data, the scheduled departure and arrival times are adjusted so that the scheduled block times are equal to the average actual block times and because of this, the average delay risk in the schedule is 50%. The journey time distributions for each origin-destination pair are derived from real flight data. The aircraft turn times and crew sit times are set to the minimum values. The input crew schedules were generated using a set partitioning model (described in [12]) solved in CPLEX. The average rate of mid duty crew aircraft changes is 0.44. The following experiments were implemented on a laptop with a 2.4GHz dual core Intel Core i7-5500U CPU, with 8Gb of RAM. All models, algorithms and the simulation were implemented in Java as single threaded applications. The validation simulation is that described in Chapter 4.

6.2.2 Experiment design

The $SPCAM$, CAM, $CAM*$, SDM and $SDM*$ are now all used in a variety of heuristics to derive reserve crew schedules. The heuristics considered are as follows:

GH: The greedy heuristic adds reserve crew one at a time to a reserve crew schedule, each time selecting the start time that reduces the objective function the most, continuing in this fashion until all of the reserve crew are scheduled.

LS: Local search starts from a randomly generated initial solution (randomly generated start time indices). In each iteration all solutions neighbouring the incumbent solution are evaluated, the solution which reduces the objective value the most is accepted. If no improving solution is available the algorithm terminates. Local search uses the cut-and-insert neighbourhood structure, that is, all solutions that have one reserve start time different to that of the incumbent solution.

GH+LS: LS starting from the GH solution.

SA: The simulated annealing [54] implementation uses the cut-and-insert neighbourhood. Each iteration randomly selects a neighbouring solution, which is accepted if it is an improving move. A non-improving move is accepted with probability $e^{-\Delta/T}$. Δ is the increase in the objective value associated with the non-improving move, T is the current temperature. The cooling scheme (value of T at any given iteration) is based on an exponential decay starting from T_0 equal to the maximum number of hub departures in a crew pairing and reaching a final temperature of 0.000001 after 20000 iterations. The cooling scheme is a function of the number of evaluations that have been performed so far.

GA1: The first genetic algorithm [43] implementation uses a population size of 50, uses four competitor tournament selection, a mutation rate of 0.001, single point cross-over which is applied with probability 1 and all parent chromosomes are replaced with children chromosomes in each generation. The genetic algorithm returns the best solution found after 20000 function evaluations.

GA2: The second genetic algorithm implementation is the same as GA1, except that the mutation operator is replaced with a single iteration of the SA algorithm (only one random neighbouring solution is considered), which is applied to each member of the population, in each generation. The SA-based mutation operator uses the same temperature scheme as the SA algorithm. GA2 is limited to a total of 20000 evaluations, which means that GA2 uses half as many generations as GA1, the other half are used evaluating SA generated mutations. GA2 is similar to a memetic algorithm [46], because of the addition of a local search based approach, to the algorithm. Only similar because, one iteration of simulated annealing is closer to a mutation operator than an application of local search.

All heuristics except for the GH are limited to 20000 function evaluations. The GH only requires around 3000 function evaluations to derive a reserve crew schedule for the given problem. 20000 function evaluations take around 10 minutes on the above described hardware software combination. This approach enables a fair test of the different heuristics in a situation where solutions are required within a set time limit. Each heuristic is repeated 10 times using each probabilistic reserve crew schedule evaluator ($SPCAM$, CAM, $CAM*$, SDM and $SDM*$). Each derived reserve crew schedule is then tested in 20000 repeat simulations to derive cancellation and delay based performance measures. Each repeat simulation uses different stochastic inputs to instantiate numbers of absent crew for each crew pairing. Firstly though, the results of this experiment are used to assess the cancellation prediction accuracy of the probabilistic crew absence models ($SPCAM$, CAM and $CAM*$) over a larger sample, compared to the single instance considered in Figure 6.2.

6.2.3 Cancellation prediction accuracy

This section compares the predicted average cancellations due to crew absence from the $SPCAM$, CAM and $CAM*$ with the average cancellation rates observed in repeat simulations. These experiments were repeated for each reserve crew schedule derived from the experiment described above in Section 6.2.2. Note that the cancellation rate predictions of the SDM and $SDM*$ match those of the CAM and $CAM*$ respectively. Figure 6.3 confirms that the CAM (and the $SPCAM$) underestimates cancellations due to crew absence, and that the $CAM*$ successfully alleviates this problem. The $CAM*$ does however systematically overestimate cancellations due to crew absence, each time in a manner similar to that demonstrated in Figure 6.2. Possible reasons for this include: 20000 repeat simulations are not enough to capture a representative sample of the worst case scenarios in which cancellation spikes occur; cumulative rounding errors; the probabilities of reserve combinations calculated based on filters 1 and 2 of Section 6.1.5 have additional factors/intricacies which have not yet been uncovered. Despite this, the $CAM*$ consistently gives the most accurate cancellation predictions, which is supported by the linear trend equation, which has a gradient close to one, an intercept close to zero and a high correlation coef-

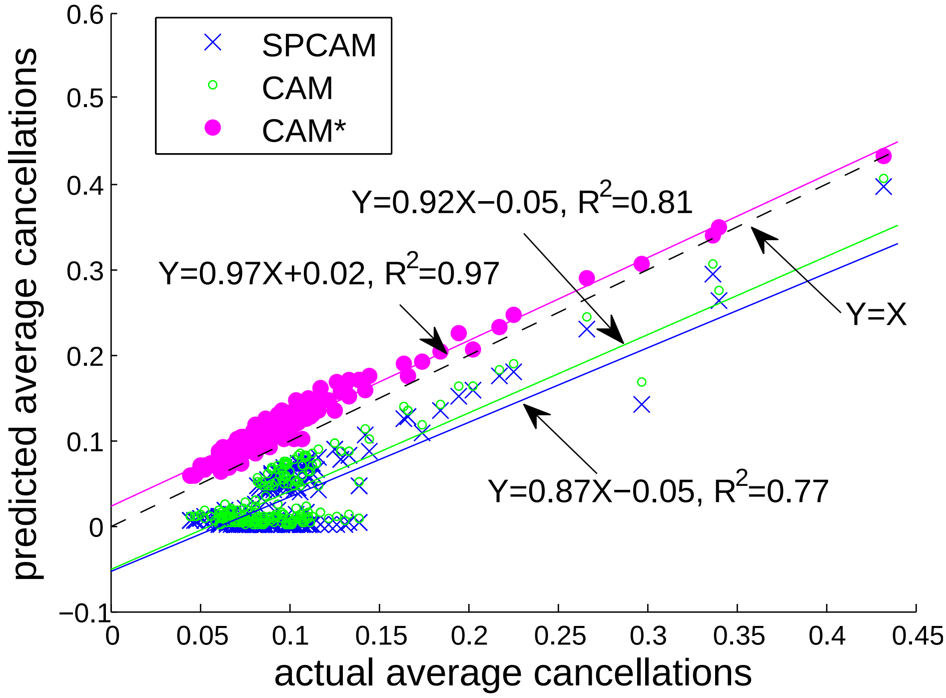

Figure 6.3: Cancellation predictions from evaluations of the $SPCAM$, CAM and $CAM*$ compared to those derived from repeat simulations

ficient compared to those of the $SPCAM$ and the CAM. On the plus side, overestimations of cancellations due to crew absence are not as potentially damaging as underestimations, because cancellations are the most severe outcome from crew absences. A systematic overestimation of cancellations due to crew absence could be beneficial, as this corresponds to a more risk averse approach to reserve crew scheduling.

6.2.4 Reserve crew scheduling application

The results of the experiment described in Section 6.2.2 are now used to show the effects that the probabilistic models and the search heuristics which were used to schedule reserve crew have on the quality of the resultant reserve crew schedules. The average cancellation measures (cancellations plus cancellation measures of delay) derived for each reserve crew schedule tested in 20000 repeat simulations are used as the measures of reserve crew schedule quality. Figure 6.4 shows the average cancellation measure of the best reserve crew schedules from 10 repeats of each heuristic used in conjunction with each evaluator. The results show that as the evaluator complexity increases ($SPCAM$ to $SDM*$) the average cancellation measure decreases. The $CAM*$ and $SDM*$ evaluators typically lead to respectively higher quality reserve crew schedules than the CAM and SDM evaluators. This result is supported by the results given in Table 6.2, which validates the model refinement of Section 6.1.7. The variance of the average cancellation measures for reserve crew schedules derived from the $SPCAM$, CAM and $CAM*$ evaluators can be explained by them not allowing for reserve-induced delay.

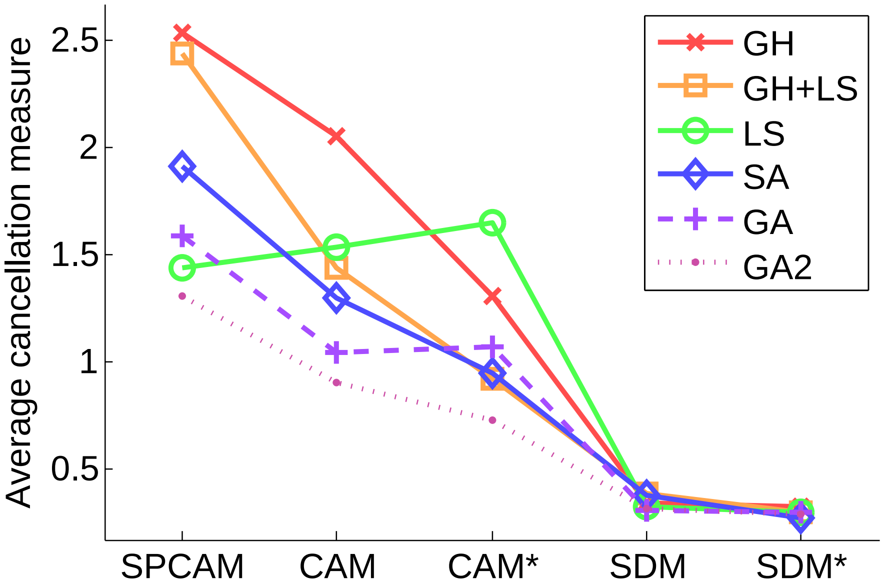

Figure 6.4: The effect of each variant of the probabilistic crew absence model and the search heuristics used to schedule reserve crew on the resultant simulation derived average cancellation measure

This is confirmed by the results of Table 6.2. The low variance of the results for the SDM and $SDM*$ evaluators can be explained by the fact that each heuristic is capable of deriving a good quality solution, provided that the evaluator used includes all of the aspects on which solution quality is judged. For the $SPCAM$, CAM and $CAM*$ evaluators, GA2 always gave the best reserve crew schedule. For the SDM evaluator, GA1 gave the best solution. For the $SDM*$, SA gave the best (overall) reserve crew schedule. The LS approach did not attain any of the best reserve crew schedules found. An explanation for this is that the cut-and-insert neighbourhood structure had a size of 3396 neighbouring solutions at any given iteration, which all had to be evaluated before accepting the best neighbouring solution. This meant that the LS approach never reached a local optimum, because only 5 full iterations could be performed within the 20000 evaluations limit. For this reason a Tabu Search was not implemented, as it would not have had the chance to exploit a tabu list. Figure 6.4 also shows that the accuracy of the evaluator is, in general, more important than the complexity of the search algorithm used to derive a reserve crew schedule.

6.2.5 Extra performance measures and alternative approaches

In this section additional performance measures are given for the best reserve crew schedules derived using each variant of the probabilistic crew absence model. They are also compared with several alternative approaches to reserve crew scheduling. The alternatives are as follows:

USR: The uniform start rate heuristic schedules the available reserve crew at times corresponding to equal intervals of hub departures. E.g. if

Method measure type	Method name	Cancellation measure: Average (best repeat)	Cancellation measure: St dev (10 repeats)	Average cancellations: due to absence	Average cancellations: due to delay	Delays $>$ delay threshold: Average (minutes)	Delays $>$ delay threshold: probability	Reserve utilisation rate	Maximum cancellation measure
-	no res	10.1668	0	10.0991	0	6.802	0.0818	0	37.02
Probabilistic	*SPCAM*	1.3119	0.2115	0.0906	0.01350	17.307	0.1067	0.4627	18.75
	CAM	0.9027	0.3615	0.0679	0.01260	14.098	0.1041	0.4632	17.18
	$CAM*$	0.7275	0.3848	0.0532	0.00555	12.557	0.0987	0.4630	16.18
	SDM	0.3136	0.0519	0.0620	0.00420	8.684	0.0926	0.4623	13.53
	$SDM*$	0.2739	0.0024	0.0903	0.00200	8.065	0.0930	0.4583	15.60
Heuristic	*USR*	1.3963	0	0.2967	0.00785	16.535	0.1020	0.4543	19.17
Simulation	*AREA*	0.9823	0	0.3366	0.00610	12.44	0.0983	0.4378	14.48

Table 6.2: Comparison of approaches to reserve crew scheduling using more performance measures

there are 25 hub departures and 5 reserve crew, $X = \{1, 6, 11, 16, 21\}$.

AREA: See Section 4.7.1.

The results in Table 6.2 correspond to the best of 10 repeats of the two alternative approaches described above. These are compared with the best reserve crew schedules derived using each of the probabilistic evaluators considered in this chapter. Note that when the best single repeats are replaced with the average of the 10 repeats of each method, the ordering of the methods, in terms of average cancellation measure, is the same. The average cancellation measure for the best repeats of each method are given to indicate the potential of each approach. The standard deviation from the 10 repeats indicates the reliability of each method.

In Table 6.2 the average cancellation measures show that the $SDM*$ results in the lowest average cancellation measure and this true for all 10 repeats as indicated by the associated low standard deviation. Table 6.2 gives the average expected number of cancellations due to crew absence and due to delays exceeding the cancellation threshold, for each method. When no reserve crew are scheduled, there are an average of 10 cancellations, all of which are due to absence. This means that all of the observed cancellations due to delay, in Table 6.2, are caused by reserve-induced delays which have been propagated and have caused delays above the cancellation threshold later on. Cancellations due to delay are highest for the $SPCAM$, which is because this approach does not penalise reserve-induced delays. The SDM and $SDM*$ reduce cancellations due to delay, because they do penalised reserve-induced delay. The $SDM*$ also lead to the lowest rate of cancellations due to delay. The $CAM*$ achieved the lowest cancellations due to absence, which can be attributed to it overestimating cancellations due to absence, resulting in reserve crew schedules which are highly risk averse in terms of this type of disruption. The delay based performance measures show that the lowest results for delays above the delay threshold and their probabilities of occurring were achieved by the approaches which penalise reserve-induced delay, i.e. SDM and $SDM*$. The reserve utilisation rate results show that reserve utilisation rate is not an indicator of the quality of a reserve crew schedule, because the maximum reserve utilisation rate corresponding to the CAM also attained a relatively high average cancellation measure. It is therefore possible to use reserve crew badly, especially if they are scheduled badly. It is interesting to note that the $SDM*$ attains the lowest average cancellation measure by finding a balanced trade-off between cancellations due to absence and reserve-induced delays. I.e. it trades cancellations due to absence for reduced delays and cancellations due to delays. The maximum cancellation measure performance attribute (last column) gives the worst case total cancellation measure from the 20000 repeat simulations. When no reserve crew were scheduled a total cancellation measure of 37 was accumulated in the worst case. The SDM has the lowest maximum cancellation measure.

6.2.6 Results summary

The results of this section have validated the improvements to the $SPCAM$ presented in this chapter. Allowing for the possibility of: simultaneous crew absence; the structure of crew pairings; the possibility of reserve-induced delay; and the variance present in the total number of crew that can be absent on any given day has resulted in a model which yields accurate cancellation predictions and reserve crew schedules which find a balanced trade-off between delay and cancellation minimisation.

Chapter 10 gives results for the CAM (and the SDM) when modified for the case of multiple fleet types, crew ranks and qualifications (see Section 6.3) and applied to the problem of reserve crew scheduling. Chapter 10 compares all approaches to reserve crew scheduling considered in this thesis in the same test instances.

6.3 Extended formulation: Including aircraft fleet types and crew ranks and qualifications

This section explains how the CAM can be extended to allow for multiple fleet types and the ranks and qualifications of crew. Note that extensions of the CAM introduced in the following sections apply equally to the $CAM*$, the SDM and the $SDM*$, but CAM is notationally more convenient.

Up until this point the focus has been limited to the simplified case of a single fleet, this meant that all aircraft had the same crew requirements and all reserve crew were qualified to operate on all aircraft. This simplification was justified by the assumption that the reserve crew scheduling problem decomposes into a separate problem for each fleet, in which reserve crew qualified for each fleet are scheduled independently. However, this is not entirely the case because reserve crew can be qualified to operate on a number of an airline's fleets. This means that a decomposition approach would have to restrict reserve crew to one of the fleets they are qualified for. However such an approach would preclude finding an optimal reserve crew schedule, because it will not exploit the extra flexibility provided by reserve crew being qualified for a number of different fleet types. In summary, the consideration of multiple fleets requires that reserve crew qualifications are also taken into account. In the following, a three fleet and three qualification example is considered. Table 6.3 shows that each qualification group is qualified to operate on a different subset of two fleets out of the three fleets.

Airline fleets are characterised by sets of aircraft of the same type. Aircraft fleets are defined by the model number (B737 for example), different fleets have different passenger capacities and distance ranges and as a result have different cabin crew requirements. The notation $FCR_{fl,rank}$ (from Fleet Crew Requirements) is used to denote the number of crew of rank $rank$ required for fleet type fl. In the following, a three fleet and two crew rank example is used where the fleet crew requirements are those given in Table 6.4. Typically the greater the passenger capacity the greater the total

$(fl \neq qual)$	Fleet		
Qualification	1	2	3
1		✓	✓
2	✓		✓
3	✓	✓	

Table 6.3: Reserve crew qualification groups and the fleets they are qualified for

	Fleet		
Rank	1	2	3
0 (low)	3	4	5
1 (high)	1	1	2

Table 6.4: Fleet crew requirements (FCR)

cabin crew requirement. Additionally for long haul flights cabin crew may require rest periods mid-flight, therefore increasing the total cabin crew requirement. In summary, the consideration of fleets requires that each fleet's crew requirements are taken into account.

Cabin crew come in a range of ranks depending upon their level of experience and training. The highest cabin crew rank is purser, all flights require at least one purser amongst the crew complement, and sometimes more for larger capacity fleet types. In this section, the CAM is also extended to the case where crew come in two ranks, referred to as low and high rank, where each fleet type requires a specified number of crew of each rank. The inclusion of cabin crew ranks means that the CAM has to allow for the possibility of different numbers of crew of each rank being simultaneously absent from a crew team. Additionally, cabin crew can fly below rank (see assumption **RC2** of Section 4.2, the *fly below rank assumption*), this means that reserve crew of high rank can be used to cover low rank crew absences if required. The following section describes what aspects of the single fleet, crew rank and qualification formulation of the CAM have to be modified to allow for the case of multiple fleets, crew ranks and qualifications.

6.3.1 Required modifications

Different numbers of absent crew of each crew rank

To allow for the possibility of different numbers of absent crew of each crew rank, the crew unavailability matrix P gains an extra dimension. Whereas before $P_{d,e}$ denoted the probability that e crew are unavailable for departure d, we now have $P_{d,e,f}$ to denote the probability that e low rank cabin crew and f high rank cabin crew are unavailable simultaneously for departure d.

The possibility of different numbers of absent crew of each crew rank requires that the len characteristic of reserve nodes used in reserve crew combination generation (Algorithm 4) becomes a vector of length 2 with len_0 denoting the number of low rank reserve crew and len_1 denoting the number of high rank reserve crew in the combination of reserve crew ending on the given reserve node.

Flying below rank

The feature that cabin crew can "fly below rank" means that a given combination of low and high rank absent crew can be covered by reserve crew in several different ways. Ranging from, covering all absent crew using high rank reserve crew, to covering all low rank absent crew using low rank reserve crew and all high rank absent crew using high rank reserve crew. In the CAM this feature can be taken into account by considering all of the combinations of low and high rank absent crew that a given combination of reserve crew can be used to cover exactly. Given a fleet type fl, a number of low rank reserve crew (lr) and a number of high rank reserve crew (hr), the notation $combos_{fl,lr,hr}$ denotes the set of combinations of numbers of low and high rank absent crew that can be covered exactly using lr low rank reserve crew and hr high rank reserve crew. Furthermore, $combos_{fl,lr,hr,cn,0}$ is the number of low rank ($rank = 0$) absent crew and $combos_{fl,lr,hr,cn,1}$ is the number of high rank ($rank = 1$) absent crew that are covered by the cn^{th} combination. Each time a feasible reserve crew combination is generated in Algorithm 4 (see Section 6.1.5) all of the ways that such a combination can used (as stored in *combos*) have to be taken into account (see Section 6.3.1). The numerous combinations of low and high rank absent crew that a given reserve crew combination can be used to cover arise from the possibility of flying below rank.

Reserve crew combination generation

In Section 6.1.5 reserve crew combinations for a given crew disrupted flight were generated by constructing a tree of reserve nodes where each path from a node back to the root node defined a reserve crew combination. The inclusion of crew ranks and the possibility of using reserve crew to fly below rank requires a modification of Algorithm 4. The problem with Algorithm 4 is that high rank crew absence can only be covered using high rank reserve crew, so reserve crew combinations containing low rank reserve crew cannot, according to *combos*, be used to cover absence involving only high rank absence crew. If low rank reserve crew are added to the tree first it is possible that no combinations will be generated involving high rank reserve crew only. One possible solution is to add all of the high rank reserve crew to the tree first, however this would interfere with the preference order reserve policy that is encoded within the CAM, as this would model a reserve order policy where high rank reserve crew are always used for roles below their assigned rank whenever this is possible. As a result, the reserve combination tree has to be generated twice, firstly using only high rank reserve crew whilst only considering their effect on the probabilities of crew absence involving only high rank absent crew. Then a second time where high and low rank reserve crew are included in the reserve crew combination tree, whilst only considering the effects of reserve crew combinations on crew absences involving at least one low rank crew. The reason why high rank only crew absences have to be treated separately is that the probabilities that high rank reserve crew are used to cover such absences do not depend on the probabilities that low rank reserve crew are available or not, which is because low rank reserve crew cannot be used to cover high rank crew absence. Excluding low rank reserve crew from reserve crew combination generation gives the correct node probabilities that high rank reserve crew are available to cover high rank only crew absences.

The second time the reserve crew combination tree is generated both low and high rank reserve crew are included, allowing for the possibility that high rank reserve crew are used as low rank reserve crew. This time the tree is used to update the probabilities that crew absence involving at least one low rank crew absence can be covered using reserve crew. As the use of high rank reserve crew in this case does depend on the probabilities that low rank reserve crew are available or not. As a result *combos* is only required for the second reserve combination tree generation. So *combos* does not include using just high rank reserve crew to cover just high rank crew absence, as these will already have been taken into account in the first reserve crew combination tree.

Probabilities that different combinations of absent crew can be replaced with reserve crew

Before the inclusion of fleets, ranks and qualifications, the notation $a_{d,l}$ denoted the probability that l reserve crew were simultaneously available at departure d to cover l absent crew. To account for the possibility of different numbers of absent crew of each rank, the extended notation $a_{d,l,m}$ is used to denote the probability that reserve crew are available to cover l low rank crew and m high rank crew that may be simultaneously absent at departure d. For the case of multiple fleets, ranks and qualifications, Line 22 of Algorithm 4 is replaced by Algorithm 6.

Algorithm 6 Algorithm for updating the probabilities that reserve crew are available to cover different numbers of absent crew of different ranks

1: **for** $cn = 1$ to $|combos_{fl,N_\xi^{len_0},N_\xi^{len_1}}|$ **do**
2: $\quad lac = combos_{fl,N_\xi^{len_0},N_\xi^{len_1},cn,0}$ (number of low rank absentees covered by combination)
3: $\quad hac = combos_{fl,N_\xi^{len_0},N_\xi^{len_1},cn,1}$ (number of high rank absentees covered by combination)
4: $\quad a_{d,lac,hac} = a_{d,lac,hac} + nProb$
5: **end for**

Algorithm 6 updates the probabilities (a) that the different combinations (line 1) of numbers of simultaneously absent crew of each rank that can be covered for using the reserve crew combination defined by the path from reserve node N_ξ back to the root node. $N_\xi^{len_0}$ gives the number of low rank reserve crew and $N_\xi^{len_1}$ the number of high rank reserve crew in the reserve crew combination corresponding to reserve node N_ξ, which are temporarily stored as *lac* and *rac* on lines 2 and 3 (for notational convenience in line 4).

Because high rank reserve crew can be used to cover low or high rank absent crew, the reserve crew combination generation algorithm (Algorithm 4) needs to allow the generation of reserve crew combinations involving more high rank reserve crew than there can be high rank absent crew from a given flight. For this purpose Constraint 6.9 replaces that of line 9 of Algorithm 4.

$$\textbf{IF} \quad \left(\begin{array}{c} rank \equiv 1 \\ \cap \\ ((FCR_{fl,0} + FCR_{fl,1}) > (len_0 + len_1)) \end{array} \right) \\ \cup \\ \left(\begin{array}{c} rank \equiv 0 \\ \cap \\ ((FCR_{fl,0} + FCR_{fl,1}) > (len_0 + len_1)) \cap \\ (FCR_{fl,0} > len_0) \end{array} \right) \tag{6.9}$$

Constraint 6.9 allows high rank reserve crew branch nodes on leaf nodes that correspond to reserve combinations where the total number of reserve crew is below the maximum total number of crew that can be absent for a flight involving fleet fl. Constraint 6.9 allows low rank reserve crew branch nodes on leaf nodes that correspond to reserve combinations where the number of low rank reserve crew is less than the maximum total number of low rank absent crew provided that the total number of reserve crew is less than the total number of crew that can be absent for a flight involving fleet fl.

Expanded solution representation

For the case of multiple fleets, crew ranks and qualifications the solution representation has to capture for each reserve crew scheduled: a reserve duty start time; a rank and a qualification. This leads to the following solution representation (Equation 6.10). Where sti_k denotes the start time index of reserve crew k, $rank_k$ denotes the rank of reserve crew k and $qual_k$ denotes the qualification of reserve crew k.

$$X = \left\{ \left\{ \begin{array}{c} sti_1 \\ rank_1 \\ qual_1 \end{array} \right\}, ..., \left\{ \begin{array}{c} sti_R \\ rank_R \\ qual_R \end{array} \right\} \right\} \tag{6.10}$$

6.4 Generalised reserve policy

As a result of extending the CAM to the case of multiple fleet types, crew ranks and qualifications (FRQs), the assumed reserve order policy of the CAM requires a modification. The problem is that when considering the case of FRQs different combinations of reserve crew with different rank and qualification compositions can have the same associated duty start times. The earliest start time order reserve policy considered so far cannot distinguish between such combinations.

This section introduces a generalised reserve policy which allows reserve crew to be considered in orders that are defined by a number of criteria, not just their start times. For example, if using reserve crew to fly below rank is considered undesirable, high rank reserve crew can be considered last so the low rank disrupted crew are replaced with low rank reserve crew if possible. Another possible order in which reserve crew can be considered for use is the time of day (modulo 24 hours) start time order (regardless of start date) as opposed to absolute start time order (time and date order). It is also possible that the reserve crew preference order can be based on which reserve crew have the estimated lowest future demand, so as to leave the largest possible remaining amount of reserve crew capacity for future disruptions. Such an approach would be adaptive to the numbers of reserve crew of each rank-qualification variety remaining at the given time.

The order based reserve policy can be generalised using a weighted sum of each of the above stated considerations. Rosenberger et al. [87] also used a weighted sum approach for reserve crew selection, this was described in Section 2.3. For each reserve crew available for a given disruption, a relative score can be calculated for each reserve order criterion. Then, the weighted sum is calculated by multiplying the order scores with the corresponding weights and taking the sum. Such an approach can be encoded within the CAM, as the resultant order can be used as the order in which reserve crew are added to the reserve crew combination tree in Algorithm 4. The generalised reserve policy can also be used in the simulation when testing both reserve crew schedules and reserve holding policies to select which combination of the available reserve crew should be used in any given situation.

6.4.1 Generalised reserve policy parameters

The generalised reserve policy (GRP) uses four parameterised criteria to determine the preference order of reserve crew use for a given departure. The criteria are: start time; absolute start time; rank (as in the desirability of flying below rank) and expected future demand. This section introduces the notation for the GRP and how relative scores are calculated for each criterion. Table 6.5 gives the notation for the relative reserve order scores,

ATS_k	: Reserve k absolute relative reserve crew start time score
STS_k	: Reserve k relative start time score
DS_k	: Reserve k relative demand score
RS_k	: Reserve k relative fly below rank score
ATW	: Absolute start time policy weight
STW	: Start time score policy weight
DW	: Reserve demand policy weight
RW	: Reserve rank policy weight
OS_k	: Total weighted relative order score of reserve k
$resD_k$	: Estimated future demand for reserve k
$maxD$	: Maximum estimated future demand for reserve crew
$p1_{rank,d}$	: Probability of one crew absence of rank $rank$ effecting departure d
ord_l	: The l_{th} reserve crew to consider for use for a given disruption
FL_d	: The fleet type assigned to departure d

Table 6.5: Generalised reserve policy notation

policy weights as well as notation used for estimating relative future demand scores for reserve crew. The relative order scores are calculated as follows.

$$ATS_k = \frac{k}{R} \tag{6.11}$$

Equation 6.11 states that the relative absolute start time score for a reserve is simply the reserve number divided by the total number of reserve crew, as reserve crew are ordered in absolute start time order initially. The relative absolute start time score is an ordinal type criterion, whereas the relative start time order score is a ratio type criterion.

$$STS_k = \left\{ \begin{array}{ll} \frac{\max\left(D_{sti_k} - D_d,\, 0\right)}{CT} & \text{if reserve } k \text{ is feasible} \\ M \text{ (large number)} & \text{if reserve } k \text{ is infeasible} \end{array} \right\} \tag{6.12}$$

Equation 6.12 states that the relative start time order score (STS_k) for a reserve crew member (k) is proportional to the delay they cause if they are used for departure d divided by the cancellation threshold. The relative start time order score for infeasible reserve crew is set to a very large number, as this ensures they end up last in the order in which reserve crew are considered for use (note that if they are infeasible they will not be considered anyway).

$$\begin{aligned}
DS_k &= \frac{resD_k}{maxD} && (6.13)\\
resD_k &= \textstyle\sum_{j \in \{\{d+1 \text{ to } n\} | Feas_{sti_k,j} = true,\ qual_k \neq FL_j\}} \left(\frac{r_k}{TFR_j} \times p1_{rank_k,j} \right) && (6.14)\\
p1_{0,j} &= \textstyle\sum_{f=0}^{FCR_{FL_j,1}} p_{j,1,f} && (6.15)\\
p1_{1,j} &= \textstyle\sum_{e=0}^{FCR_{FL_j,0}} p_{j,e,1} && (6.16)\\
TFR_j &= \textstyle\sum_{m \in \{\{1 \text{ to } R\} | Feas_{sti_m,j} = true,\ qual_m \neq FL_j\}} (r_k) && (6.17)\\
maxD &= \max_k (resD_k) && (6.18)
\end{aligned}$$

Equation 6.13 gives the relative future demand score (DS_k) for a reserve crew member (k), which is the estimated future demand ($resD_k$) for the reserve crew divided by the maximum estimated future demand ($maxD$) for a reserve crew member (Equation 6.18). The estimates of $resD_k$ are calculated, using Equation 6.14 by summing for each future flight the relative availability of reserve k (if feasible) compared to the total availability of feasible reserve crew (TFR_j, Equation 6.17) multiplied by approximate demand for reserve crew, which is taken as the probability of 1 crew absence of the same rank as reserve k ($p1_{rank_k,j}$). Equations 6.15 and 6.16 show how P is used to calculate the probabilities of 1 crew absence of each rank. Using P and r to determine reserve orders during an evaluation of the CAM means that GRP is responsive to the particular reserve crew schedule being evaluated in the CAM.

$$RS_k = rank_k \tag{6.19}$$

Equation 6.19 states that the relative fly below rank score (RS_k) of a reserve crew is simply their rank, 0 if low rank and 1 if high rank. So if the fly below rank policy weight (RW) is non-zero, this will have the affect of moving reserve crew of higher rank towards the end of the order in which reserve crew will be considered for a disruption.

$$\begin{aligned} OS_k &= (ATW \times ATS_k) + (STW \times STS_k) \\ &\quad + (DW \times DS_k) + (RW \times RS_k) \end{aligned} \tag{6.20}$$

$$ord = \text{Reserve crew numbers in lowest to highest order of } OS \tag{6.21}$$

Equation 6.20 shows how to calculate the total weighted reserve crew order scores (OS_k). Equation 6.21 gives the order in which reserve crew are to be considered for use for the given disruption.

6.4.2 Generalised reserve policy parameter space

In this section the parameter space of the generalised reserve policy is defined which is then investigated in Sections 6.4.3 and 6.4.4. The parameter space of the generalised reserve policy is the set of all possible values of the tuple $\{ATW, STW, DW, RW\}$. For the following investigation the parameter space is constrained as follows.

$$\begin{aligned} ATW + ASW + DW + RW = \;& 1 \\ 0 \leq ATW \leq \;& 1 \\ 0 \leq STW \leq \;& 1 \\ 0 \leq DW \leq \;& 1 \\ 0 \leq RW \leq \;& 1 \end{aligned} \tag{6.22}$$

Constraint set 6.22 states that the sum of policy weights must equal 1 and each individual policy weight must be between 0 and 1 inclusive. The investigation of the GRP parameter space will test a sample of possible $\{ATW, STW, DW, RW\}$ tuples from a reserve crew scheduling perspective (used in the $SDM*$ to schedule reserve crew) and a day of operation reserve policy perspective (used as the policy in the validation simulation) and also look at the interaction between both perspectives. The aim is to find which policy parameters work best from a scheduling perspective and which work best from an online perspective. Note that in the validation simulation P, which is required for Equations 6.15 and 6.16, is updated as realisations of crew absence become known. I.e. the appropriate elements of P are set to 1.

6.4.3 Experimental design

To explore the effect of different sets of policy parameters used for reserve scheduling in the $SDM*$ and used online as the reserve order policy in the validation simulation, the following systematic sample of parameter sets (Equation 6.23) will each be used to generate a reserve crew schedule using the $SDM*$ as the evaluator in a simulated annealing algorithm (Section 6.2.2). Then each reserve crew schedule will be tested in simulation used in conjunction with each of the parameter sets used as the weights for the online (validation simulation) application of the GRP.

$$
\begin{array}{lllllll}
1. & \{ & 1 & 0 & 0 & 0 & \} \\
2. & \{ & 0 & 1 & 0 & 0 & \} \\
3. & \{ & 0 & 0 & 1 & 0 & \} \\
4. & \{ & 0 & 0 & 0 & 1 & \} \\
5. & \{ & 0.5 & 0.5 & 0 & 0 & \} \\
6. & \{ & 0.5 & 0 & 0.5 & 0 & \} \\
7. & \{ & 0.5 & 0 & 0 & 0.5 & \} \\
8. & \{ & 0 & 0.5 & 0.5 & 0 & \} \\
9. & \{ & 0 & 0.5 & 0 & 0.5 & \} \\
10. & \{ & 0 & 0 & 0.5 & 0.5 & \} \\
11. & \{ & 0.33 & 0.33 & 0.33 & 0 & \} \\
12. & \{ & 0.33 & 0.33 & 0 & 0.33 & \} \\
13. & \{ & 0.33 & 0 & 0.33 & 0.33 & \} \\
14. & \{ & 0 & 0.33 & 0.33 & 0.33 & \} \\
15. & \{ & 0.25 & 0.25 & 0.25 & 0.25 & \}
\end{array}
\tag{6.23}
$$

The experiment will be repeated for six different input airline schedules derived from real airline schedule data. The six test instances are those that will be used in Chapter 10 to compare all reserve crew scheduling and policy approaches considered in this thesis. The schedules are of length 1, 3 and 7 days respectively. The properties of the six test airline schedules are given in Table 10.1.

6.4.4 GRP parameter experiment results

15 reserve crew schedules were derived from using the $SDM*$ (using a simulated annealing algorithm, see Section 3.5.4), one reserve crew schedule for each set of policy parameters in Equation 6.23. Each reserve crew schedule was then tested in a simulation 15 times, one for each set of policy parameters in Equation 6.23 used as the online reserve policy. For each of the 225 simulation tests, 20000 repeat simulation runs were used to derive average cancellation measure performance measures. The result is 225 data points, where each data point consists of the sets of policy parameters used to schedule reserve crew and those used online in simulation testing and a resultant (response) average cancellation measure derived from the repeat simulations.

The best policy parameter combinations for each of the six schedules are shown in Table 6.6. Table 6.6 shows that the lowest average cancellation measures were achieved using different policy parameter combinations for each schedule. For schedule 1 the best policy parameter combination gave full weight to the absolute start time consideration both in the scheduling phase and in online testing of the resultant reserve crew schedule, this corresponds to assuming the default policy when scheduling reserve crew and using the default policy on the day of operations. In contrast the best policy parameter combination for schedule 2 involved creating reserve crew schedules whilst assuming that reserve crew being used to fly below rank is avoided if possible, the corresponding online policy parameters gave

Schedule	Policy parameter							
	Offline				Online			
	ATW	STW	RW	DW	ATW	STW	RW	DW
1	1	0	0	0	1	0	0	0
2	0	0	1	0	0.25	0.25	0.25	0.25
3	0	0	0.5	0.5	0.25	0.25	0.25	0.25
4	0	0.5	0.5	0	0.25	0.25	0.25	0.25
5	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
6	0.33	0	0.33	0.33	0.33	0.33	0.33	0

Table 6.6: The GRP policy parameter combinations that minimises average cancellation measure

equal weight to each aspect of the policy. The best online policy parameter combinations for schedules 2 to 5 all involved giving equal weights to each aspect of the policy, and, as will be shown in Figures 6.5 and 6.6, the equal weights online policy is amongst the best policy parameter combinations for schedules 1 and 6 as well. This can be taken as evidence that in general a good choice of online policy parameters involves giving equal weight to each aspect of the policy. Table 6.6 contains the policy parameters that will be used in Chapter 10 to compare all of the approaches to reserve crew scheduling and reserve policies considered in this thesis.

Figures 6.5 and 6.6 show more details for the GRP experiment results for schedules 1 and 6.

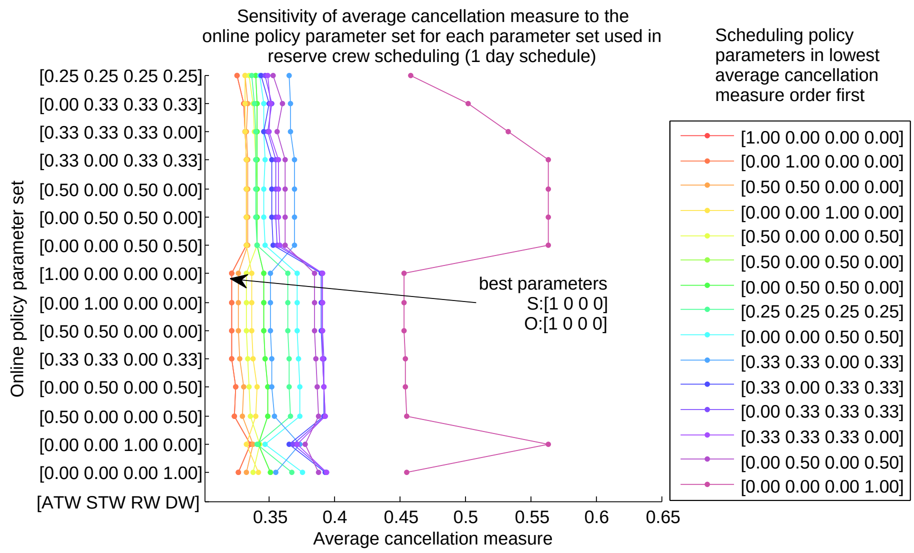

Figure 6.5: Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 1

Figure 6.5 shows the average cancellation measures from each of the 225 simulation experiments. In Figure 6.5 each data series corresponds to a reserve crew schedule derived whilst assuming one of the policy parameter sets tested in conjunction with each policy parameter combination used online. The data series corresponding to each scheduling policy parameter set have been colour coded in a rainbow order where red corresponds to the lowest average cancellation measure and purple the highest. The y-axis gives the online parameter sets corresponding to the data points in each row. The online policy parameters have also been ordered on the y-axis according to lowest to highest average cancellation measure order first, so that the best policies on average online and offline are listed first on the y-axis and in the key respectively. Note that the best on average does not necessarily correspond to the best single combination of scheduling and online policy parameter sets (see annotation). Figure 6.5 shows that for schedule 1 the average cancellation measure is significantly influenced by both the parameters used offline and those used online. This can be seen as variance in the average cancellation measure for each data series (offline policies) and each online policy on the y-axis. Figure 6.5 also shows that the worst policy both online and offline is to give full weight to the expected future demand element of the policy. This can be taken as evidence that in general the benefit of using reserve crew to cover a disruption that has actually occurred outweighs the benefit of saving reserve crew for disruptions that might occur in the future. Additionally, reserve crew with low future demand will typically be those scheduled later, when these reserve crew used in place of reserve crew scheduled at earlier times will typically cause more reserve-induced delay.

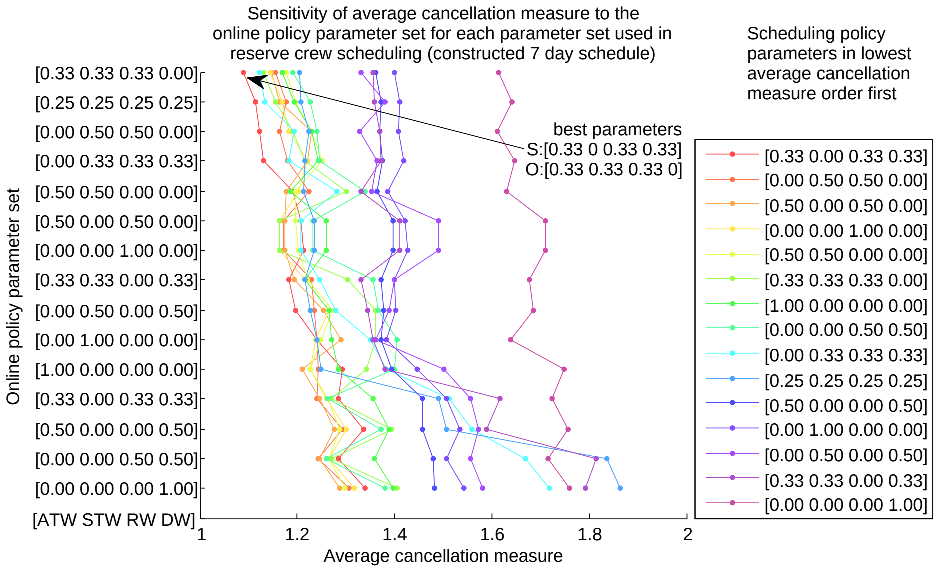

Figure 6.6: Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 6

Figure 6.6 shows the average cancellation measures corresponding to each combination of offline and online parameter sets for the case of schedule 6. Figure 6.6 shows that for schedule 6 (the constructed 7 day schedule) the average cancellation measure becomes much more sensitive to the online policy parameters compared to that demonstrated in Figure 6.5 for the real 1 day schedule. This indicates that when there are both more reserve crew and more disruptions for which those reserve crew can be utilised, the decisions regarding which reserve crew are used for each crew related disruption, become more important. In general, the online policy becomes more important when the number of ways in which reserve crew can be used increases. Figure 6.6 also shows that the worst reserve crew schedules occur when the fly below rank parameter is given a low weight, i.e. when the assumed reserve policy during reserve crew scheduling is to never try to avoid using reserve crew below their assigned rank. The reason this has a negative impact on the reserve crew schedule is that this encourages reserve crew schedules where high rank reserve crew are rarely available at times when they are likely to be used to fly below rank (because this is wasteful given their limited availability). This has the knock-on effect of reducing the efficiency with which high rank crew related disruptions can be covered. This was not the case for schedule 1 because it is a much shorter schedule and such dilemmas occur less frequently.

Appendix B contains the graphs for schedules 2 to 5 equivalent to those of Figures 6.5 and 6.6.

6.5 Chapter summary

In this chapter, the *SPCAM* of Chapter 5 has been improved by allowing for: the possibility of multiple crew being absent simultaneously from each crew team; the structure of the crew schedule with respect to the total potential disruption caused by each instance of crew absence; the possibility of reserve-induced delay; and the variance that exists in the total number of crew that may be absent on any given day of operation. Experimental results validated the proposed improvements.

The *CAM* was then extended to allow for the case where there are multiple fleet types and reserve crew are each qualified for a specified subset of all fleet types. The extended model also allows for crew ranks and the possibility that high rank reserve crew can be used in roles below their assigned rank. A result of this extension was that the assumed earliest start time reserve order policy had to be able to distinguish between combinations of reserve crew with equivalent start times but different rank-qualification compositions. A weighted sum of start time, rank and expected future reserve demand (GRP) was introduced for this purpose. An investigation into the effects of the weights for the weighted sum was carried out to select good sets of weights.

6.6 The *CAM* used in subsequent chapters

In essence, the *CAM* is a model of crew absence uncertainty and reserve crew recovery. The model can be used to evaluate reserve crew schedules in search algorithms for the purpose of reserve crew scheduling, or used to evaluate alternative reserve use decisions online as a reserve policy.

In subsequent chapters, the CAM is also used as a building block. Chapter 8 utilises the CAM to provide probabilities that crew absence is covered and teams of reserve crew are available to replace delayed crew at different times.

Chapter 7

Probabilistic crew delay model

This chapter describes a probabilistic model of the occurrence of crew-related delays, and how such delays propagate through an airline's schedule. The model evaluates the effect that any given reserve crew schedule will have on reducing crew-related delays and their knock-on effects. The model is used to search for a delay minimising reserve crew schedule. The work presented in this chapter corresponds to a conference paper [15] published during this research. The approach is analogous to the probabilistic crew absence models of Chapters 5 and 6. The probabilistic crew delay model (CDM) of this chapter ignores crew absence disruptions. The justification for this is that crew absence and delay disruptions are largely independent of one another, because crew are either absent or delayed, never both. The aim was to then combine/integrate the probabilistic crew absence and delay models to give a single model that could be used to schedule reserve crew with the objective of minimising the expected levels of delay and uncovered crew absence disruptions. This (integration) goal is reached in Chapter 8 with development of the statistical delay propagation model which uses the CAM of Chapter 6 to provide probabilities of reserve crew availability. This chapter represents an initial probabilistic model devoted to scheduling reserve crew in anticipation of crew related delays, i.e. delays which occur when an aircraft has to wait for crew who are delayed on an incoming flight, which could be avoided by replacing the delayed crew with reserve crew. Chapter 8 presents a vastly improved model to that presented in this chapter.

The CDM uses a learning phase to derive: the probabilities of crew-related delays; the probabilities that those delays are propagated; and the expected durations of those delays. The simulation used in the learning phase includes swap recovery actions from delays, and this provides the mechanism through which the CDM (indirectly) models the effects of swap recovery actions. The simulation which is used was introduced in Chapter 4, although this chapter is based on that simulation when it was at an earlier stage of its development (reserve crew teams were modelled as indivisible teams as opposed to being constructed from individual reserve crew at times when reserve crew teams are required). One of the outputs of the learning phase is a probability matrix containing the probabilities that crew-related delays propagate from one flight to another. This probability matrix (P)

is analogous to the crew absence probabilities of Chapters 5 and 6. An evaluation procedure is developed for the effect that a reserve crew schedule has on the probabilities of propagating crew-related delays and therefore the overall expected level of delay propagation. This is again analogous to the evaluation procedures of Chapters 5 and 6.

7.1 Motivation for a probabilistic model of crew-related delay

Airline schedules can become infeasible due to the disruptive effects of uncertainty in an airline's operating environment. Moreover, delays (including crew delays) can propagate through the schedule due to the presence of resource connections, so dependencies exist between different aircraft rotations and crew pairings (allocations to aircraft) in the schedule. Even when resource connections are not present, a crew's pairing can become infeasible if the crew is delayed enough so that their maximum flying time would be exceeded. If crews are absent or delayed, reserve crew can be used to restore the schedule's feasibility. Most airlines therefore also schedule reserve crew in addition to their regular crew.

Previous work on modelling delay propagation was discussed in Section 2.5.2, AhmadBeygi et al. [3] introduced the concept of delay propagation trees to track how delays propagate through an airline's schedule. In contrast, this work uses a probability matrix to model the propagation of crew-related delay through an airline's schedule, the probability matrix is then used to evaluate the quality of candidate reserve crew schedules in terms of delay minimisation.

7.2 Model overview

The model introduced here is concerned with crew-related delay propagation in a single hub and spoke network. The aim of the model is to assign the standby duty start times for a fixed number of reserve crew duties such that the total expected crew-related delay propagation is minimised.
The proposed method involves a simulation parameter generation phase used to derive probabilities of crew-related delay and their associated expected delay durations. The parameter generating simulation simulates recovery from delays by searching for crew and aircraft swaps that absorb delays. The probabilities of crew-related delay are therefore independent of the effects of reserve crew and dependent on resource swaps. Assuming that swaps are a cheaper recovery action and are preferred over the use of reserve crew, further reductions in delays can then be achieved by optimising the reserve crew schedule. The parameters generated in the simulation phase are stored in matrices which record the causal relationship of propagating crew-related delays. The parameter matrices are then used as the inputs for the probabilistic evaluation procedure which evaluates the effect that a given reserve crew schedule has on the expected level of crew-related delay.

The evaluation procedure is then used to search for a reserve crew schedule that minimises the total expected crew-related delay. The experimental results indicate that the proposed model minimises crew-related delay in comparison to a variety of alternative methods of reserve crew scheduling for the problem instances considered later on in this chapter.

Chapter structure

The outline of this chapter is as follows. Table 7.1 defines all of the notation used in the model. The CDM is formulated in Section 7.3. Section 7.4 sets out an experiment to test the methods put forward and gives the experimental results with interpretations. Section 7.5 concludes with a summary of the main findings of this chapter.

P	:	Matrix of probabilities of crew-related delay
$p_{i,j}$	:	Element i, j of the matrix P. The probability that departure i is delayed by crew-related delay propagating from flight j
$p_i = \sum_{j=1}^{ND} p_{i,j}$	:	The total probability that departure i is delayed due to crew
Q	:	Hard copy of P after phase 1 (crew-related delay probabilities independent of the effects of a reserve crew schedule)
L	:	Matrix of mean crew-related delays
$l_{i,j}$	:	Element i, j of the matrix L. The mean crew-related delay of departure i propagated from flight j
D_j	:	List of departures with a non-zero probability of crew delay *originating in* flight j
E_i	:	List of departures with a non-zero probability of crew delay *propagating to* flight i
NS	:	Number of simulations used for input parameter derivation
ND	:	Number of departures
CT	:	Cancellation threshold
DT	:	Delay threshold
DL	:	Crew duty length
MS	:	Minimum sit period (rest time) for crew between consecutive flights
TT	:	Minimum aircraft turn time between consecutive flights
R	:	Number of reserve crew
X	:	Reserve crew schedule vector. Decision variable
x_k	:	Element k of the reserve crew schedule vector X. Duty start time of reserve k
Dep_i	:	Scheduled departure time of flight i
Arr_i	:	Scheduled arrival time of flight i
A_i	:	Aircraft scheduled to flight i
C_i	:	Crew scheduled on flight i
$ceta_n$	:	Estimated time of arrival of crew n
LDN_n	:	Last flight assigned to crew n
$aeta_r$	:	Estimated time of arrival of aircraft r
a_i	:	Relative importance of flight i in objective function evaluation
cd_i	:	Crew-related delay experienced by flight i
lcd_n	:	Crew-related delay experienced by the previous flight that crew n operated
r_k^i	:	Probability of reserve k being available to cover crew-related delay affecting flight i

Table 7.1: Notation

7.3 The probabilistic crew delay model

This approach for reserve crew scheduling consists of three sequential steps: (1) *input generation* through *simulation*, (2) *probabilistic crew delay optimisation*, and (3) *validation*. Phase 1 estimates the delay probabilities and expected durations for each individual flight in the schedule using simulation and assumes that no reserves are available for recovery. This information is then used as input for the CDM (Phase 2), which generates a reserve schedule that minimises crew delay for the given inputs. The resulting reserve schedule is validated through simulation in phase 3, during which reserves can be used for recovery. The individual components of the approach are described in more detail in Sections 7.3.1, 7.3.2, and 7.3.3, respectively.

7.3.1 Phase 1: Input Generation

Simulation is used to generate two matrices P and L that contain, respectively, the delay probabilities and the expected delay durations for each individual flight in the schedule. The two-dimensional structure of matrices is used to model the correlations that exist between the crew-related delays that are experienced by each flight. An element $p_{i,j}$ of the matrix P stores the probability that flight i experiences a crew-related delay that was propagated from an earlier flight j. Similarly $l_{i,j}$ stores the average duration of the crew-related delay that was propagated from flight j. Typically the matrices P and L will have a lower triangular structure, which is because flights are usually listed in scheduled departure time order and delay typically propagates from earlier scheduled flights to later ones. In this chapter, the matrices P and L are derived during a simulation learning phase. After this, P and L are used as input data for a model that evaluates the effect that any given reserve crew schedule has on reducing the expected level of crew-related delay in an airline's schedule. The matrix structure of P and L is used to evaluate the effect that any given reserve crew schedule will have reducing the probabilities that crew-related delays will occur, as well as any knock-on effects that may result from such delays. In summary, the matrix structure allows the modelling of the propagation of crew-related delays. The evaluation model is then used within a number of search heuristics that are used to find crew-related delay minimising reserve crew schedules.

The simulation model that is used to learn P and L uses delay distributions which were derived from historic data and applies aircraft swaps and crew swaps (but no reserves) to recover from disruptions. Crew-related delay can be divided into two categories: *root delays* where the cause cannot be traced back to preceding flight(s) and *propagated delays* which can be traced back to preceding flights. In the simulation, each flight i in the schedule has two variables, cd_i, which is the crew-related delay affecting flight i and lcd_{C_i}, which is the crew-related delay experienced by the previous flight flown by the crew assigned to flight i (crew C_i). So for flight i, lcd_{C_i} represents the root crew-related delay and cd_i represents the crew-related delay that propagated and affected flight i. For any two flight legs j followed by i which are operated by the same crew, $lcd_{C_i} = cd_j$ (i.e. these are the two ways of referring to the same crew-related delay from different perspectives, crew perspective and flight number perspective respectively). If $cd_i > 0$ (flight i is delayed by crew) and $lcd_{C_i} > 0$ (the assigned crew C_i caused a delay on their previous flight), then at least some of the crew-related delay is propagated delay. The exact value of cd_i is defined by Equation 7.1, and is equal to the delay that can be attributed to connecting crew discounting any delay due to the connecting aircraft and not including delays which are less than the delay threshold. The value of cd_i is equal to the delay that could be avoided/absorbed if a reserve crew were available for flight i. To reiterate, during phase 1 all other recovery actions except for reserve crew are considered and that the calculation of cd_i is performed after all other recovery actions have been considered and applied. This ensures that further reduction in delay is likely to be only achieved through the consideration of reserve crew use.

$$cd_i = max\left(0, ceta_{C_i} + MS - max\left(aeta_{A_i} + TT, dep_i + DT\right)\right) \tag{7.1}$$

If cd_i exceeds zero during the simulation learning phase, it means that a crew-related delay has occurred and the matrices P and L are updated to reflect this. Whenever a crew-related delay occurs ($cd_i > 0$), the corresponding entries in P and L are updated using Algorithm 7. The probability weighting of each update to P and L is a, where $a = 1/NS$ (NS = the number of repeat simulations used in the learning phase). Algorithm 7 determines whether or not, and if so how much of the crew-related delay can be attributed to crew-related delay that was propagated from an earlier flight. If a crew-related delay cannot be attributed to crew-related delay that was propagated from an earlier flight, then the crew-related delay is a root delay. Root delays are modelled by the diagonal elements of P and L (i.e., $p_{i,i}$ and $l_{i,i}$). Conversely, if the crew-related delay (or a portion of it) was propagated from an earlier flight then such delays are modelled by the off-diagonal elements of P and L (i.e. $p_{i,j}$ and $l_{i,j}$ with $i \neq j$). $p_{i,j}$ gives the probability that flight i suffers crew-related delay propagated from flight j, whilst $L_{i,j}$ gives the expected duration of the crew-related delay propagated from flight j to flight i.

Algorithm 7 Procedure for populating P during simulation

```
if cd_i > 0 then
    if (cd_i − lcd_{C_i}) > 0 then
        p_{i,i} = p_{i,i} + a ((cd_i − lcd_{C_i}) / cd_i)
        p_{i,j} = p_{i,j} + a (lcd_{C_i} / cd_i)
    else
        p_{i,j} = p_{i,j} + a
    end if
end if
```

Algorithm 7 states that when a crew-related delay occurs at flight i, flight i can only be identified as a root cause of crew-related delay if the crew-related delay of flight i exceeds that of the crew-related delay of flight j (lcd_{C_i}). Otherwise all of the crew-related delay is propagated crew-related delay. For example, if $p_{i,j}$ is equal to 0.01 ($i \neq j$), flight i has 1% chance of suffering crew-related delay propagation caused by a crew-related delay that propagated from flight j. This implies that flight j itself also suffered from a crew-related delay (i.e. $cd_j \neq 0$ and $p_{j,j} \neq 0$). Finally, if $lcd_{S_i^C} \leq cd_i$, not all of the crew-related delay of flight i can be can be traced back to the crew delay of flight j (cd_j). In summary, Algorithm 7 attributes crew-related delay probabilities according to the relative amounts of root and propagated delay crew-related delay minutes.

7.3.2 Phase 2: Probabilistic Crew Delay Optimisation

The CDM is used as an evaluator in search algorithms to generate a reserve crew schedule that minimises the probabilistic crew delay for the delay probabilities and expected delay durations in P and L. The following assumptions are mostly simplifications of those stated in Section 4.2, and are given in curly braces below. It is assumed that: (a) reserves are based at the hub station only, $\{RP7\}$; (b) reserves have a zero response time,$\{RC8\}$; (c) each reserve crew can cover exactly one disrupted crew per duty, $\{RC9\}$; and (d) each flight requires exactly one team of crew, $\{C2\}$. Furthermore, it is assumed that, if a disruption occurs, reserves are allocated in earliest start time order, $\{RP5\}$. This maximises the remaining reserve crew duty time. Reserve crew are assumed to be used as demand occurs in earliest start time order (default policy) and never held in anticipation of larger crew-related delays. Holding policies are considered in Chapters 8, 9 and 10.

Objective Function

The objective function used in the CDM quantifies the effects that a reserve crew schedule has on a disruption by iteratively calculating the probability that a reserve crew is still available (i.e., that they have not been used to handle previous disruptions) and can be used to cover a given crew disruption. Let p_i denote the probability that departure i ($i \equiv x_k+m$ in Algorithm 8) is delayed (sum of row i), and r_k denote the probability that reserve k is available to cover a crew-related delay. The probability that the reserve is still available to cover the next flight is then given by Equation 7.2.

$$r_k^{i+1} := r_k^i(1 - p_i) \tag{7.2}$$
$$p_i := p_i(1 - r_k^i) \tag{7.3}$$

Equations 7.2 and 7.3 underpin Algorithm 8, which calculates the objective value of a reserve schedule (the expected total weighted crew-related delay propagation for flights when the reserve schedule is enacted) for the delay probabilities and delay durations given by P and L. Notice also the similarity between Equations 7.2 to 7.3 and those of Equations 5.1 to 5.2 in Chapter 5. The only difference is that in this case p_i corresponds to the sum of row i of the matrix P, which includes the sum of the probabilities that flight i suffers crew-related delay which are propagated from previous flights.

Algorithm 8 Objective function evaluation

1: $P = Q$
2: **for** $k = 1$ to R **do**
3: $\quad r_k^{x_k} = 1$
4: $\quad m = x_k$
5: $\quad$ **while** $Dep_m < Dep_{x_k} + DL$ **do**
6: $\quad\quad$ **if** $Arr_{LDN(C_m)} \leq Dep_{x_k} + DL$ **then**
7: $\quad\quad\quad r_k^{m+1} = r_k^m(1 - p_m)$
8: $\quad\quad\quad$ Evaluate the immediate benefit of reserve ($immedEval(r_k^m, m)$)
9: $\quad\quad\quad$ Evaluate knock-on effects of reserve ($knockOnEval(r_k^m, m)$)
10: $\quad\quad$ **end if**
11: $\quad\quad m = m + 1$
12: $\quad$ **end while**
13: **end for**
14: $objVal = \sum_{i=1}^{ND} \sum_{j=1}^{ND} a_i p_{i,j} (l_{i,j})^b$

Algorithm 8 iterates through the R reserves in the schedule (X) in start time order (line 2), one at a time. Just as in Chapter 5, the order in which reserve crew are considered reflects the assumed policy for the order in which reserve crew are used. In this case the assumed reserve policy is to use reserve crew in earliest start time order first. For each reserve crew, the initial probability of availability is initialised to 1 (line 3). Lines 5 to 12 iterate through all flights that can be covered by the given reserve crew feasibly within the reserves duty length (lines 5 and 6). Line 7 updates the probability that reserve k remains available for subsequent crew-related delays. Line 8 evaluates the effect that the probability that reserve k remains available has on the probabilities that flight m still experiences crew-related delay, i.e. the immediate impact of the reserve (see Algorithm 9). Line 9 evaluates the knock-on effects that using the reserve crew to absorb crew-related delay affecting flight m has on future crew-related delays. Algorithm 10 is the procedure for line 9. The objective value for the entire schedule is calculated on line 14, and is equal to the weighted sum of the expected crew delay durations. The weight a_i denotes the relative importance of flight i and may be derived from factors such as passenger numbers or the availability of alternative flights for re-routing passengers.

The *immedEval* (Algorithm 9) and *knockOnEval* (Algorithm 10) procedures used in objective function evaluation (Algorithm 8) calculate how the probabilities of crew-related delays are reduced by a given reserve crew schedule. These reduced probabilities of crew-related delay are used in line 14 to calculate the weighted sum of expected crew-related delay which is to be minimised by manipulating the reserve schedule (X).

Algorithm 9 Procedure for calculating the effect a reserve crew has on the probability that flight i is delayed for crew-related reasons

$immedEval(r_k^i, i)$
for $j = 1$ to $|E_i|$ **do**
$\quad p_{i,E_{i,j}} = p_{i,E_{i,j}}(1 - r_k^i)$
end for

Algorithm 9 evaluates the effect that a reserve crew k, available with probability r_k^i, has on the probability that crew-delay propagated from a previous flight or a root crew-related delay still delays flight i. E_i is a list of flights with a non-zero probability of propagating crew-related delay to flight i.

The purpose of accounting for knock-on delays is to determine the effect that a reserve crew member who is available for flight j has on future delays. The probability of the crew-related delay of future flight i originating at flight j is reduced proportionally to the probability that a reserve is available for flight j and to the probability that flight i is delayed due to flight j. The assumption is made that the probability that flight i is delayed due to delay propagated from flight j is proportional to the ratio of the expected duration of the delay that propagates from flight j to flight i and the total expected delay that propagates from flight j. E.g. if flight 1 has an expected delay of 1 hour and flight 2 which suffers delay propagated from flight 1 has an expected delay of an hour, this implies that delay effecting flight 1 is directly proportional to the delay experienced by flight 2. This reasoning is applied recursively in Algorithm 10.

Algorithm 10 Procedure for calculating the effect a reserve crew has on the probabilities of future crew-related delays still occurring

1: $knockOnEval(r_k^j, j)$
2: **for** $i = 1$ to $|D_j|$ **do**
3: $\quad knockOnEval\left(r_k^j\left(\frac{p_{(D_{j,i}),j} \times l_{(D_{j,i}),j}}{\sum_{k=1}^{|D_j|}\left(p_{(D_{j,i}),k} \times l_{(D_{j,i}),k}\right)}\right), D_{j,i}\right)$
4: $\quad p_{(D_{j,i}),j} = p_{(D_{j,i}),j}(1 - r_k^j)$
5: **end for**

The outer loop in Algorithm 10 considers each flight i ($\in D_j$) that has a non-zero probability of experiencing a delay propagated from flight j. Within the loop the recursive call of the procedure is made. The probability that a reserve prevents delay propagating to a subsequent flight is reduced proportionally to the probability that the reserve is used and the probability that the delay would have propagated. Line 4 in the procedure evaluates the probability that a crew-related delay still occurs given that the reserve is available at flight i.

Search Algorithm

In this chapter, the model is solved using a greedy heuristic which proceeds by adding one reserve crew to the schedule at a time, choosing a duty start time (discretised according to scheduled departure times) such that the reduction in the objective value is maximised. A local search approach is also considered, where the neighbourhood structures which are used include single swap, cut and insert and "power set shift". The "power set shift" neighbourhood is defined as each possible lateral shift of each substring of a binary string solution bounded by 1's such that the shift does not overwrite any surrounding 1's. The results are shown in Section 7.4.

7.3.3 Phase 3: Validation

Simulation was also used to validate the CDM described above, this time considering reserves as one of the recovery actions. In contrast to the simulation of Chapter 4 each departure from the hub station is modelled as an out-and-back cycle, so from the hubs perspective each departure results in an arrival after visiting a single spoke station. This model therefore uses aggregated block time distributions for each spoke destination to account for the time between leaving the gate at the hub to the time the aircraft arrives back at the gate of the hub.

7.4 Results

7.4.1 Data Instances

The schedules considered here cover a time span of 24 hours, contain 300 flights carried out by 37 aircraft and approximately 120 teams of crew. To simplify the analysis, all aircraft are assumed to be of the same fleet type. All instances were generated using a custom developed instance generator that is underpinned by real world data (described in Section 4.5.1).

This model is designed for airline schedules that contain a risk of crew-related delay propagation. So airline schedules that contain a risk of crew-related delay are required to show the method's effectiveness. An airline schedules inherent risk of delay increases with the number of aircraft changes by crew and with decreasing crew connection times. Without either of these characteristics, delays are a direct consequence of other uncertain events.

In order to vary the level of risk of crew-related delay propagation, the schedule instance generator has two parameters that can be altered. The first $OnTime\%$ chooses the allocated block time (gate to gate) such that the specified percentage of flights according to journey time distributions are completed on time. Aircraft routings are generated by assuming that each aircraft serves exactly one remote destination and shuttles back and fourth with a ground time equal to the minimum between each flight leg. As a result $OnTime\%$ effectively determines the aircraft routing and all scheduled departure and arrival times. Increasing the value of $OnTime\%$ has the effect of increasing the chance connecting crew will be able to make the connection without causing a delay to the waiting aircraft. The second parameter $PofAC$ (probability of aircraft change) is the rate of aircraft changes in crew schedules and also controls the risk of crew-related delay in an airline schedule.

25 schedules were generated using each pairwise combination of the following parameter sets. $PofAC = [0, 0.1, 0.2, 0.3, 0.4]$ and $OnTime\% = [55, 60, 65, 70, 75]$. These 25 schedules are used in Section 7.4.2 and Section 7.4.4. These parameters were chosen to generate airline schedules with a risks of crew-related delay ranging from low to high. The variety of test instances will also give some degree of confidence that the method will work in a variety of real world situations.

7.4.2 Convergence

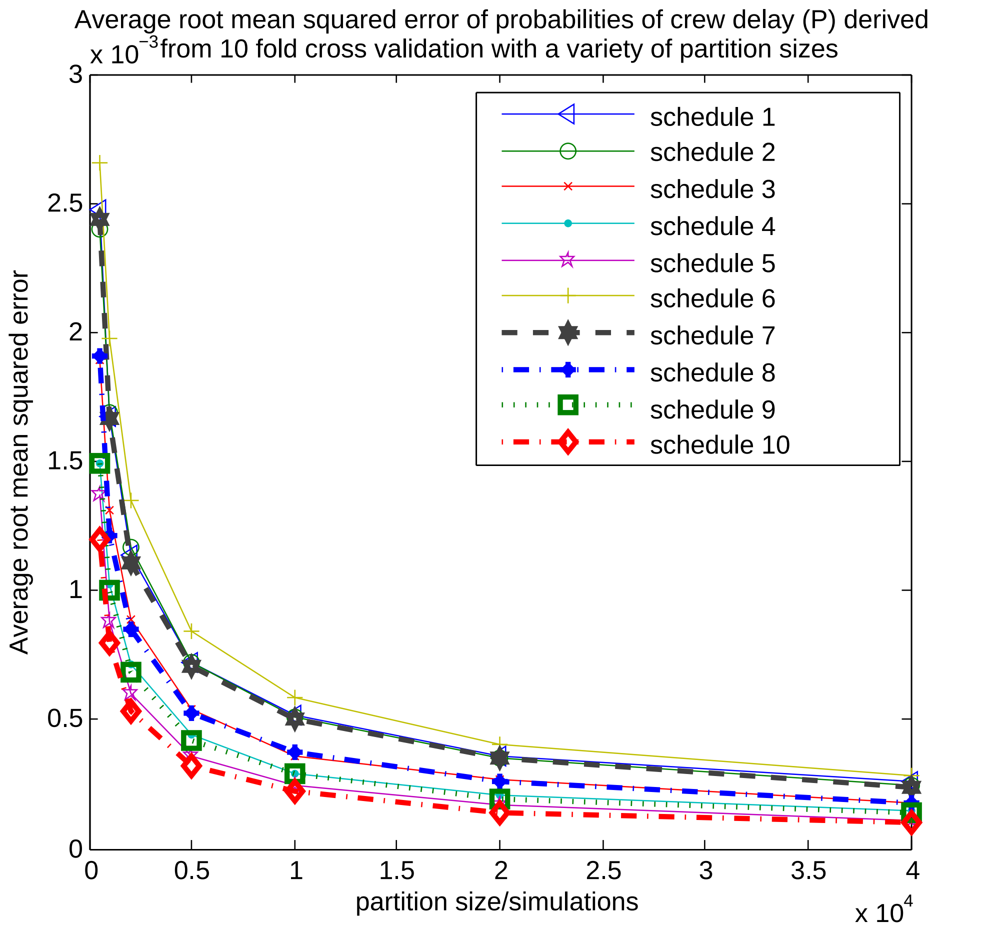

Figure 7.1: Average root mean squared error of P elements derived from simulation using 10-fold cross validation for a variety of sample sizes

Since P and L were approximated using simulation in phase 1 (§7.3.1), it was necessary to ensure that sufficient iterations had been performed to allow these matrices to converge. In the following, the overall convergence of the elements of P and L are considered. K-fold cross validation [47] was used to find the appropriate number of repeat simulations to train P and L. A partition (a fold) in this case corresponds to a P or L matrix derived from a number of simulations, the number of simulations defines the sample size. The results are illustrated in Figure 7.1, showing the average root mean squared errors over all departures from a sample of 10 schedule instances. Seven different simulation sample sizes were considered.

It can be observed from Figure 7.1 that the average root mean squared error for P falls to approximately 0.2×10^{-4} for a sample size of 40,000. Full convergence to zero may require a number of simulations many orders of magnitude larger than the sample sizes considered here. The reason is that the number of reachable simulation configurations or states is excessively large. Based on a trade-off between parameter accuracy and the time required to run the simulations, 20000 simulations was considered to be sufficient to derive useful estimates of P and L for any given schedule instance. L has a very similar convergence rate to P.

7.4.3 Accuracy of the model

The 25 problem instances generated above were solved using the three-phase approach described in Section 7.3. The resulting objective values, i.e. the expected crew-related delay propagation, were recorded for each. The reserve schedules that were generated were simulated to estimate the total expected crew-related delay. The values of the simulation model were then compared against the objective function values obtained by the CDM. The results are shown in Figure 7.2. In order to conclude that the CDM accurately predicts the crew-related delay, there should be a linear relationship with gradient 1 and intercept 0 between the values obtained by the CDM and the simulation. It can be observed from Figure 7.2 that a nearly perfect linear relationship with gradient close to 1 and intercept slightly above zero was found. This shows that the CDM is consistent and accurately predicts how reserves are used, as well as the impact they have.

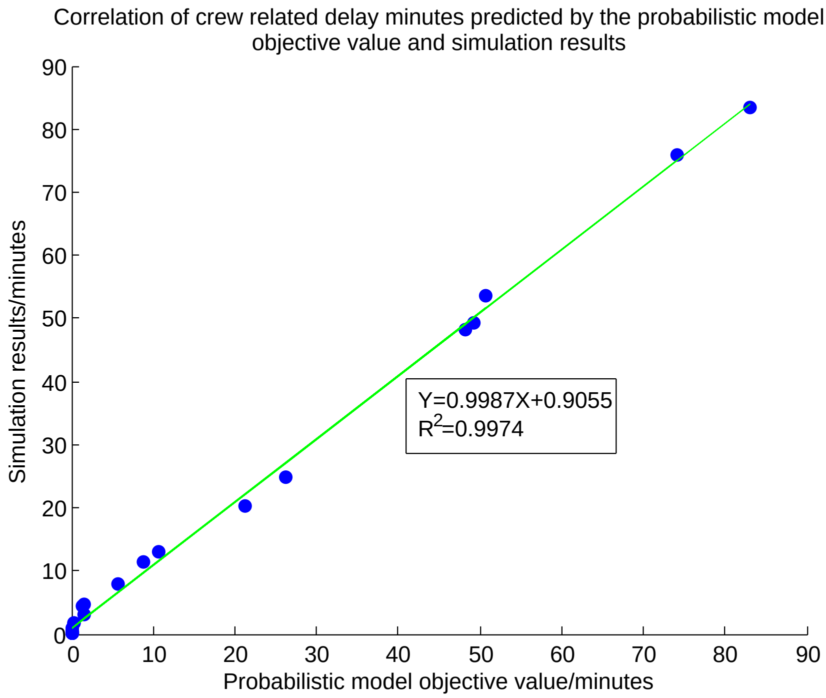

Figure 7.2: CDM predicts crew-related delay observed in validation simulations

7.4.4 Comparison with other approaches

To validate the CDM as an effective approach for scheduling airline reserve crew taking into account other recovery actions and operational uncertainty, a comparison was made between the approach presented here and other alternative approaches (briefly introduced below). Simulation was used to validate the reserve crew schedules obtained by each of the individual approaches and to derive performance measures that reflect their quality. The performance measures of interest are: the cancellation rate; the reserve utilisation rate; the average crew-related delay propagation; and the average total delay propagation. Approaches which required simulation derived input parameters were limited to 20000 runs to determine a reserve schedule for a given planned day of operations. For each schedule, 8 teams of reserve crew were to be scheduled.

Models

Various models are evaluated including variants of the CDM and two other types of method that are not probabilistic.

Probabilistic models: Different variations of the CDM were developed and compared to the model introduced above. The first two variations of the approach, denoted by *Prob 1* and *Prob 2*, have a delay exponent (b in Algorithm 8) equal to 1 and 2, respectively. The idea is that having a delay exponent that is greater than 1 will lead to reserve crew schedules that provide more coverage for departures associated with larger expected delays. *Prob** is the same as *Prob 1* except that Algorithm 10 (which accounts for knock-on effects) is removed from Algorithm 8. The purpose of this is to verify whether this aspect of the model, when applied in a search method, results in improved reserve crew schedules. In the last variation of the CDM, denoted by *Prob LS*, the greedy algorithm used in *Prob 1* is replaced with a local search algorithm, in order to determine the influence of the search algorithm on the quality of the resulting reserve crew schedules.

Area under the graph: Area 1. This simulation based approach was described in Section 4.7.1.

Iterative area under the graph: Area 2. This approach is an iterative variant of *Area 1.* The reserves are used in the simulation to iteratively derive when reserves are likely to be needed. This approach is divided into stages where each stage derives a reserve schedule (as in *Area 1*) based on the results from the previous stage. The next stage uses the new reserve schedule to derive more information about when reserves are required. A possible pitfall of this approach is that it may not converge towards a good or optimal solution or it may not converge at all.

Uniform start rate: Uniform. This solution approach starts reserve crew duties at equal times intervals throughout the scheduled day of operations. Described in Section 3.5.3.

Observations

The simulation which was used to validate the approaches described above implements a recovery policy which first considers crew and aircraft swaps, and subsequently reserve crew in conjunction with aircraft swaps. The validation simulation also assumes that once reserves are used they are treated as regular crew and once regular crew are replaced by reserves, the regular

Solution method	Cance-llation rate	Reserve utili-sation rate	Avg. crew delay (mins)	Avg. total delay (mins)	Prob. of delay > 30 min	Solution time (sec)
No res	4.18E-6	0.0000	0.3465	2.0846	2.58E-3	0
Prob 1	2.20E-6	0.7890	0.1394	1.5511	7.95E-4	1226
*Prob **	2.40E-6	0.7863	0.1414	1.5735	7.95E-4	1219
Prob 2	2.41E-6	0.7797	0.1407	1.5769	7.83E-4	1232
Prob LS	2.51E-6	0.7894	0.1399	1.5528	8.04E-4	1400
Area 1	2.41E-6	0.7332	0.1542	1.6383	8.22E-4	1215
Area 2	2.43E-6	0.7331	0.1537	1.6376	8.23E-4	1216
uniform	2.35E-6	0.6771	0.1827	1.5420	1.29E-3	1

Table 7.2: Simulation derived performance measures for a variety of solution methods

crew cannot be used as reserves. However, the simulation is flexible enough to allow for changes to these assumptions. Table 7.2 shows cancellation rates, reserve utilisation rates average crew delays, average delays, probabilities of crew delays over 30 minutes, and solution times for 8 methods of reserve crew scheduling. The results are based on 20000 repeat simulations for each of the 25 schedule instances for each method. This means that each method is tested on 500,000 days of operations with a total of 150,000,000 simulated departures. The first row *No res* shows the results corresponding to no scheduled reserves. The results show that cancellations due to delays are very rare and that *Prob 1* minimises cancellations and gives the lowest average crew-related delay. *Prob LS* gives the highest reserve utilisation rate. *Prob 1* and *Prob 2* have the lowest probabilities of crew delays over 30 minutes with *Prob 2* having the lowest, this can be attributed to the delay exponent of 2 used in the objective function. All of the methods that are based on simulation (*Prob 1,*,2,LS* and *Area 1,2*) took over 20 minutes to find reserve schedules for all of the 25 schedule instances, almost all of this time is spent running simulations to derive parameters for the respective models. *Prob 2* took slightly longer than *Prob 1*, the only explanation for this is that the presence of the delay exponent of 2 in the objective function made the objective function computationally more demanding. The *Prob ** results indicate that removing the *knockOnEval* procedure given in Algorithm 10 from function evaluation (Algorithm 8) results in marginally higher average crew delay and total delay however the cancellation rate is reduced, however cancellations are negligible in all cases.

In general the *CDM* based approaches were most effective. Of the four variants of the *CDM* the method *Prob 1* dominated the other variants on nearly all performance measures considered including the main objective criterion used in this investigation which is to minimise crew-related delay. Appendix A gives the results for *Prob 1, Area 1 and Uniform* for each individual airline schedule instance and each performance measure, they show that the averaged results of Table 7.2 are representative.

7.5 Chapter summary

A probabilistic model for scheduling airline reserve crew in anticipation of crew-related delays was presented. The CDM takes journey time uncertainty into account as well as the availability of other recovery actions. Simulation provides the mechanism which makes this possible. The CDM is also able to anticipate the future impact of the reserve crew it schedules in terms of the absorption of crew delays that are likely to propagate further in a schedule. The method was tested over a range of schedule instances in which the likelihood of crew-related delay propagation was controlled with schedule generation parameters. It was shown that the CDM accurately models the expected crew-related delay associated with a given flight schedule and reserve crew schedule combination. In comparison with a range of alternative approaches to reserve crew scheduling the CDM proved most effective overall and using a delay exponent in the objective function greater than 1 leads to reserve crew schedules that minimise the probabilities of longer crew delays. Additionally, results are given which suggest that the recursive procedure (Algorithm 10) for factoring in the effect of knock-on crew delays absorbed by reserves scheduled previously appears to reduce average delays at the expense of a marginal increase in cancellation rate. The solution times are also reasonable from a practical point of view.

Chapter 8

Statistical delay propagation model

In this chapter a statistical delay propagation model ($SDPM$) is introduced. This model is motivated by the 1^{st}, 3^{rd} and 7^{th} bullet points of Section 3.2, which are repeated here. Firstly, *reserve crew demand is influenced by journey time uncertainty*, this is important because crew who are delayed on a connecting flight can be replaced with reserve crew. Secondly, *other recovery actions may reduce reserve crew demand*, this is important because swap recovery actions may be available which make the use of reserve crew unnecessary. Thirdly, *the structure of an airline's schedule dictates how disruptions may propagate through the schedule*, this is important because the scheduled departure times and the scheduled connections of all individual crew and aircraft determine which flights suffer direct or indirect knock-on effects from any given disruption. The $SDPM$ accounts for these factors as it is a model of how delay-uncertainty propagates through an airline's schedule which also models the effects of airline recovery actions for delays.

The $SDPM$ is (just as the models of Chapters 5, 6 and 7) designed as an evaluator of reserve crew schedules. A reserve crew schedule can have a significant influence on how delay propagates. Firstly, reserve-induced delays (see Section 6.1.3) have the potential to propagate. Secondly, reserve crew teams have the potential to prevent delays which may have a high chance of propagating to subsequent flights (see Chapter 7).

Early in this project, attempts were made at implementing a statistical delay propagation model (See Figure 1.1). The initial attempt was abandoned due to difficulties that occurred when trying to model swap recovery actions in such a framework. This difficulty is overcome in this chapter by explicitly allowing for the possibility that crew and aircraft can be assigned to different crew pairings and aircraft routings.

The $SDPM$ supersedes the CDM of Chapter 7, which was an attempt to circumvent the difficulty of modelling the effects of swap recovery actions using a statistical delay propagation approach. The $SDPM$ has several advantages over the CDM, including: it does not require a simulation learning phase, so it is highly convenient (fast) for use as an online reserve holding policy; it accounts for aircraft, crew and reserve-induced delays as opposed to crew-related delays only; and it allows for crew unavailability, because it takes as input the probabilities that crew absences are covered by reserve crew, which are an output of the CAM of Chapter 6. So the $SDPM$ also integrates the separate models for crew absence and crew delay uncertainty.

As described in Section 2.5.2 approaches that are conceptually similar to the $SDPM$ have been explored previously. Berger et al. [21] use a similar approach to calculate departure and arrival time distributions for a rail network. However, the problem specific details of the problem tackled in this work and that of Berger et al. make the two approaches structurally very different. Berger et al. calculate departure distributions as a function of train delays and waiting time rules for connecting passengers, whereas this work calculates departure time distributions as a function of aircraft delays, connecting crew delays and airline recovery actions (including swaps and reserve crew use). Berger et al. evaluate their model in terms of delay prediction accuracy, whereas this work considers scheduling and decision support applications in addition to delay prediction accuracy.

Chapter structure

The remainder of this chapter is structured as follows. Table 8.1 defines the notation used by the $SDPM$ model. Section 8.1 introduces the $SDPM$. Section 8.2 gives experimental results for the $SDPM$. Section 8.3 describes how the $SDPM$ extends to case of multiple fleet types, crew ranks and qualifications. Section 8.4 concludes with a summary of the main findings from this chapter.

8.1 The Statistical Delay Propagation Model

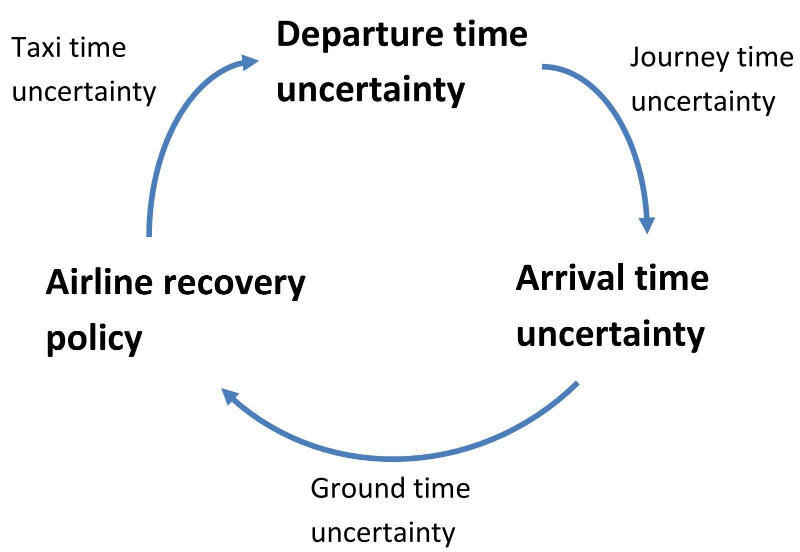

Figure 8.1: Delay propagation cycle

Journey times have inherent uncertainty which can be captured in the form of statistical distributions (see Section 2.5.4). Journey times directly influence arrival times, and arrival times can influence the departure times of subsequent flights involving the same crew or aircraft. This is the case if there is insufficient slack time between the arrival from the previous flight and the scheduled departure time of the next flight. If resources split, e.g. the crew does not follow the aircraft, multiple flights will be affected (see Figure 2.2). Hence, journey time uncertainty leads to arrival time uncertainty and arrival time uncertainty, coupled with an airline recovery policy, leads to departure time uncertainty, which through journey time uncertainty leads back to arrival time uncertainty. Figure 8.1 depicts this delay propagation cycle. In the $SDPM$, this process is modelled in the form of delay distributions propagating through an airline's schedule, as illustrated in Figure 8.2. It shows how there may be no departure time uncertainty

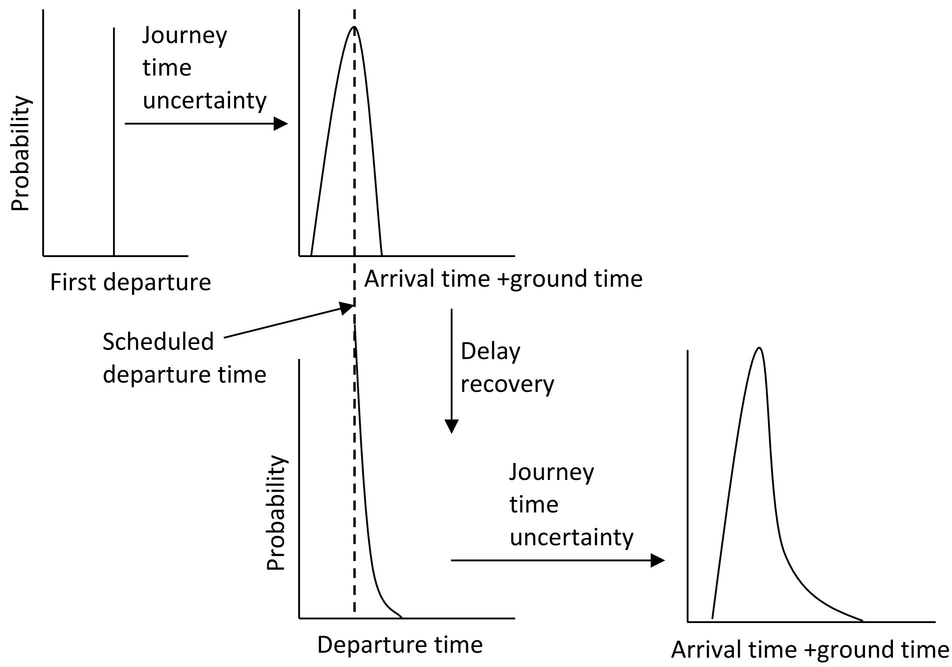

Figure 8.2: Propagating delay distribution

for the first flight of the day (top left). Then, after the first journey, the arrival time exhibits uncertainty (top middle) as a result of journey time uncertainty. Airline recovery actions may reduce the expected departure time of the next flight (bottom middle). The arrival time distribution of the second flight (bottom right) attains a complex form as it is a function of two journeys, scheduled slack, and airline recovery.

Hereafter, both aircraft and crew teams are sometimes referred to as airline resources and both crew pairings and aircraft routings as lines of flight. This is useful when discussing the details of the $SDPM$ because crew and aircraft are modelled in a very similar way.

8.1.1 Assumptions

On the day of operation, delayed arrivals can lead to delays of subsequent departures. The following assumptions are made regarding recovery actions for preventing delayed departures. The $SDPM$ structurally relies upon assumptions 1, 4, 5 and 7. The other assumptions represent details which can be easily be changed in the $SDPM$. For each assumption, the corresponding assumptions of Section 4.2 are given in braces. As a result of the matching assumptions between the simulation and the $SDPM$, the $SDPM$ is effectively a theoretical solution for the simulation of Chapter 4.

Assumption 1: *A preference order exists for the application of recovery actions for disrupted departures from the hub.* I.e. airline recovery actions are determined sequentially for disrupted flights according to a predefined preference order. In this work the earliest departure time order is used as the preference order. An extension might be to consider a preference order based on the relative financial importance of different flights, i.e. apply recovery actions for the high revenue disrupted departures first. {RP1 (*sequential recovery assumption*)}.

Assumption 2: *In the event of delays greater than a delay threshold of 15 minutes, swap recovery actions and reserve teams use are considered for replacing delayed crew.* A delay threshold of 15 minutes is based upon the on-time performance measure which is used by the FAA and which was also used by Sohoni et al. [97] to rate airline performance, however this is a parameter that can be changed. {RP2 (*delay threshold assumption*), RA1 (swaps can absorb delays assertion)}.

Assumption 3: *Feasibility for swap recovery requires that swapped crew can finish each other's duties within their respective duty shifts and can swap back to their original lines of flight before their next scheduled duty shifts.* Assumption 3 protects crew from swaps which can possibly lead to illegal overtime and swaps which cannot be undone before the beginning of their next scheduled duty shifts. {C9 (crew swap assumptions), L4 (line of flight swap assumptions)}.

Assumption 4: *For each departure there exists a preference order for the lines of flight which can be used for swap recovery actions.* Assumption 4 provides a preference order for the lines of flight from which resource swaps can be obtained. The probability calculations for resource swap recovery actions in the $SDPM$ require a preference order to calculate conditional probabilities. In this work, the largest common ground time is used as the preference order in which to consider swappable lines of flight. This approach provides the largest time window in which swap recovery actions can be implemented (Ageeva [2] actively scheduled overlaps to increase swap availability and as a result also increased schedule robustness). {RA6 (*swap recovery selection assumption*), L4 (line of flight swap assumptions)}.

Assumption 5: *A feasible resource swap is beneficial (reduces delay) if the replacement resource is available before the delayed resource and is not delayed for its own next scheduled departure, whilst at the same time the delayed resource does not delay the next flight of the replacement resource's line of flight.* Assumption 5 when combined with assumptions 1 to 4 for the case of a single swap, provides the necessary conditions for a swap which reduces overall delay. Section 8.1.4 discusses this assumption in more detail in relation to how the $SDPM$ models beneficial swap recovery actions. See Appendix C for a proof of a theorem based on assumption 5. {L4c (*later flights are delayed less by delayed resources assumption*), L4d (*the replacement resources must not be delayed for their own next scheduled flight assumption*)}.

Assumption 6: *Airline recovery occurs at the hub station and delays occurring at spoke stations are propagated.* Assumption 6 reflects the fact that the proposed $SDPM$ is geared towards a single hub airline where all flights from the hub station visit a spoke station and then return to the hub station. An extension of this assumption would be to include the modelling of recovery actions that may be available at spoke stations. {RP6 (*deadheading not viable for solving delay and unexpected crew absence disruptions assumption*), RP7 (*low spoke station flight volume assumption*)}.

Assumption 7: *The recovery action which best minimises delay, of those available, is the one that is implemented.* In the event of a tie, it is assumed that crew swaps are preferred to reserve crew team use, as those reserve crew can then be saved for later use. Tied crew swaps and/or aircraft swaps are broken using Assumption 4. {RA6 (*swap recovery selection assumption*)}.

Assumption 8: *Flights are cancelled if their departure is delayed beyond the flights cancellation threshold after the consideration of delay recovery actions.* This cancellation threshold assumption serves as a ceiling on delays over which a flight is cancelled. In this work a constant cancellation threshold of 3 hours for all flights is used. The value of 3 hours is the same value which was used by Rosenberger et al. [85] in their aircraft re-routing model. Another possibility would be to use different cancellation thresholds for different flights and even for different crew, such an approach could be used to allow for maximum working hour regulations, such as maximum flight hours per day, week, month and year rules. Appendix H considers a variable cancellation threshold formulation. {RP3 (*cancellation threshold assumption*)}.

8.1.2 Overview of the $SDPM$

The $SDPM$ follows a similar procedure to that of the simulation flow diagram of Figure 4.1. It considers each scheduled departure in turn. For each departure, the simulation determines a single deterministic departure time, whereas, the $SDPM$ discretises time and considers the probability that the departure takes place within each time interval between the departure time and cancellation threshold. The $SDPM$ considers all possible assignments of crew and aircraft which may enable the departure to take off within each time interval, including the probabilities of all possible swap recovery actions. Once the simulation has determined a departure time, a journey time is instantiated from the corresponding journey time distribution. Analogously, the $SDPM$ adds the entire journey time distribution to the departure distribution just derived (using the convolution procedure defined in Section 8.1.8). The resultant arrival time distribution is used for calculating departure time distributions for subsequent departures. In

ETA	:	Estimated time of arrival
C, (A)	:	Crew (aircraft) ETA matrices
$C_{i,j,t}$, $(A_{i,j,t})$	:	Probability element of the crew (aircraft) ETA matrix corresponding to crew (airframe) j assigned to crew pairing (aircraft routing) i arriving at the hub at time interval t
$\Box^-$	:	A matrix storing negative contributions of the corresponding elements of the probability matrix $\Box$. For example C^- stores the decrement of the corresponding elements of C as a result of those elements contributing to the probability that the departure under consideration takes place at a given time interval
$\Box^+$	:	A matrix storing positive contributions of the corresponding elements of the probability matrix $\Box$. For example C^+ stores the increment of the corresponding elements of C as a result of delayed resources being swapped to different lines of flight when calculating the probability that the departure under consideration takes place at a given time interval
T_d	:	true if hub departure d is a scheduled through flight, false otherwise
Dt_d	:	Scheduled departure time interval of hub departure d
Dc_d	:	Cancellation threshold time interval of hub departure d
St_d	:	Scheduled departure time interval of a spoke departure, that is the return flight for hub departure d
F_d, (E_d)	:	Crew pairing (aircraft routing) assigned to hub departure d
$G_{d,m}$, $(H_{d,o})$	:	m^{th} (, o^{th}) feasible crew team (airframe) for hub departure d
$J_{d,u}$, $(I_{d,s})$	:	u^{th} (, s^{th}) preferred swappable crew pairing (aircraft routing) for hub departure d
$L_{d,i}$, $(K_{d,i})$	:	The next hub departure of crew pairing (aircraft routing) i at the departure time of hub departure d
V_d	:	Destination of hub departure d
$\gamma_{t,j}$, $(\alpha_{t,j})$	:	Probability that crew team (aircraft) j is used for the given departure at departure time interval t
ϵ, (θ)	:	Probability of crew (aircraft) availability at a given time interval for a given departure
$\epsilon^{'}$, $(\theta^{'})$	:	Probabilities of crew (airframe) availability at a given time interval for a given departure excluding the probability that the available crew and airframe were already used at an earlier time interval for the given departure
β, (λ)	:	Probability of crew (airframe) swap availability at a given time interval for a given departure
$\beta^{'}$, $(\lambda^{'})$	:	Probability of crew (airframe) swap availability at a given time interval for a given departure excluding the probability the same crew (airframe) swaps are used at an earlier time interval for the given departure
$\phi_{\rho,\nu,\mu}$	:	Probability the ρ^{th} feasible crew assigned to the given crew pairing swaps with the ν^{th} feasible crew for the given departure which is assigned to the μ^{th} swappable crew pairing for the given departure excluding the probability the pairwise swap was used at a previous time interval for the same departure
$\sigma_{\nu,\mu}$	:	Probability element corresponding to the ν^{th} feasible crew for the given departure who is assigned to the μ^{th} swappable crew pairing and can adopt the given crew pairing at the given time interval to replace delayed crew
$\tau_{\rho,\mu}$	:	Probability element corresponding to the ρ^{th} feasible crew for the given departure being late and assigned to the μ^{th} swappable crew pairing at the given time interval
X	:	Reserve crew schedule
X_k	:	start time index of the k^{th} reserve scheduled to begin a reserve block
R	:	Number of reserve crew in a given reserve crew schedule
CT	:	Cancellation threshold (maximum delay before a flight is cancelled)
DT	:	Delay threshold (minimum delay for which delay recovery actions are considered)
CM_d	:	Cancellation measure of delay affecting departure d
$f(t,d)$	:	Denotes the function used to convert a delay to a measure of cancellation, which is a function of the departure time interval t and the departure d under consideration. Returns 1 for delays exceeding the cancellation threshold (as these are cancelled) and 0 for delays below the delay threshold
b	:	Delay exponent used in the delay cancellation measure function
W	:	Width of time intervals used in the ETA matrices
n	:	Number of hub departures in the airline schedule

Table 8.1: Definitions

effect, the $SDPM$ simulates all possible simulation outcomes in a single evaluation.

Algorithm 11 Outline of the $SDPM$

```
1: Inputs: Airline schedule, airline delay recovery policy, minimum
   ground times, journey time distributions, probabilities of crew avail-
   ability
2: Outputs: Departure distributions, probabilities of flight cancellations
3: ObjVal = 0
4: for each hub departure d do
5:   CAM (Chapter 6) provides the probabilities that crew are available
     at different time intervals after replacing absent crew with reserve
     crew and the probabilities that reserve crew teams are available to
     replace delayed crew at different time intervals (Section 8.1.4)
6:   for each time interval t between the departure time and the cancel-
     lation threshold of departure d do
7:     Calculate the probabilities for the availability of crew and aircraft
       (Section 8.1.5)
8:     Calculate the cumulative probability cumuP of departure and the
       probability incP that the departure takes place at time interval t
       (Section 8.1.6)
9:     ObjVal = ObjVal + (incP × f(t, d)) (delay contribution)
10:    Update the departure time distributions (Section 8.1.7)
11:  end for
12:  ObjVal = ObjVal + (1 − cumuP) (cancellation contribution)
13:  Provide the CAM with the probabilities that reserve crew teams were
     used to replace delayed crew (Section 8.1.4)
14:  Update the ETA matrices for resources used for departure d (Section
     8.1.7)
15:  Apply journey time uncertainty to the departure distributions to up-
     date the ETA matrices ready for the next departure (Section 8.1.8)
16: end for
```

Algorithm 11 gives an overview of the $SDPM$ when it is used as an evaluator of reserve crew schedules, to derive an objective value ($ObjVal$) for a given reserve crew schedule. $ObjVal$ is a measure of the expected delay and cancellation disruptions. For each hub departure d (line 4), the $SDPM$ calculates the probabilities of departure during each time interval t between the departure time and the cancellation threshold (line 6). To do this, the $SDPM$ must calculate the probabilities that crew and aircraft are available, including those obtained through swaps. These probabilities are calculated from the crew and aircraft ETA matrices (see Section 8.1.3) using the procedure explained in Section 8.1.5. The contributions to the crew ETA matrix corresponding to reserve crew used to replace absent crew and reserve teams available to replace delayed crew are derived from the CAM on line 5, this is explained in Section 8.1.4. On line 7, the cumulative probability that the departure takes place at or before the given time interval is calculated. How it is calculated is the subject of Section 8.1.6. The increment ($incP$) in the cumulative probability of departure at time interval

t, compared to time interval $t-1$, is the probability that the departure takes place during time interval t. This provides the weight for the delay contribution to the objective value for a departure at time interval t (line 9). The function $f(t,d)$, is referred to as the delay cancellation measure function, and maps the delay, due to departure d departing at time interval t, to a measure with the units of cancellations. This function is explained in Section 8.1.2. The next step of Algorithm 11 step 10 updates the departure distributions with respect to the probabilities that each crew and aircraft is used for the departure during time interval t. The details for this part of the procedure are the subject of Section 8.1.7. After calculating the probability of departure during each time interval, line 12 adds the probability that the flight is cancelled to the objective value, which is one minus the cumulative probability ($cumuP$) of departure. The reasons for flight cancellations are discussed in Section 8.1.4. On line 13, the $SPDM$ provides the CAM with the probabilities that reserve crew teams were used for departure d, this is explained in Section 8.1.4. On line 14, the $SDPM$ updates the ETA matrices to account for the crew and aircraft used for departure d. This part of the procedure is explained in Section 8.1.7. Line 15 applies journey time uncertainty to the derived departure distributions to update the ETA matrices ready for subsequent departures. The procedure for this is the subject of Section 8.1.8.

As indicated in Algorithm 11 the $SDPM$ requires input parameters which are obtained from the CAM (Chapter 6). The $SDPM$ and the CAM are therefore evaluated in a synchronised way, they each consider all scheduled hub departures in earliest departure time order. The $SDPM$ requires the CAM to be evaluated for each departure before the $SDPM$ considers each departure. Section 8.1.4 explains the points in both the $SDPM$ and the CAM at which information is exchanged between the two models. One way to view the relationship between the $SDPM$ and CAM is that the CAM is the reserve crew engine for the $SDPM$, because the CAM is devoted to the fine details of the use of individual reserve crew. In contrast, the $SDPM$ takes a much more aggregated approach to modelling crew, i.e. crew consist of indivisible teams which may be available at different times.

Delay cancellation measure

The delay beyond the delay threshold (DT), for departure d at time interval t, depends on the departure time interval (Dt_d). The actual delay duration depends on the width (W) of the time interval used in the $SDPM$. The delay for departure d at time interval t is given by Equation 8.1.

$$delay = \max\left(0,\ \left((t - Dt_d) \times W\right) - DT\right) \tag{8.1}$$

$$f(t,d) = \begin{cases} \left(\frac{delay}{CT-DT}\right)^b & \text{if } delay < CT \\ 1 & \text{otherwise} \end{cases} \tag{8.2}$$

To retain the simplicity of a single objective model, delay contributions to ($ObjVal$) the objective function are converted to an equivalent measure of cancellations, using Equation 8.2. For a discussion of the delay cancellation measure function see Section 3.5.1. Just as in Chapter 6 $b = 2$ is assumed in this chapter.

8.1.3 ETA matrices

The *SDPM* uses a grid-like structure to store the probabilities associated with all possible arrival events resulting from previous departures, at a given scheduled departure time. These structures are referred to as ETA (Estimated Time of Arrival) matrices. This section defines the structure of ETA matrices and a number of associated conventions that are used later on.

There is one ETA matrix for crew (C) and one for aircraft (A). An element, $A_{i,j,t}$ of the aircraft ETA matrix stores the probability that aircraft j is assigned to aircraft routing i, has arrived, and is available for a subsequent flight at time interval t. $C_{i,j,t}$ is defined similarly. Time is discretised into W minute intervals.

When the ETA matrices are used to calculate the departure time distribution for a given flight, an additional set of matrices A^-, A^+, C^- and C^+ store the changes to the probabilities of A and C that result from particular resources being used for that departure. The changes are not stored directly in C and A because some calculations still depend on the initial values of C and A from immediately after the previous departure. A^- (C^-) store negative changes to A (C) corresponding to airframes (crew) being used for a given departure. A^+ (C^+) store positive changes to A (C) corresponding to airframes (crew) being swapped and assigned to a different line of flight. The changes stored in A^- (C^-) and A^+ (C^+) are applied to A (C) after the departure distribution has been calculated for the departure under consideration, before moving on to the next departure. This step occurs on line 14 of Algorithm 11.

Similarly $A^{\ominus}$ ($C^{\ominus}$) is used to store changes to A^- (C^-) that cannot be applied until after the calculation of the probability of departure within a given time interval during the calculation of a departure time distribution. This step is demonstrated in Section 8.1.7 on line 36 of Algorithm 16.

In general, any probability matrix $\square$ has a counterpart matrix $\square^-$ which stores changes (decrements) to the corresponding elements of $\square$, the changes are stored because they cannot be applied until all the calculations that depend on the initial values of $\square$ have been carried out. For an example, see σ and τ on line 20 of Algorithm 15.

In the following, a probability such as $C_{i,j,t} + C^-_{i,j,t}$ is referred to as a *discounted probability*. A *discounted probability* is the probability that a given resource is available for a departure considering (and deducting) the probability that the resource has already been used for the departure at an earlier time interval. For an example, see line 5 of Algorithm 15.

In the crew ETA matrix, the resource layers sum to 1 minus the probability of crew unavailability. Additionally, the corresponding line of flight layer sums to 1 minus the probability that the initially assigned crew is unavailable.

8.1.4 Modelling aspects of the $SDPM$

This section discusses how the $SDPM$ models: cancellations; swap recovery actions; and the use of reserve crew to replace absent and delayed crew.

Modelling flight cancellation

In the $SDPM$, the probability that a given flight is cancelled is equal to the probability that the flight does not depart before the cancellation threshold. I.e. one minus the cumulative probability of departure before the cancellation threshold. Flights can be cancelled for two reasons, crew unavailability and delays over the cancellation threshold. To account for cancellations due to crew unavailability (for any given flight), the CAM provides the $SDPM$ with the probabilities that reserve crew are available, at different times, to replace absent crew (see Section 8.1.4). This means that when calculating a departure time distribution, the cumulative probability of departure is capped by the probability that crew are available and not absent. The probability of cancellation due to a delay exceeding the cancellation threshold, is the probability of crew availability minus the cumulative probability of departure before the cancellation threshold.

Modelling resource swaps

Swap recovery actions can be used to replace delayed airline resources with those available at the hub station, which are themselves assigned to flights with later departure times. In this work, it is assumed that swap recovery actions are considered for departures which are delayed beyond the delay threshold. The swap recovery actions considered are all combinations of single crew and single aircraft swaps involving the delayed resources, the same which is used in the simulation of Chapter 4, explained in Section 4.4.1. In the $SDPM$, swaps are simulated by assigning delayed resources to other lines of flight and the non-delayed resources to the delayed line of flight, with some probability. The effect is that the delayed departure has some probability of departing earlier than it would have done, without the consideration of swap recovery actions. For each departure, the lines of flight which may provide feasible swap recovery actions are considered in a preference order (see Assumption 4). The calculation of the probabilities that swap recovery actions are available for a given departure at a given time interval are the subject of Section 8.1.5.

Figure 8.3 illustrates the general approach for modelling swap recovery actions in the $SDPM$. In the example there are two crew and two lines of flight, the flight under consideration involves line of flight 1 and the crew are swappable for that line of flight. The example considers the probability that crew are available for a departure at time interval 2. Crew 1 has a 0.1 probability of being delayed beyond time interval 2. Crew 2 has a 0.8 probability of being available at or before time interval 2. Given this, there is a 0.08 probability that crew 1 and crew 2 can swap lines of flight in order to reduce the delay of the flight associated with line of flight 1. In the example, assuming that the aircraft have no chance of being delayed beyond

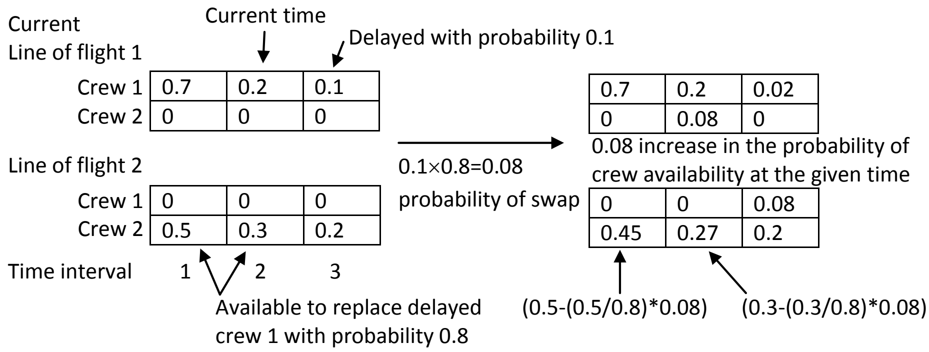

Figure 8.3: Resource swaps in an ETA matrix context

time interval 2, the swap is implemented with a probability of 0.08. The result is that crew 2 end up with a probability of 0.08 of being used for the delayed flight at time interval 2, crew 1 ends up with a probability of 0.08 of being assigned to line of flight 2 and a 0.02 probability of remaining available at time interval 3 for line of flight 1. The probabilities that crew 2 remains assigned to line of flight 2 available at time intervals 1 and 2 are reduced proportionately by a total of 0.08. Figure 8.3 shows the relevant elements of the crew ETA matrix before (left) and after (right) the (probabilistic) swap recovery action.

The assumptions of Section 8.1.1 define what, in this work, constitutes a feasible and beneficial swap recovery action. In the $SDPM$, Assumption 5 sets bounds on the time intervals for which resources assigned to different lines of flight are considered as beneficial swap recovery actions (see w on lines 13 to 17 of Algorithm 15). I.e. the delayed resources must not cause a delay of any other departure, whilst the replacement resources must reduce the delay of the affected departure. Assumption 5 accounts for all beneficial swaps for the case where only one of the (alternative) resources (crew or aircraft) on a swappable line of flight are delayed for their next flight, see Appendix C.

It must be noted that Assumption 5 does not account for all possible beneficial resource swaps when both of the resources assigned to a swappable line of flight are delayed, but by different amounts. In this case it is possible that the less-delayed resource can be swapped and still reduce the overall delay. This is possible if the delayed resource of the current flight is available before the most delayed resource of the other flight. Accounting for this exception is not considered in this work, because such exceptions: are rare; are inherently limited in terms of the amount of delay that they can absorb; and would require an extended ETA matrix structure to retain the information about which resources shared a previous flight.

Modelling reserve crew use

The $SPDM$ is primarily a model of delay propagation and requires as input the probabilities that crew are available for flights at different times (which are stored in C). The $SDPM$ relies upon the CAM to provide probabilities that reserve crew are available to cover for crew absence at different times and the probabilities that teams of reserve crew are available to replace delayed crew. This step corresponds to line 5 of Algorithm 11. This section describes the details of how the probabilities that reserve crew are available to replace absent and delayed crew, as calculated by the CAM, are taken into account in the $SDPM$, and how the $SDPM$ notifies the CAM of the probabilities that reserve crew teams are actually used to replace delayed crew, after calculating a departure time distribution for a given departure.

Crew ETA matrix contributions corresponding to reserve crew being used to cover for absent crew are provided by the CAM. When the CAM is used in conjunction with the $SDPM$ this step replaces line 24 of Algorithm 4 of Section 6.1.5, because the $SPDM$ replaces the rudimentary approach to modelling reserve-induced delay used in Chapter 6. Line 24 of Algorithm 4 is replaced with the following:

$$t = floor\left(\frac{D_d + delay - D_1}{W} + 0.5\right) \tag{8.3}$$

$$C_{F_d,F_d,t} = C_{F_d,F_d,t} + g \tag{8.4}$$

Equations 8.3 and 8.4 show how the CAM provides the $SDPM$ with crew ETA matrix contributions corresponding to reserve crew being used to replace absent crew. Equation 8.3 calculates the time interval at which the particular reserve crew will be available to replace absent crew, accounting for reserve duty start times (Equation 6.5 without the EDT_d term). The addition of half an interval is because the actual times associated with time intervals correspond to the centres of time intervals and time 0 is the centre of the first time interval. Equation 8.4 updates the crew ETA matrix for the probability of availability of crew team F_d assigned to their own line of flight at time interval t due to the reserve combination ($ResCom$) being used to replace absent crew. g is the probability that reserve crew are used to cover crew absence, see line 23 of Algorithm 4.

The $SDPM$ models the use of reserve teams to replace delayed crew as a crew swap between the delayed crew, who are assigned to the given line of flight, and a reserve team, who are assigned to an empty line of flight. After the swap, the replaced delayed crew are assigned to the empty line of flight with some probability. If reserve crew are used to replace delayed crew with some probability, after calculating the departure time distribution for a given flight, the $SDPM$ passes this information to the CAM (see Algorithm 14).

If on line 24 of Algorithm 4, $ResCom$ is a full team of reserve crew, which could also be used to replace delayed crew, crew ETA matrix contributions are added corresponding to the probabilities that reserve crew teams are assigned to the empty line of flight and are available to replace delayed crew assigned to line of flight F_d. In this case the following procedure applies:

Algorithm 12 Reserve crew team ETA matrix contributions

1: $t = floor\left(\frac{D_d + delay - D_1}{W} + 0.5\right)$ (time interval)
2: **for** each $k \in ResCom$ **do**
3: $\quad RI_{k,t} = RI_{k,t} + nodeProb$
4: **end for**
5: $RA_t = RA_t + nodeProb$
6: $C_{TC+1,TC+F_d,t} = C_{TC+1,TC+F_d,t} + nodeProb$
7: $RE \leftarrow C_{TC+1,TC+F_d,t}$

Algorithm 12 shows how the crew ETA matrix is updated with respect to the probability that a team of reserve crew is available to replace delayed crew at time interval t. TC is the total number of crew and crew pairings, so $C_{TC+1,TC+F_d,t}$ corresponds to a reserve crew team constructed to replace crew assigned to line of flight F_d, assigned as crew team number $TC + F_d$, who are assigned to the empty line of flight (i.e. line of flight $TC + 1$). This means that the second dimension of the crew ETA matrix has a length which is twice the total number of crew teams, because for each crew team, a reserve crew team can be constructed to replace them. $nodeProb$ is the probability that the reserve combination is available (see Algorithm 4). Line 1 of Algorithm 12 calculates the time interval t at which the reserve team will be available to replace delayed crew. The purpose of lines 2 to 5 is to store the probabilities that individual reserve crew are available as part of reserve teams at different times between the scheduled departure time and the cancellation threshold, which are required later (see Algorithm 14) to update the CAM with respect to the probabilities that reserve crew remain available for subsequent disruptions.

For each individual reserve in the reserve crew team (line 2), the probability ($RI_{k,t}$) that reserve crew k is available as part of a reserve crew team at time interval t is updated on line 3. Line 5 updates the total probability (RA_t) that reserve crew teams are available at time interval t. Line 6 updates the crew ETA matrix for the probability that the given reserve team is available to replace delayed crew. Line 7 stores the updated elements of the crew ETA matrix in a list RE (required for Algorithm 14).

In the $SDPM$ using teams of reserve crew to replace delayed connecting crew is treated as being less preferable than any crew swap (see Assumption 7), this is achieved by including the empty line of flight last in the preference order of swappable lines of flight for each flight. However, alternative policies can be modelled by altering the preference order.

Once the departure distribution is calculated for departure d, C^- contains the probabilities that crew are actually used for this departure. The following two procedures (Algorithms 13 and 14) are used to update the probabilities that individual reserve crew remain available for subsequent departures (Algorithm 14), taking into account their possible use for replacing delayed crew (Algorithm 13). These procedures constitute line 13 of Algorithm 11.

Algorithm 13 The probabilities that reserve teams are used RU

1: **for** each $m \in RE$ **do**
2: $\quad i = i(RE_m),\ j = j(RE_m),\ t = t(RE_m)$
3: $\quad RU_t = -C^{-}_{i,j,t}$
4: $\quad C_{i,j,t} = 0$ and $C^{-}_{i,j,t} = 0$
5: **end for**
6: Clear RE

Algorithm 13 collects from the crew ETA matrix (elements stored in RE) the probabilities (RU_t) that reserve crew teams are actually used. Line 2 defines functions which return the indices corresponding to a given element of an ETA matrix. Line 3 stores the probability that reserve crew are used at each time interval. Line 4 resets the crew ETA matrix element probability to zero, because reserve crew teams only remain in the crew ETA matrix if they are actually used, i.e. if they are assigned to a flight other than the empty line of flight. Line 6 clears RE ready for use in subsequent departures.

The CAM then uses RU to update the probabilities that individual reserve crew remain available for subsequent use. The following procedure is applied for this. Algorithm 14 shows how the probabilities ($u_{d,k}$) that

Algorithm 14 The procedure for updating the probabilities that individual reserve crew are used for departure d

for each $t \in \{St_d\ to\ Ct_d\}$ **do**
$\quad$**for** each $kin\{1\ to\ R\}$ **do**
$\quad\quad u_{d,k} = u_{d,k} + \left(\frac{RI_{k,t}}{RA_t}\right) \times RU_t$
$\quad$**end for**
end for
Clear RE

individual reserve crew are used for departure d are updated to account for the probabilities (RU) that they are used as part of reserve teams to replace delayed crew. The probabilities that individual reserve crew are used depends on the probabilities they were available at each time interval relative to the total probabilities that reserve crew teams were available at the same time multiplied by the probability that reserve crew teams were used at that time interval. Once algorithm 14 is applied, $u_{d,k}$ on line 15 of Algorithm 3 of Section 6.1.4 allows for the probabilities that reserve crew are used to replace delayed crew.

8.1.5 Calculating the probabilities of resource availability and resource swap availability

As illustrated in Algorithm 11, for each departure, the $SDPM$ considers each time interval between the departure time and the cancellation threshold. For each time interval, the $SDPM$ computes the cumulative probability of departure on or before the given time interval. This calculation depends on: the probabilities that airline resources are available during or before that time interval (see Section 8.1.4); the probabilities that resource swaps are available (see Section 8.1.4); and the probabilities that reserve crew teams are available to replace delayed crew (see Section 8.1.4).

Algorithm 15 Calculation of crew swap availability probabilities for a departure d at time interval t

---

Inputs: Crew and aircraft ETA matrices, sets defining swappable lines of flight and swappable resources, departure number, time interval
Outputs: Probabilities that crew and aircraft are available for the departure at the given time interval, probabilities of the availability of feasible and beneficial pairwise resource swaps
// Calculate the probability of (incumbent) crew availability for departure d at time interval t
$\epsilon = \sum_{\rho \in G_d} \sum_{w=1}^{t} C_{F_d,\rho,w}$
$\epsilon' = \sum_{\rho \in G_d} \sum_{w=1}^{t} \left(C_{F_d,\rho,w} + C^{-}_{F_d,\rho,w} \right)$
// The probability of preferred crew availability before considering crew swaps and the same excluding the probabilities the incumbent crew have already been used for the departure
$pca = \epsilon$
$pca' = \epsilon'$
$\beta = 0$, $\beta' = 0$
Reset σ and τ
for $\mu \in J_d$ (swappable crew pairings for flight d) **do**
  $B = G_d \cap G_{L_{d,\mu}}$ (set of swappable crew)
  // Calculate probabilities that replacement crew on the μ^{th} swappable crew pairing are available on time
  $\sigma_{\nu,\mu} = \sum_{\nu \in B} \sum_{w \in 1 \ldots t} C_{\mu,\nu,w}$
  $\sigma^{-}_{\nu,\mu} = \sum_{\nu \in B} \sum_{w \in 1 \ldots t} C^{-}_{\mu,\nu,w}$
  // Calculate probabilities that delayed crew are feasible to adopt crew pairing μ
  $\tau_{\rho,\mu} = \sum_{\rho \in B} \sum_{w \in \{\max(t+1,\ St_d+DT) \ldots St_{L_{d,\mu}}\}} C_{F_d,\rho,w}$
  $\tau^{-}_{\rho,\mu} = \sum_{\rho \in B} \sum_{w \in \{\max(t+1,\ St_d+DT) \ldots St_{L_{d,\mu}}\}} C^{-}_{F_d,\rho,w}$
  // Calculate the incremental probabilities of pairwise swaps involving crew pairing μ
  $\phi_{\rho,\nu,\mu} = \left(\sigma_{\nu,\mu} + \sigma^{-}_{\nu,\mu}\right)\left(\tau_{\rho,\mu} + \tau^{-}_{\rho,\mu}\right)\left(1 - pca'\right),\ \forall \nu \in B,\ \rho \in B$
  // Calculate cumulative and incremental probabilities of a crew swap involving crew pairing μ
  $csu = 0$, $csu' = 0$
  $csu = \sum_{\nu \in B} \sum_{\rho \in B} \sigma_{\nu,\mu} \tau_{\rho,\mu} (1 - pca)$
  $csu' = \sum_{\nu \in B} \sum_{\rho \in B} \phi_{\rho,\nu,\mu}$
  $\beta = \beta + csu$, $\beta' = \beta' + csu'$
  $pca = pca + csu$, $pca' = pca' + csu'$
end for

Algorithm 15 shows how the probability of crew availability is calculated for a given departure d at time interval t, including crew obtained through crew swaps or by reserve crew used to absorb delay. The same algorithm applies to aircraft availability probabilities when all of the crew variables are replaced with the equivalent aircraft variables.

Firstly, Algorithm 15 calculates (on line 4) the probability (ϵ) that crew who are already assigned to the given line of flight (F_d) are available during or before time interval t. Line 5 calculates the same probability, excluding the probability the same crew contributed to the probability the departure took place before time interval t.

Next, for each swappable line of flight (J_d), in preference order, Algorithm 15 calculates the probability (β) of the availability of crew swaps, which allow the departure to take place during or before time interval t. The probabilities (σ) that feasible crew assigned to swappable lines of flight are available at time interval t are calculated on line 14. On line 15, the same probabilities are calculated excluding the probabilities those crew were used in previous time intervals. The probabilities (τ) that crew assigned to the given line of flight are not available until after time interval t, are calculated on line 17. On line 18, the same probabilities are calculated excluding the probabilities that the same crew were replaced in an earlier time interval. The probabilities (ϕ) that each pairwise swap is a possible recovery action in time interval t are then calculated on line 20. The calculation takes into account: the probability ($1-pca'$) that crew swaps involving preferred lines of flight are not available at the same time; the *discounted probabilities* (first defined near the end of Section 8.1.3) (σ+σ^-) that replacement crew are available at time t; and the *discounted probabilities* (τ+τ^-) that the crew assigned to the given line of flight are delayed. This algorithm includes the use of reserve crew teams, because the empty line of flight, to which reserve crew are initially assigned, is included last in the list of swappable lines of flight (J_d). Lines 22 to 26 keep track of the total probability of crew swap availability and the probability that crew swaps involving preferred lines of flight are available at the same time.

Algorithm 15 provides the probabilities of resource and swap availabilities which are used in Equation 8.5 to calculate the cumulative probability of departure during or before time interval t.

8.1.6 Calculating the cumulative probability of departure during a given time interval

In Algorithm 11, line 8 stated that the cumulative probability of departure during or before the t^{th} time interval is calculated. The calculation of the probability of cumulative departure depends on whether or not the hub departure (d) involves a crew connection, i.e. whether the crew's last flight was on a different aircraft. If the crew were on the same aircraft for the previous flight, the probabilities that the crew and aircraft are available for the next departure are dependent events, because they depend on a common preceding event. However, if the crew connect from a different aircraft, the probabilities that the crew and aircraft are simultaneously available for a departure can be treated as independent events. Equation 8.5 shows how the cumulative probabilities of departure on line 8 in Algorithm 11 are calculated for the cases of crew connections and when crew stay on the same aircraft as they were on for the previous flight (also called a "through" flight, the name used in the airline industry to describe passenger/crew connections that require no aircraft change).

$$cumuP = \begin{cases} \begin{pmatrix} \min(\theta, \epsilon) \\ +\beta \max(0, \theta - \epsilon) \\ +\lambda \max(0, \epsilon - \theta) \\ +(\lambda\beta) \end{pmatrix} & \text{, if } T_d = \text{true} \\ ((\theta + \lambda) \times (\epsilon + \beta)) & \text{, if } T_d = \text{false} \end{cases} \tag{8.5}$$

The first case of Equation 8.5 gives the cumulative probability of departure when the crew are scheduled to stay on the same aircraft. The first term accounts for the probabilities of crew and aircraft availability being dependent on the same event. The second and third terms account for the probabilities that the crew and aircraft, who shared the same preceding flight, are split up by resource swaps for earlier departures occurring after their previous arrival. In particular, the second term accounts for the probability that the assigned crew were swapped during the recovery for a previous delayed departure (implied by the $max\,(0, \theta - \epsilon)$ term) and then the swapped delayed crew being swapped with a different crew, during the delay recovery for the current departure (with probability β). The third term is the aircraft equivalent of the second term. The last term accounts for the probability of departure where both the crew and aircraft are from resource swaps.

The second case of Equation 8.5 applies if the given departure d involves a crew connection. The cumulative probability of departure during or before the given time interval is the product of the probabilities of aircraft and crew availability including those obtained through swap recovery.

8.1.7 Updating departure distributions

The previous section considered the calculation of the cumulative probability of departure during or before time interval t, as a function of resource availability and swap recovery availability. The increment in the cumulative probability from time interval $t-1$ to time interval t is the probability ($incP$) that the departure takes place during time interval t. This section shows: how the crew departure time distribution is updated for time interval t, with respect to the probabilities that each feasible crew is used for departure d during time interval t; how the relevant elements of C^- are updated, with respect to the probability that each crew contributes to the probability of departure during time interval t; and how the relevant elements of C^+ are updated, with respect to the probabilities that delayed resources are swapped and assigned to different lines of flight. The equivalent operations for aircraft are obtained by replacing crew variables with the equivalent aircraft variables.

Algorithm 16 Increase in cumulative probability of departure, updating C, and departure distribution γ for departure d at time interval t

1: **Inputs:** C, d, t, $newCumuP$, $cumuP$, $\epsilon^{'}$, $\beta^{'}$, $\theta^{'}$, $\lambda^{'}$, ϕ
2: **Outputs:** $incP$, updated: γ, C, $cumuP$
3: $incP = newCumuP - cumuP$
4: $cumuP = newCumuP$
5: $cau = incP\left(\frac{\epsilon^{'}}{\epsilon^{'}+\beta^{'}}\right)$ (Probability of incumbent crew use)
6: $csu = incP\left(\frac{\beta^{'}}{\epsilon^{'}+\beta^{'}}\right)$ (Probability of crew swap use)
7: // The use of incumbent crew who were already assigned to the given crew pairing
8: **for** $\rho \in G_d$ **do**
9:     **for** $w \in \{1...t\}$ **do**
10:         $p_{\rho,w} = cau\left(\frac{C_{F_d,\rho,w}+C^{-}_{F_d,\rho,w}}{\epsilon^{'}}\right)$
11:         $C^{\ominus}_{F_d,\rho,w} = -p_{\rho,w}$
12:         $\gamma_{(t-Dt_d),\rho} = p_{\rho,w}$
13:     **end for**
14: **end for**
15: // The use of crew swapped from other lines of flight
16: **for** $\mu \in J_d$ **do**
17:     $B = G_d \cap G_{L_{d,\mu}}$
18:     // Update the probabilities that replacement crew are used (swaps)
19:     **for** $\nu \in B$ **do**
20:         **for** $w \in \{1...t\}$ **do**
21:             $p_{\nu,w} = csu\left(\frac{\sum_{\rho\in B}(\phi_{\rho,\nu,\mu})}{\beta^{'}}\right)\left(\frac{C_{\mu,\nu,w}+C^{-}_{\mu,\nu,w}}{\sigma_{\nu,\mu}+\sigma^{-}_{\nu,\mu}}\right)$
22:             $C^{\ominus}_{\mu,\nu,w} = -p_{\nu,w}$
23:             $\gamma_{(k-Dt_d),\nu} = \gamma_{(k-Dt_d),\nu} + p_{\nu,w}$
24:         **end for**
25:     **end for**
26:     // Update the probabilities that delayed crew are swapped
27:     **for** $\rho \in B$ **do**
28:         **for** $w \in \{\max(t+1,\, St_d + DT)\, ...St_{l_{d,\mu}}\}$ **do**
29:             $p_{\rho,w} = csu\left(\frac{\sum_{\nu\in B}(\phi_{\rho,\nu,\mu})}{\beta^{'}}\right)\left(\frac{C_{F_d,\rho,w}+C^{-}_{F_d,\rho,w}}{\tau_{\rho,\mu}+\tau^{-}_{\rho,\mu}}\right)$
30:             $C^{\ominus}_{F_d,\rho,w} = -p_{\rho,w}$
31:             $C^{+}_{\mu,\rho,w} = C^{+}_{\mu,\rho,w} + p_{\rho,w}$
32:         **end for**
33:     **end for**
34: **end for**
35: // Update the probabilities that crew assigned to different lines of flight available at different time intervals have been used for the given departure on or before time interval t
36: $C^{-} = C^{-} + C^{\ominus}$, $C^{\ominus} = 0$

In Algorithm 16, line 3 calculates the increase in the cumulative probability ($incP$) of departure compared to the previous time interval. The algorithm then attributes $incP$ to the elements of C proportionately to the relative probabilities of crew availability and crew swap availability. Line 5 determines the contribution that can be attributed to incumbent crew, and line 6, that to crew swaps. Line 10 calculates the proportions of the contribution which are due to incumbent crew that can be attributed to each feasible crew, available during or before time interval t. Line 11 stores the changes to the total probabilities that incumbent crew are used for departure d during or before time interval t. Line 12 updates the crew departure distribution with the same probabilities. Then, the algorithm attributes the probability that a crew swap is used to enable departure during time interval t to the relevant elements of C. For each swappable line of flight (Line 16), line 21 calculates the contribution due to a crew swap involving crew swapped from other lines of flight. This is proportionate to: firstly, the ratio of the probability that a pairwise crew swap takes place involving the given crew, relative to the total probability a crew swap is used (the first bracket); and secondly, the ratio of the *discounted probability* that the alternative crew is available relative to the total *discounted probability* of alternative crew availability (the second bracket). The probabilities calculated on line 21 are used on line 22 to store the change in the probabilities that each element of C contributes to the cumulative probability of departure d during or before time interval t. Line 23 updates the crew departure distribution to account for crew swaps enabling departure during time interval t. Lines 29 to 31, are analogous to lines 21 to 23, except that they deal with the probabilities that incumbent delayed crew are replaced in crew swaps. Whereas line 23 updated the crew departure time distribution, line 31 updates the probabilities that incumbent crew will be assigned to different lines of flight when subsequent departures are considered. After performing all calculations which depend on the probabilities of crew being used for departure d before time interval t (i.e. C^-), the accumulated changes ($C^\ominus$) for departure at time interval t are added to C^- (line 36).

After considering all time intervals, the final result is a set of fully detailed departure time distributions, for both crew and aircraft. These distributions are used in the procedure outlined in Section 8.1.8 to update the ETA matrices ready for subsequent departures. Either of these distributions can be used to derive a point estimate departure time. In Section 8.2, such predictions are used to assess the prediction accuracy of the $SDPM$ compared to predictions derived from the corresponding simulation model.

8.1.8 Calculating an arrival time distribution for a hub-spoke-hub cycle

Arrival time distributions are required by the $SDPM$ to update the ETA matrices after calculating a departure time distribution. Algorithm 17 demonstrates how to calculate a discrete arrival time distribution for an out and back cycle (hub-spoke-hub) from a discrete departure time distribution and discrete journey time distributions for the out and return flights, whilst allowing for the possibility that delay may be absorbed by slack at the spoke station. Let DD denote the departure time distribution, AD the arrival distribution, $Qout$ the hub to spoke journey time distribution, Qin the spoke to hub journey time distribution, MS the minimum rest for crew between consecutive flights and TT the minimum turn time for aircraft. Note that MS and TT are assumed to be constant here, but can vary according to fleet types and whether or not crew are scheduled to connect from a different aircraft. Superscripts p and t denote the probability and time

Algorithm 17 Procedure for calculating arrival time distributions

1: **Inputs:** DD, $Qout$, Qin, MS, TT
2: **Outputs:** AD
3: **for** $\forall k \in DD$, $\forall l \in Qout$, $\forall m \in Qin$ **do**
4: $\quad w = \max(St_d,\, DD_k^t + Qout_l^t + \max(MS,\, TT)) + Qin_m^t$
5: $\quad AD_w^p = AD_w^p + (DD_k^p \times Qout_l^p \times Qin_m^p)$
6: **end for**

associated with a particular element of a discrete statistical distribution, whilst w denotes the arrival time index of the discrete arrival time distribution corresponding to a particular combination of $\{k, l, m\}$. The total time for a hub-spoke-hub cycle depends upon the initial hub departure time ($DDist_k^t$), how much delay propagates through the spoke station given the minimum ground time required before the return journey and the scheduled (spoke) departure time (St_d) of the return journey.

8.2 Experimental results

The $SDPM$ is now validated in terms of modelling accuracy and delay prediction accuracy. The $SDPM$ is also assessed in reserve crew scheduling (see Section 8.2.1) and reserve policy applications (see Section 8.2.1). The investigation is divided as follows. In Section 8.2.3 the $SDPM$ is validated in terms of the accuracy of its modelling of delay propagation, swap recovery actions and reserve team use. In Section 8.2.4, time interval sizes (W) are investigated in terms of their effect on delay prediction accuracy, reserve crew schedule quality (reserve scheduling application of the $SDPM$) and reserve policy quality (reserve policy application of the $SDPM$). In Section 8.2.5 the $SDPM$ is used in reserve crew scheduling and reserve policy applications in a range of different configurations, each configuration corresponding to a different assumed recovery policy. The aim is to find the best combination of $SDPM$ configurations for the reserve crew scheduling and reserve policy applications.

8.2.1 Description of $SDPM$ applications

Reserve scheduling application of the $SDPM$

The $SDPM$ is used as an evaluator of reserve crew schedules in a simulated annealing algorithm [54] used to schedule reserve crew. The same algorithm was used in Chapter 6, see Section 6.2.2. The cooling scheme (value of T at any given time since time zero) is based on an exponential decay, starting from T_0 =the maximum number of hub departures in a crew pairing ($maxCP$) and reaching a final temperature of 0.000001 after 20 minutes. I.e. $T_t = maxCP \times e^{-bt}$, where $b = -ln(\frac{0.000001}{maxCP})/20$. Note that the cooling scheme in this case is a function of time as opposed to evaluation number, as it was in Section 6.2.

Reserve policy application of the $SDPM$

The $SDPM$ is also used as a reserve policy, within the validation simulation, to evaluate alternative reserve use decisions (see Section 3.3 for the original motivation for investigating reserve policies). The idea here is that it may sometimes be beneficial to hold back reserve crew, if it is likely that they can be used to greater effect later on. Reserve holding can be beneficial if the current disruption: is a small delay and larger disruptions are expected later, for which those reserve crew could be used for recovery; or, crew are absent, but replacing them with reserve crew induces a large amount of reserve-induced delay and other large disruptions can be solved more efficiently using the same reserve crew. The $SDPM$ evaluates the alternative decisions, starting from the same (initial) conditions as those in the validation simulation plus the decision being evaluated. Actual arrival times are used to update the corresponding entries of the ETA matrices, whose probabilities are set to one. The $SDPM$ recommends decisions which minimise the overall expected level of disruption, i.e. that minimise the value of $ObjVal$.

8.2.2 Test instance

The experiments are based on the test instance that was described in Section 6.2.1. The following experiments were implemented on a laptop with a 2.4GHz dual core Intel Core i7-5500U CPU, with 8Gb of RAM. All models, algorithms and the simulation were implemented in Java as single threaded applications.

8.2.3 Modelling accuracy of the $SDPM$

This section demonstrates that the $SDPM$ successfully models: delay propagation (Figure 8.4); the effect of swap recovery actions (Figure 8.5); and the use of reserve crew to absorb delays (Figures 8.6 and 8.7). In the $SDPM$ and the simulation, the modelling/use of swap recovery actions and reserve crew used to absorb delays, can each be switched on or off. In the following these features are switched on and off in both the simulation and the $SDPM$ to demonstrate that the $SDPM$ models the swap recovery and reserve crew used for delays correctly. The results in this section are based on the smallest possible interval size, $W = 1$. Figure 8.4 shows the predicted delays for each hub departure, when swap recovery actions and reserves which could be used to absorb delays are both switched off (i.e. not modelled in the $SDPM$ and the simulation). The aim is to demonstrate that delay propagation has been modelled correctly. Figure 8.4 shows that the $SDPM$ delay predictions agree very closely with those derived from the

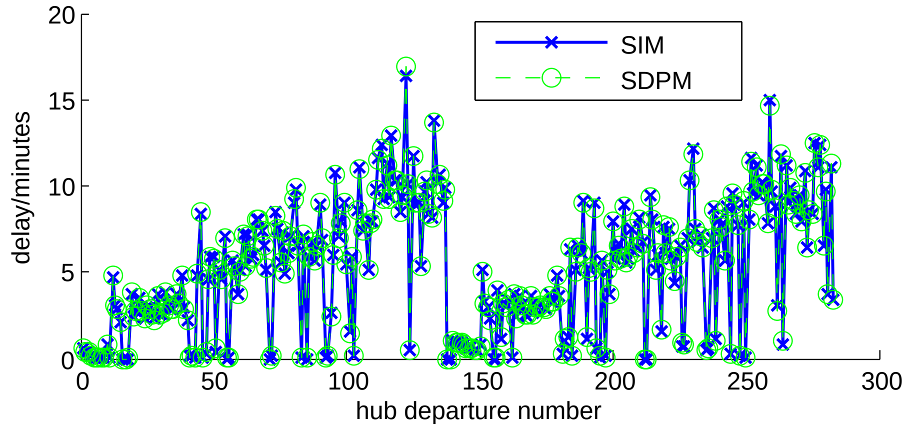

Figure 8.4: $SDPM$ delay predictions compared to those derived from simulation

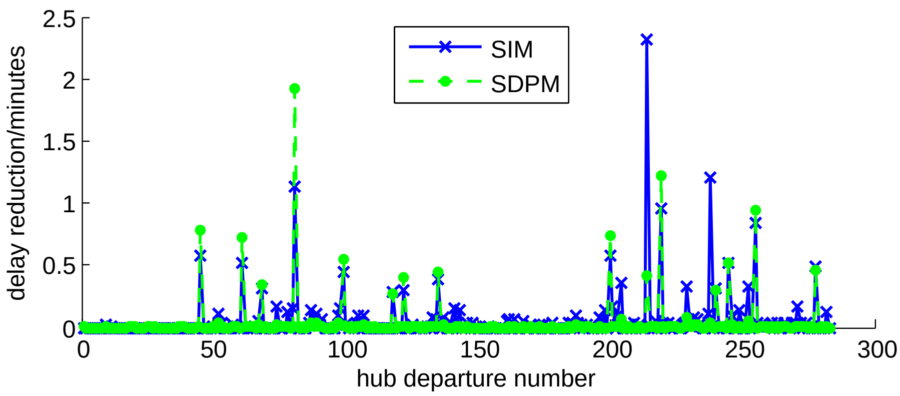

Figure 8.5: The predicted average delay reductions due to allowing swap recovery actions

simulation (Chapter 4). Figure 8.4 also shows how average delay increases over the course of each day of the two-day schedule. Appendix Section D.1 gives an equivalent graph (Figure D.1) to that of Figure 8.4 based on $W = 5$ for a test instance which is described in that Appendix. Next, the qualitative correctness of the $SDPM$'s modelling of swap recovery actions is considered. Figure 8.5 shows the average delay reductions due to allowing swap recovery actions, as predicted by the simulation and the $SDPM$. The results were obtained by switching on the modelling/use of swap recovery actions in the $SDPM$ and the simulation, and subtracting the resulting delay predictions from those displayed in Figure 8.4. Figure 8.5 shows that the $SDPM$ correctly predicts which departures benefit the most (on average) from swap recovery actions, and does so with a reasonable level of accuracy. The $SDPM$ does not however predict the delay reductions for departures which experience a very small amount of delay reduction due to swap recovery actions. The differences between the $SDPM$'s and the simulation's predictions for the effect of swap recovery actions can be attributed to the simplifying assumptions (discussed in Section 8.1.4) made by the $SDPM$

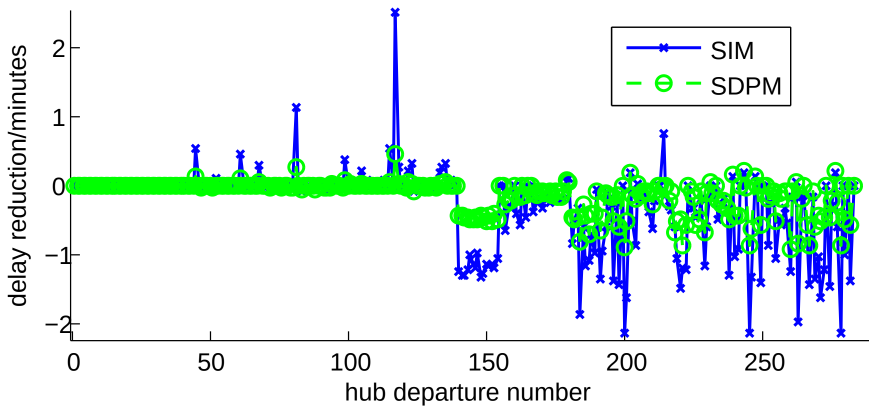

Figure 8.6: The predicted average delay reductions due to using reserve crew to absorb delays

on the conditions under which swaps are beneficial. In general, the effects of swap recovery appear to be relatively small on average, but this is because beneficial swap recovery actions are relatively rare. Appendix Section D.1 gives an equivalent graph (Figure D.2) to that of Figure 8.5 for another test instance which is defined in Appendix D.

Next, the qualitative correctness of the $SDPM$'s modelling of reserve crew used to absorb delay is considered. Figure 8.6 shows the predicted reduction in average delays that result from allowing reserve crew to absorb delays. The results are based on switching on the modelling/use of reserve crew to absorb delays in the $SDPM$ and the simulation, and subtracting the resulting delay predictions from those displayed in Figure 8.4.

Figure 8.6 shows that contrary to expectation, always allowing reserve crew to absorb delay can actually lead to an overall increase of delay (negative reductions). The explanation for this is that, if reserve crew are used to reduce delay, this increases the risk that future flights which are affected by crew absence, will be delayed longer when waiting for later starting reserve crew to become available, because, typically, earlier starting reserve crew will be used to absorb delay. This means that it is best to avoid always using reserve crew to absorb delay. However, it is still plausible that in some isolated circumstances using reserve crew to absorb delay will be beneficial both immediately and in the long run. Determining whether or not one of those exceptional circumstances is in progress is the main purpose of the policy application of the $SDPM$ (see Section 8.2.5). The results of Figure 8.6 show that both the simulation and the $SDPM$ predict this effect. The $SDPM$ tends to underestimate the impact of using reserve crew on average delays by up to a minute, this can be attributed to the simplifying assumptions used to model beneficial swap recovery actions (discussed in Section 8.1.4), which do not account for all possible beneficial swap recovery actions. As a result of this, the rate of utilisation of reserve crew for replacing delayed crew is also marginally underestimated. A knock-on effect of this is demonstrated in Figure 8.7. Figure 8.7 shows that cancellation rates increase when allowing reserve crew to replace delayed crew. This result,

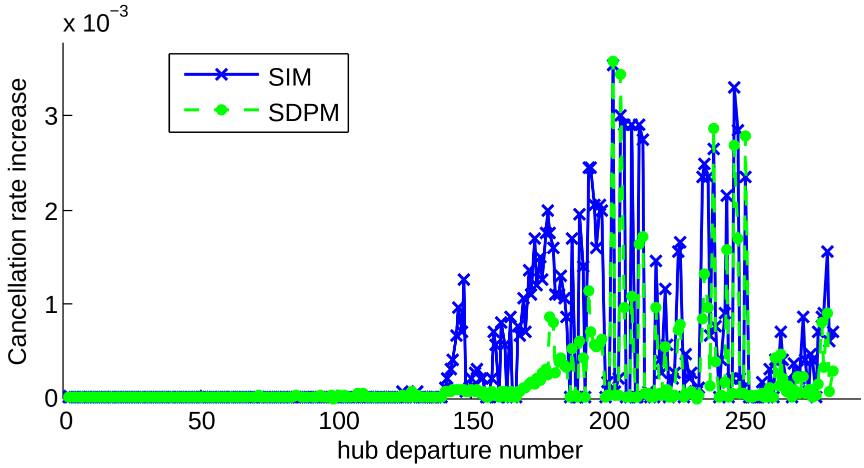

Figure 8.7: The predicted average delay reductions due to using reserve crew to absorb delays

Prediction		Configuration			
type	source	no delay reserve use		delay reserve use	
		no swaps	swaps	no swaps	swaps
delay (minutes)	SIM	5.221	5.161	5.499	5.300
	SDPM	5.212	5.174	5.347	5.242
	RMSE	0.07797	0.1601	0.3970	0.2884
cancellation (rate)	SIM	5.913E-4	5.913E-4	12.40E-4	9.898E-4
	SDPM	13.90E-4	13.90E-4	16.86E-4	15.61E-4
	RMSE	9.276E-4	9.268E-4	8.721E-4	8.307E-4

Table 8.2: Overall delay and cancellation prediction accuracy results for various simulation and $SDPM$ configurations

together with those of Figure 8.6 confirms the idea that using reserve crew to replace delayed crew is best treated as an exception, rather than the rule.

In terms of the accuracy of the predictions derived from the $SDPM$, there is a clear qualitative agreement on which departures suffer an increased cancellation rate. However, the $SDPM$ tends to underestimate the increased cancellation rate by a magnitude up to 2×10^{-3}, which can be attributed to the $SDPM$ underestimating the use of reserve crew for delay absorption, the same explanation as for the marginal underestimation of the impact of using reserve crew to replace delayed crew on average delays. Appendix Section D.1 gives graphs (Figures D.3 and D.4) which are equivalent to Figures 8.6 and 8.7 for a test instance which is defined in that appendix section. They also support the idea that always using reserve crew to replace delayed crew has negative consequences.

Table 8.2 gives average delay and cancellation predictions from the simulation (SIM) and the $SDPM$, as well as root mean squared error (RMSE) values derived from the predictions for individual departures, where the simulation predictions are treated as the target values. Table 8.2 shows that the RMSE's for the $SDPM$ predictions are very small, for both delay

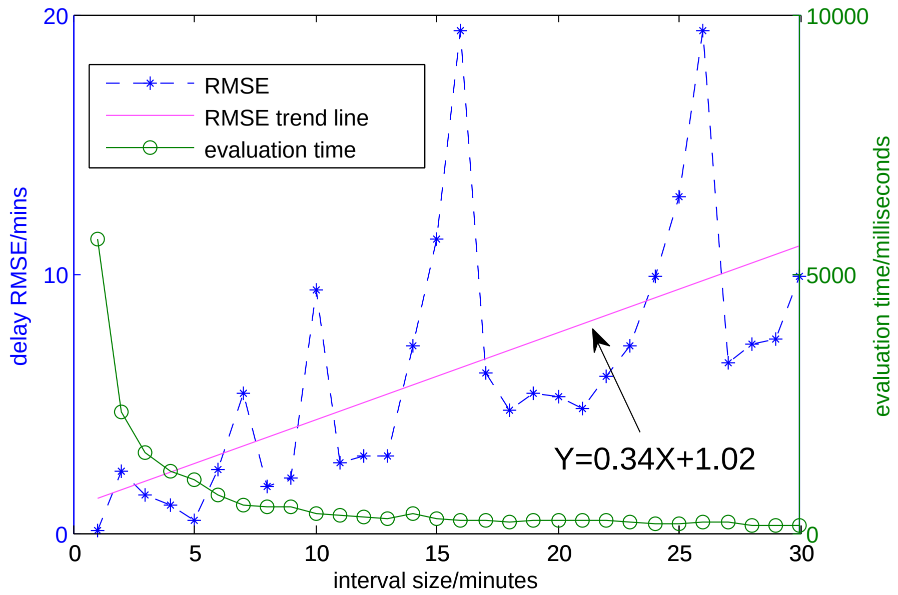

Figure 8.8: The effect of interval size (W) on prediction accuracy and evaluation time

and cancellation predictions. The results are given for four different configurations of the $SDPM$, representing all four combinations of swap recovery actions and reserve crew used to absorb delays, each switched either on or off (see the last four column headings of Table 8.2). The results support the results described above that both the simulation and the $SDPM$ predict that swap recovery actions lead to reduced delays, but allowing reserve crew to absorb delays—whenever this is possible—actually increases overall delay. The $SDPM$'s cancellation predictions are typically slight overestimates, the reason for this has to do with the CAM, see Section 6.2.3 for the explanation of this. However, overestimating cancellations due to crew absence is not as potentially damaging as underestimating them.

In summary, the $SDPM$ gives delay predictions which are very accurate, as evidenced by small RMSE's, and also correctly predicts the qualitative effects of swap recovery actions and the consequences of using reserve crew to absorb delays. Next, the effect of the time interval size W on prediction accuracy, reserve crew schedule quality and reserve policy quality is considered.

8.2.4 The effect of interval size

The prediction accuracy results of the previous section were based on an interval size of 1 minute. Figure 8.8 displays the RMSE for delay predictions derived from the $SDPM$ for a range of interval sizes. The times required to derive the predictions from the $SDPM$ using each interval size are also given. Figure 8.8 shows that prediction accuracy decreases as the interval size increases. This is because, when interval sizes are large, the interval centres become less accurate approximations of the exact times of the events that fall within those time intervals. These errors will accumulate over the course of a schedule. The trend line for the delay RMSE has a gradient which is less than 0.5, a gradient of 0.5 corresponds to the maximum error for a single event time being represented by the nearest interval centre. The fluctuations of the RMSE can be attributed to the variance of the average difference between the actual event times and the nearest interval centres for different interval sizes. I.e. some interval sizes lead to interval centres which are closer to a larger number of the actual event times, which reduces the RMSE, whilst for other interval sizes the opposite is true.

Figure 8.8 shows that evaluations of the $SDPM$ can takes as long as 5 seconds, when using a 1 minute interval size. The simulation takes around 20 seconds for 20000 repeat simulations[1], so the $SDPM$ is always faster, even with an interval size of 1 minute. Evaluation times drop rapidly as interval size increases, because the memory requirements (sizes of the ETA matrices) are inversely proportional to the interval size used, and the number of time intervals which need to be considered for each departure also drop. Evaluation times are important in the scheduling and policy applications of the $SDPM$, because decisions are usually required in a timely manner. A balanced trade-off between prediction accuracy and evaluation time is required. Appendix Section D.1 gives a graph (Figure D.5) equivalent to that of Figure 8.8 for a test instance which is defined in that appendix section.

Next, the effect of interval size on the quality of the reserve crew schedules which can be derived within a 20 minute time limit, using the simulated annealing algorithm, is considered. Figure 8.9 shows the average total cancellation measure derived from 20000 repeat simulations for each reserve crew schedule derived from the $SDPM$, using a variety of interval sizes. Each reserve crew schedule is tested using an absence only reserve policy (**abs only**), which only uses reserve crew to cover for absent crew and never to replace delayed crew. The **abs only** policy approximates an optimal reserve policy, as will be shown in Section 8.2.5.

Figure 8.9 shows that for very small interval sizes, reserve crew schedule quality is, on average, low. This is because the evaluation time is such that the simulated annealing algorithm (Section 8.2.1) cannot perform a sufficient number of iterations within the 20 minute time limit in order to exploit the high modelling accuracy associated with a small interval size. Conversely, for very large interval sizes, reserve crew schedule quality—as indicated by the average cancellation measure—deteriorates. This can be explained by the reduced level of modelling accuracy associated with larger interval sizes, despite being able to perform more simulated annealing iterations within the 20 minute time limit.

Figure 8.9 also shows the effect of interval size on the performance of the reserve policy application of the $SDPM$. The reserve policy application results are based on the single best reserve crew schedule from the reserve scheduling experiments, which occurred for an interval size of 20 minutes. The policy experiment results show that the $SDPM$ policy improves reserve crew schedule performance compared to the *abs only* policy, and that a slow

[1]For a 10-fold cross validation [47], the average cancellation measure RMSE, converges to approximately 0.004, for a simulation sample size of 20000.

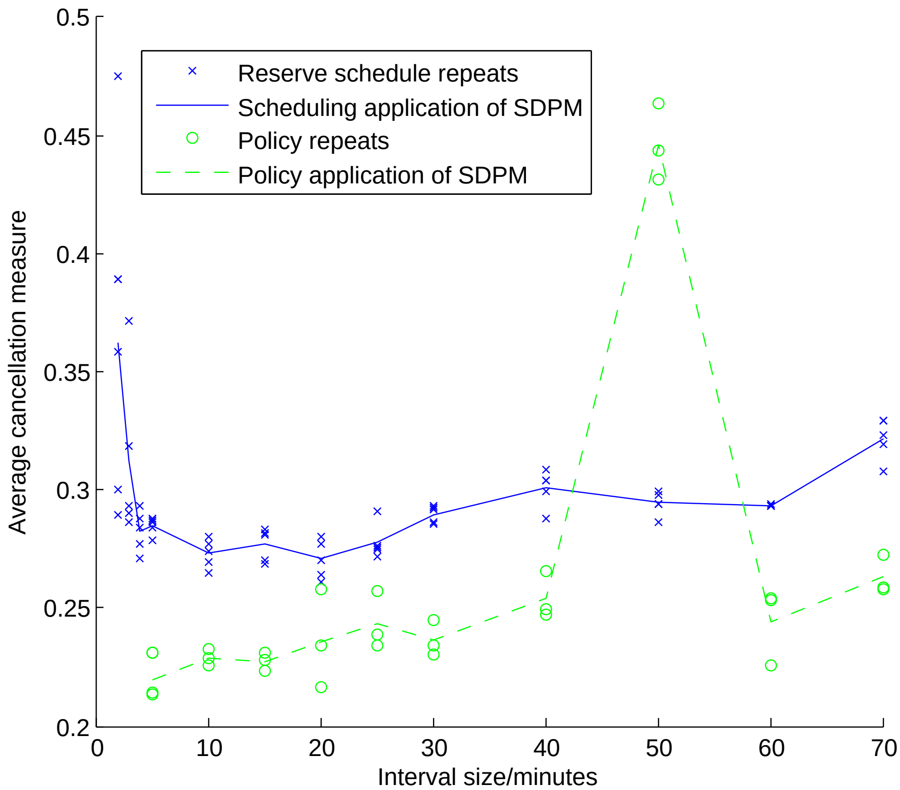

Figure 8.9: The effect of interval size (W) on reserve crew schedule quality and reserve policy quality

deterioration in policy quality occurs as interval size increases.

For an interval size of 50 minutes, a massive increase in average cancellation measure occurs. The explanation for this is that the low modelling accuracy associated with an interval size of 50 minutes leads to systematic errors which lead to recurring poor reserve use decisions. A time interval of 50 minutes also has a large delay RMSE and corresponds to a large fluctuation similar to one of the fluctuations shown in Figure 8.8.

The results of Figures 8.9 and 8.8 indicate that an efficient choice of interval size is 20 minutes. This interval size will be used in the following section.

8.2.5 Scheduling and policy applications of the $SDPM$

This section investigates the effect of different $SDPM$ modelling configurations, which are used in scheduling and policy applications, on the expected level of day of operation disruption. The $SDPM$ has the options of modelling swap recovery actions and reserve crew used for delays. Different configurations can be achieved by switching these features either on or off. Different configurations correspond to assuming different airline recovery policies and reserve policies. This section tests the effects of the following configurations.

SDPM1: Models neither swap recovery or reserve crew used for delays.
SDPM2: Models swap recovery, but not reserve crew used for delays.

SDPM3: Models reserve crew used for delays, but not swap recovery.
SDPM4: Models both swap recovery and reserve crew used for delays.
The **SDPM1** and **SDPM2** configurations correspond to assuming the **abs only** policy (Section 3.5.2), which always uses reserve crew to cover for absent crew, but never to absorb delays.
The **SDPM3** and **SDPM4** configurations correspond to assuming the **default** heuristic reserve policy (see Section 3.5.2), which always uses reserve crew to cover for absent crew, and also to absorb delays whenever this is immediately beneficial.
This section also considers the effect of using the CAM and the SDM (Chapter 6) to schedule reserve crew and as reserve policies.

The following experimental results are based on using each configuration of the $SDPM$, the CAM and the SDM as the evaluator of reserve crew schedules in the simulated annealing algorithm (Section 8.2.1), which is used to derive reserve crew schedules. 10 repeats are performed for each, the best reserve crew schedule, according to the abs only policy, is then used to test the effect of using each configuration of the $SDPM$, the CAM, and the SDM as a reserve policy. The results for the heuristic **abs only** and **default** policies are also given.

The results are given for the average total cancellation measure accumulated in 2000 repeat simulations, for each reserve crew schedule, tested in conjunction with each reserve policy. 2000 repeat simulations represents a compromise between the the amount of time required to perform all of the experiments and ensuring that a large enough number of repeats are performed to derive representative performance measures.

For a 10-fold cross validation, the average cancellation measure RMSE is approximately 0.0137, for a simulation sample size of 2000. So the cancellation measures of Table 8.3 are approximately accurate to within 0.0137. The experiment took nearly 3 days to complete. The random simulation inputs for instantiating uncertain outcomes of journey times and crew absence were the same for each combination of reserve crew schedule and reserve policy tested, which ensures that the comparison is a fair one.

Schedule	Policy								Average
	CAM	SDM	SDPM1	SDPM2	SDPM3	SDPM4	default	abs only	
CAM	1.6593	2.0725	1.0847	1.0771	1.0847	1.0736	2.0057	1.4449	1.4378
SDM	0.2724	0.5350	0.2339	0.2330	0.2339	0.2499	0.7941	0.2750	0.3534
SDPM1	0.2799	0.5119	0.2443	0.2450	0.2443	0.2659	0.7912	0.2711	0.3567
SDPM2	0.2821	0.5404	0.2505	0.2496	0.2505	0.2697	0.7670	0.2682	0.3598
SDPM3	0.3948	0.5480	0.3154	0.3159	0.3154	0.3213	0.6668	0.4117	0.4112
SDPM4	0.3948	0.6398	0.3123	0.3125	0.3123	0.3227	0.6670	0.3965	0.4197
Average	0.5472	0.8079	0.4068	0.4055	0.4068	0.4172	0.9486	0.5112	

Table 8.3: Average cancellation measures for different combinations of configurations of the $SDPM$ used for reserve crew scheduling and as a reserve policy

reason	schedule	policy	CM	Cancellations		average	Reserve utilisation		Max CM
				absence	delay	delay	absence	delay	
worst	CAM	SDM	2.0725	0.0005141	0.00003710	24.233	0.4588	0.09471	21.25
best and increasing	SDM	SDPM2	0.2330	0.0004876	0	7.516	0.4566	0.00313	10.33
scheduling model	SDPM1	SDPM2	0.2450	0.0005159	0	7.515	0.4563	0.00167	10.12
complexity	SDPM2	SDPM2	0.2496	0.0005159	0	7.644	0.4564	0.00167	10.12
Assumed	SDPM2	default	0.7670	0.0016364	0.00001873	9.079	0.4431	0.13567	14.64
and used	SDPM4	default	0.6670	0.0009898	0.00000689	10.307	0.4481	0.09913	15.15
policies	SDPM2	abs only	0.2682	0.0003936	0.00000442	7.815	0.4579	0	13.52
	SDPM4	abs only	0.3965	0.0005913	0.00000053	8.878	0.4534	0	14.98

Table 8.4: Extra performances measures for the interesting results of Table 8.3

Table 8.3 shows how the various approaches to reserve crew scheduling and reserve policies compare to one another on average (last column and last row). The average results for the approaches to reserve crew scheduling indicate that the SDM, $SDPM1$ and $SDPM2$ approaches perform best overall, and attain similar average cancellation measures when compared with one another. These approaches to reserve crew scheduling assume that the **abs only** policy will be used on the day of operation. The SDM, $SDPM1$ and $SDPM2$ approaches to reserve crew scheduling, in this order, represent increasing levels of complexity for modelling delay. SDM is a static delay model, $SDPM1$ is dynamic model of delay propagation, and $SDPM2$ allows for swap recovery actions as well as delay propagation. When these reserve crew schedules were used in conjunction with the policy they assumed during scheduling (**abs only**), this increasing model complexity appears to pay off. However, when these reserve crew schedules were used in conjunction with the $SDPM2$ policy, the relatively simple approach to reserve crew scheduling, the SDM, gave the best overall result. However, this result may be on the limit of statistical significance (+/-0.0137 see above), and also, the $SDPM2$ obtained better reserve crew schedules (as low as 0.21), when it was used in the interval size experiments of Figure 8.9, so the reliability of the simulated annealing algorithm is also a factor.

Assuming that the results of Table 8.3 are representative, an explanation for the good reserve crew schedules derived from the SDM is similar to the explanation for why always using reserve crew to absorb delay led to increased overall delay. Just as using reserve crew for delays is best treated as an exception, the benefits offered by swap recovery actions should not be relied upon when scheduling reserve crew. This is because swap recovery actions are an opportunistic and unreliable method of recovery.

The $SDPM2$ configuration, which allows for swap recovery actions, works best as a reserve policy, because real time information about the availability of swap recovery actions is less uncertain and therefore can be exploited more reliably in an online context.

The average results for policies indicate that the $SDPM$ based reserve policies lead to significantly lower average cancellation measures compared to the other alternatives considered. $SDPM2$ is marginally the best policy overall, however each of the $SDPM$ based policies attains similar results. The results also indicate that the CAM should not be used to schedule reserve crew, and the SDM should not be used as the reserve policy. Table D.2 of Appendix Section D.2 contains the equivalent results for Table 8.3 based on the average cancellation measures from 10 repeats of each configuration used for reserve crew scheduling, as opposed to the best reserve crew schedule from each configuration according to the **abs only** policy. The results support the conclusion that the SDM is a good approach to reserve crew scheduling and that the $SDPM$ is a good approach for a reserve holding policy. Table D.3 of Appendix Section D.2 gives the equivalent results for Table 8.3 when the two day test instance uses the actual scheduled event times as opposed to the tightened test schedule used in this section. These additional results also support those given above.

Next, more performance measures are given for several of the combinations of approaches to reserve crew scheduling and reserve policies. The focus here is on explaining why the various combinations of approaches achieved the average cancellation measures they did.

Firstly, Table 8.4 shows why scheduling reserve crew with the CAM and using the SDM as a reserve policy results in the largest cancellation measure, which is because of very high levels of delay. The CAM does not account for delays, only cancellations due to crew absence, which are accordingly low, because cancellation rate can be decreased at the expense of increased delays. On the other hand, when the SDM is used as a reserve policy, it is unresponsive to real time information about delays, and so performs badly as a reserve policy.

Secondly, Table 8.4 shows extra performance measures for the best overall approach, and the effect of increasing the complexity of the model which is used to schedule reserve crew. The results of Table 8.4 show that, for the best reserve policy ($SDPM2$), allowing for delay propagation ($SDPM1$) and swap recovery ($SDPM2$) during reserve crew scheduling, marginally increases the risk of cancellation due to crew absence. The reason for this is that $SDPM1$ and $SDPM2$ give more weight to delay in the objective function compared to the SDM, which only allows for reserve-induced delays. Additionally, the $SDPM2$ policy eliminates cancellations due to delay and also minimises the worst case scenario (Max CM), compared to the rule of thumb policies (**default** and **abs only**). The combination of the SDM used for scheduling and $SDPM2$ used as the reserve policy, works well because the SDM schedules reserve crew whilst ignoring uncertainties such as delays and the availability of swap recovery actions, whilst the $SDPM2$ considers these uncertainties, but only when real time information becomes available about them.

Thirdly, Table 8.4 shows the effect of assuming either, the **default** policy ($SDPM4$), or the abs only policy ($SDPM2$) during reserve crew scheduling, and that the result depends on whether or not those policies are used on the day of operation. It turns out to be the case that the **default** policy works best if it was also the assumed policy, during reserve crew scheduling. Similarly, the **abs only** policy works best if it was the assumed policy during reserve crew scheduling. In general, the performance of a reserve crew schedule depends on which policy was used to derive it, and in this case whether the reserve policy used on the day of operation matches it.

Scheduling reserve crew whilst assuming the **default** policy leads to low quality reserve crew schedules, where reserve crew are scheduled such that they are unlikely to be available to absorb delays. This has the unfortunate consequence that reserve crew are also less able to cover all crew absence disruptions.

Table 8.4 also shows the reserve utilisation rates for absence and delay disruptions. In general, the policies which minimise the average cancellation measure are very conservative with respect to using reserve crew to absorb delays. This makes sense because absorbing delays using reserve crew is expensive in term of manpower, but the rewards are limited in comparison to the penalty of not being able to replace absent crew with reserve crew.

In this way the abs only reserve policy approximates the optimal reserve policy.

Appendix Section D.2 contains additional results for the various configurations used for reserve crew scheduling, in conjunction with the $SDPM1$ policy, the **default** policy and the **abs only** policy, for the 4 test instances defined in that appendix. These additional results support the conclusion that the SDM is a good approach to reserve crew scheduling and that the $SDPM$ works well as a reserve holding policy.

8.3 Including Aircraft fleet types, crew ranks and qualifications

The CAM was extended to the case of aircraft fleets, crew ranks and qualifications in Section 6.3. As a result, the task of extending the $SDPM$ is mostly complete. In fact the CAM absorbs all of the detail regarding crew ranks and qualifications, because crew are viewed as inseparable teams in the $SDPM$. The only difference that including fleet types has on the $SDPM$ is on which crew and aircraft swaps are feasible. In fact the only additional constraint is that for a pair of crew (aircraft) to be swappable they have to be associated with the same fleet type.

8.3.1 Experiment results for the case of multiple fleet types, crew ranks and qualifications

In Chapter 10 the $SDPM$ is applied to a range of realistic problem instances for the case of multiple fleet types, crew ranks and qualifications. Experiments were performed to validate delay and cancellation prediction accuracy for the 6 test instances (summarised in Table 10.1) that feature in Chapter 10. Appendix E.1 demonstrates the high level of prediction accuracy attained for the 6 test instances.

8.3.2 The effect of interval size on solution quality

In preparation for the experiments of Chapter 10 the effect of interval size on reserve crew schedule quality is investigated. The investigation is repeated for each of the 6 test instances which will be used in Chapter 10, the results are based on a simulated annealing algorithm (see Section 8.2.1) using $SDPM1$ as the evaluator with repeat experiments for each interval size tested. Each repeat experiment is limited to 10 minutes (using the same hardware and software described in Section 6.2.1). The experiments are analogous to those of Figure 8.9 of Section 8.2.4. The aim is to find an interval size for each test instance that represents an optimal trade-off between model accuracy and solution quality, given that a solution is required within a specified amount of time.

The experiment results for test instances 1 to 6 are given in Appendix F. Test instances 1 and 6 represent the two extremes of the test instances, the results for these test instances are discussed in detail in Appendix F.

Test instance	1	2	3	4	5	6
Optimal tradeoff interval size	10	20	80	15	35	55

Table 8.5: Optimal trade-off interval sizes for the Chapter 10 test instances

Table 8.5 gives the optimal trade-off interval sizes for each test instance according to a 50/50 weighted sum of the average and minimum cancellation measures achieved in the repeat experiments for each interval size. In general, smaller schedules permit the simulated algorithm to exploit the increased accuracy provided by using a smaller interval size, whereas for larger schedules evaluation times inhibit this effect because of the reduced number of iterations that can be performed within a fixed time limit.

8.4 Chapter summary

In this chapter, a statistical model of delay propagation in an airline network has been introduced based on the concept of delay distributions propagating through an airline's schedule. The *SDPM* was based on assumptions in line with those made for the simulation of Chapter 4, which made it effectively a theoretical model of that simulation. The *SDPM* superseded the probabilistic crew delay model of Chapter 7 as it allows for delays in general and not just crew related delays. The *SDPM* incorporates crew absence uncertainty, which means it integrates delay and crew absence uncertainty within a single model. The *SDPM* was validated in terms of delay and cancellation prediction accuracy compared to the predictions derived from the simulation. It was demonstrated that the *SDPM* models delay propagation, swap recover actions and the use of reserve crew to replace delayed crew correctly. The prediction accuracy tests were repeated for a number of different configurations of the model, which were also implemented in the validation simulation. The features that modelled swap recovery actions and the use of reserve crew to replace delayed connecting crew were each switched on or off. It was found that contrary to expectation, always allowing reserve crew to replace delayed crew can actually lead to an overall increase in delay. This was caused by an increase in the average time that crew absence affected flights had to wait for replacement reserve crew.

An investigation of the time interval size used in the *SDPM* revealed that although small interval sizes lead to more accurate predictions, larger interval sizes can be as effective in reserve crew scheduling and reserve policy applications whilst also providing results faster. The *SDPM* was then implemented in a number of different configurations, applied to schedule reserve crew and as a reserve policy. Each configuration corresponded to a different assumed airline recovery policy. It was found best to schedule reserve crew using the *SDM* evaluator and the *SDPM* as the reserve holding policy. The reason for this was that the *SDPM* is better able to exploit real time information about arrival times from previous flights and therefore the availability of unreliable and opportunistic recovery actions such as swaps. The *SDM* works best offline because it does not attempt to exploit these.

The *SDPM* easily extends to the case where there are multiple fleet types, crew ranks and qualifications, because the CAM absorbs the vast majority of the additional modelling details. In Chapter 10, the $SDPM$ is applied in a range of test instances where there are multiple fleet types, crew ranks and qualifications, experiments were performed to validate the modified $SDPM$ and find the optimal trade-off interval sizes for those test instances.

Chapter 9

Mixed integer programming simulation scenario model

This chapter explores a scenario-based approach to modelling crew-related disruptions and reserve crew recovery, which represents an alternative to the probabilistic approaches described in Chapters 5 to 8. In contrast to the probabilistic models, the scenario-based approach accounts for reserve crew used for both crew absence and crew-related delay disruptions from the outset, as opposed to developing and then integrating two independent models.

Much of the content of this chapter was published in [16]. This chapter describes a scenario-based approach to the reserve crew scheduling problem called the **Mixed Integer Programming Simulation Scenario Model** ($MIPSSM$) which uses information from repeat simulations of an airline network. The simulation data is used to generate disruption scenarios which are used to form the constraints and coefficients of the $MIPSSM$ formulation. The $MIPSSM$ formulation is then solved to find the reserve crew schedule that would have minimised the level of delay and cancellation that would have occurred in the original simulations which were used to derive the disruption scenarios.

Influences for this work

The $MIPSSM$ described in this chapter was influenced by approaches such as robust optimisation (see Section 2.7.2), stochastic programming (see Section 2.7.1) and recoverable robustness (see Section 2.7.3).

The $MIPSSM$ is most similar to a robust optimisation approach, in which solutions are found which are optimal over an uncertainty set, or a set of realisations of the uncertain parameters for a given problem. In relation to the work in this chapter, an element of an uncertainty set corresponds to a disruption scenario. In robust optimisation, robust solutions have the quality that they are stable under uncertainty, i.e. work well for a wide range of possible outcomes of the uncertain parameters, including worst case outcomes, as opposed to performing very well in some cases and very poorly in others.

The $MIPSSM$ described in this chapter is similar to a recoverable robustness approach. Recoverable robustness is similar to robust optimisation except that recoverable robustness includes a model of recovery. The model of recovery is analogous to the recourse problem in multi-stage stochastic programs.

The structure of the reserve crew scheduling problem does not allow a direct application of recoverable robustness or stochastic programming. Firstly, the recovery phase of recoverable robustness is a fixed algorithm, whereas in the reserve crew scheduling problem, the recovery actions are the variables to be optimised. Secondly, the assumptions required by stochastic programming also prevent a direct application of this approach, namely because of the assumption that the recovery decisions in one stage do not influence the model of uncertainty of future stages. In fact, reserve crew used now, do influence future disruptions which are related directly or indirectly to the current disruption.

In comparison to previous chapters

The $MIPSSM$ was developed over the same time period as the CAM of Chapter 6, and as a result, the first formulation given applies to the case of a single aircraft fleet type, and a single crew rank and qualification, with crew teams consisting of a number of individual crew. The single fleet, crew rank and qualification formulation simplifies the initial development and analysis of the $MIPSSM$. The initial formulation is then modified for the case of multiple fleet types, crew ranks and qualifications (Section 9.8). The development of the $MIPSSM$ also motivated the development of the simulation tool described in Chapter 4, as simulation is required to generate the input disruption scenarios for the $MIPSSM$ formulation.

A fundamental difference between the $MIPSSM$ and the probabilistic approaches is that in the $MIPSSM$ reserve crew are scheduled in such a way that they can be used optimally over a given set of scenarios. This means that the $MIPSSM$ approach implicitly assumes an optimal reserve policy, whereas, the probabilistic models assume a preference order based policy with no reserve holding. The optimal policy assumed by the $MIPSSM$ relies on full knowledge of future outcomes, which is not available on the day of operation. As a compromise, a simulation based learning approach is proposed for learning the optimal reserve policy corresponding to a given reserve crew schedule. The learned policy takes the form of a look-up table.

Although the $MIPSSM$ is an alternative approach to the probabilistic approaches of Chapters 5 to 8, the $MIPPSM$ is similar to the probabilistic crew delay model of Chapter 7. This similarity lies in the use of simulation to provide data on when to schedule reserve crew in anticipation of crew-related delays whilst assuming that swap recovery actions are preferred to reserve crew recovery actions. In both approaches, the data regarding the possible benefit of using reserve crew to cover delayed crew is calculated after swap recovery actions have been applied. This provides the mechanism through which both approaches (indirectly/implicitly) take swap recovery actions into account. Additionally, both approaches model the potential for knock-on delays, that reserve crew, once used to cover crew-related delays, may prevent or reduce.

One of the potential advantages of a scenario-based approach in comparison to the probabilistic approaches is that scenarios are much simpler to model, as each scenario corresponds to a single sequence of events. However, a possible disadvantage of the $MIPSSM$ approach is that it is not a structural model of airline operations—which the $SDPM$ of Chapter 8 is. As a result, the $MIPSSM$ will not be able to capture the indirect effects of reserve crew use, only the direct ones. This is a typical limitation of multi-stage stochastic programming formulations, they often require the assumption that the uncertainty of future outcomes does not depend on the decisions made in previous stages. So the difference between the quality of the solutions from the $MIPSSM$ approach and the $SDPM$ approach should (all else being equal) indicate the value of allowing for indirect (network) effects.

Chapter structure

The remainder of this chapter is structured as follows. Section 9.1 gives an overview of the proposed $MIPSSM$ approach. Section 9.2 introduces the simulation used to generate disruption scenarios and how disruption scenarios are derived from the simulation. Section 9.3 presents the formulation of the $MIPSSM$ and Section 9.4 introduces several alternative objective functions for the $MIPSSM$ formulation and a scenario selection heuristic. Section 9.5 describes how a look up table reserve policy can be derived for a reserve crew schedule using an adapted version of the $MIPSSM$ formulation. Section 9.6 gives experimental results for a comparison of the $MIPSSM$ approach with alternative methods of reserve crew scheduling. Section 9.7 presents an investigation into what makes a good set of input scenarios for the $MIPSSM$ formulation with respect to solution reliability and the quality of the resultant reserve crew schedule. Section 9.8 modifies the initial $MIPSSM$ formulation to accommodate multiple aircraft fleet types and the ranks and qualifications of crew. Section 9.9 describes possible future work. Section 9.10 gives a summary of the main findings of this chapter.

9.1 Overview of the MIPSSM

This section describes the sequence of stages involved in the $MIPSSM$ approach. Additionally, an explanation is given of how the objective values of using reserve crew to absorb crew-related delays are calculated from the delay cancellation measure function, a function which was first introduced in Chapter 3.5. The use of the delay cancellation measure function means that the $MIPSSM$ formulation can be formulated as a single objective problem.

9.1.1 Stages of the MIPSSM approach

Figure 9.1 illustrates the stages that are required to be performed sequentially in the proposed $MIPSSM$ approach, from input data through to

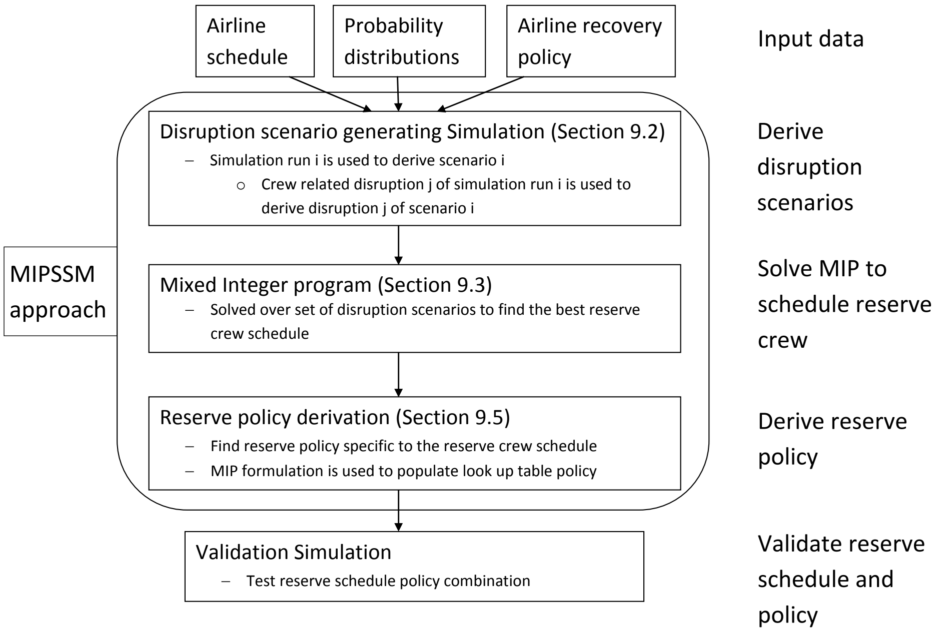

Figure 9.1: Sequential stages of the $MIPSSM$ approach to scheduling airline reserve crew

validation. Note that the input data and validation simulation stages are not part of the $MIPSSM$ approach to reserve crew scheduling, but have been included in Figure 9.1 to illustrate the full cycle of deriving and testing reserve crew schedule and policy combinations. The $MIPSSM$ approach to reserve crew scheduling involves three main stages:

1) A simulation stage is used to derive disruption scenarios. 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 airline's schedule in the considered time horizon from start to finish once. A disrupted flight in the simulation results in a disruption added to the disruption scenario. For each disruption in a disruption scenario there is a corresponding record of all of the reserve crew start times (discretised to match the scheduled departure times) which, if scheduled, would allow the corresponding reserve crew to be used to remove completely, or reduce, the given disruption.

2) A $MIPSSM$ formulation is solved to find the best reserve crew schedule for the set of disruption scenarios generated in the first stage. In the $MIPSSM$ formulation there are 2 types of variables: x the reserve crew schedule and y the reserve use variables. For each disruption scenario there is a corresponding subset of the reserve use variables. The reserve use decisions made for each disruption scenario have to be feasible with respect to the overall reserve schedule x (i.e. reserve crew can only be used if they are scheduled). The difficulty is finding a reserve schedule that allows disruptions in many of scenarios to be covered in an efficient manner. Solving the $MIPSSM$ formulation over a set of input disruption scenarios in an appropriate solver finds both the reserve crew schedule x and the reserve use decisions y that minimise delay and cancellations over all of the

input disruption scenarios.

3) Lastly a reserve policy is derived corresponding to the reserve crew schedule found in the $MIPSSM$ formulation stage, which defines the conditions on the day of operation under which reserve crew use is permitted. The policy takes the form of a look up table which specifies the minimum number of reserve crew that should be available at each departure time if reserve crew are to be permitted to be used to absorb crew-related delay affecting a given departure. Reserve crew are always used to cover for absent crew and never held.

9.1.2 Cancellation measure of a delay

cm_h	:	Delay cancellation measure of flight h
td_h	:	Total delay affecting flight h
CT	:	Cancellation threshold
$rd_{h,l}$	:	The delay that occurs when using reserve crew with start time index l to recovery from a delay or absence disruption affecting departure h
D_h	:	Departure time of flight h
$aeta_h$	:	Estimated time of arrival of the aircraft whose next flight is flight h
$ceta_h$	:	Estimated time of arrival of the crew whose next flight is flight h
TT	:	Minimum turn time of aircraft between consecutive flights
MS	:	Minimum sit or rest for crew between consecutive flights
cd_h	:	The delay that can be attributed to the crew assigned to flight h over and above any delay caused by the aircraft which assigned to same flight

Table 9.1: Delay cancellation measure related notation

Table 9.1 lists and defines the notation used for calculating delays and delay cancellation measures. The goal of the $MIPSSM$ approach is to schedule reserve crew to minimise delay and cancellation disruptions. To retain the simplicity of a single objective problem in the $MIPSSM$ formulation, Equation 9.1 converts delay into a measure of cancellation. Equation 9.1 is based on Equation 3.3, which was first introduced and explained in Section 3.5.1. Equation 9.1 is used to determine the cancellation measure of a delay before reserve crew recovery actions have been taken into account. In this case the total delay is calculated using Equation 9.3. If a flight is delayed due to delayed connecting crew, i.e. Equation 9.4 returns a non-zero value, reserve crew can be used to absorb the delay by replacing the delayed crew. In which case, Equation 9.2 replaces Equation 9.3 as the numerator of Equation 9.1 to obtain the cancellation measure of a delay if reserve crew with start time index l (start time=D_l) are used to replace the delayed connecting crew of flight h.

$$cm_h = \left(\frac{td_h}{CT}\right)^n \tag{9.1}$$

$$rd_{h,l} = \max(0,\ \max(D_l,\ aeta_h + TT) - D_h) \tag{9.2}$$

$$td_h = \max(0,\ \max(aeta_h + TT,\ ceta_h + MS) - D_h) \tag{9.3}$$

$$cd_h = \max(0,\ ceta_h + MS - \max(D_h,\ aeta_h + TT)) \tag{9.4}$$

In this thesis the delay exponent of the delay cancellation measure function is set to 2 ($n = 2$), see Section 3.5.1. The disruption scenario generation stage collects information about the possible objective value (i.e. the associated cancellation measures) of using reserve crew scheduled at different times for different disruptions in each disruption scenario. The equations described above are used for this purpose.

9.1.3 Disruption scenarios

The proposed $MIPSSM$ approach uses the concept of disruption scenarios. A disruption scenario corresponds to a set of crew-related disruptions that could occur during the implementation of an airline's schedule. In the $MIPSSM$ approach disruption scenarios are collected from a simulation of an airline. The simulation has stochastic crew absence and journey time inputs instantiated from corresponding statistical distributions. For each disruption in a disruption scenario information must be maintained about the disruption size (in the form of a cancellation measure), the number of reserve crew required to cover the disruption, and the benefits of using reserve crew scheduled at different times to cover the disruption. The information about the benefit of using reserves scheduled at different possible times is stored in the form of sets of *feasible reserve instances* corresponding to each disruption in each disruption scenario (see Section 9.1.4)

9.1.4 Feasible reserve instances

In the simulation which generates disruption scenarios, information regarding the benefit of using reserve crew scheduled at different times to absorb a given disruption is collected. For each reserve start time that is feasible to absorb a given disruption, a *feasible reserve instance* is generated. A feasible reserve instance therefore corresponds to a combination of a reserve crew duty start time and a disruption that could be absorbed by using a reserve crew with such a duty start time. For each feasible reserve instance there is a cancellation measure that replaces the cancellation measure of the disruption when no reserve crew were available, if the reserve is used (in the $MIPSSM$ formulation of Section 9.3). The use of feasible reserve instances means that the $MIPSSM$ formulation only contains binary variables corresponding to feasible instances of reserve crew use and therefore reserve feasibility constraints are not required.

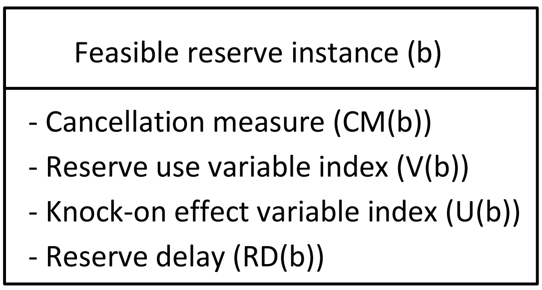

Figure 9.2: Feasible reserve instance attributes

Let b denote a given feasible reserve instance. For each feasible reserve instance (b) there is:

1) A corresponding cancellation measure $(CM(b))$ which is calculated in the disruption scenario generation stage. This is the cancellation measure that applies in the $MIPSSM$ formulation if reserve crew with the duty start time (corresponding to b) are used to cover the disruption (corresponding to b).

2) An associated unique reserve use variable index $(V(b))$ which identifies the binary reserve use variable in the $MIPSSM$ formulation associated with the feasible reserve instance.

3) A unique (knock-on effect) reserve use variable index $(U(b))$ corresponding to feasible reserve instances which can absorb a root delay that subsequently propagates, hence reducing the secondary delay.

4) A reserve delay $(RD(b))$ caused by waiting for the given reserve crew to start their standby duty before they can be used for the disruption associated with the given feasible reserve instance.

Feasible reserve instances generated in the disruption scenario generation phase are each stored in two sets. In one set containing all of the feasible reserve instances which were generated for the same disruption and the same disruption scenario, and in a second set containing all of the feasible reserve instances generated with the same reserve start time and for the same disruption scenario. These sets are then used to form the constraints of the $MIPSSM$ formulation (Section 9.3).

9.2 Disruption scenario generation simulation

This section explains the disruption scenario generation stage of the *MIPSSM* approach. Table 9.2 introduces the schedule notation and Table 9.3 introduces the notation of disruption scenarios. Section 9.2.1 describes how the single hub airline simulation of Chapter 4 is used firstly for disruption scenario generation and then later reused for experimental validation of reserve crew schedules. Section 9.2.2 defines what is meant by a disruption scenario and how the information it stores is collected from the simulation.

C_h	:	Crew team number scheduled to flight h
A_h	:	Aircraft number scheduled to flight h
$crewSize_h$	:	Number of crew in crew team scheduled to flight h
$\|P_n\|$	:	Number of hub departures in crew pairing n
$P_{n,m}$	:	Departure number of the m^{th} hub departure of crew pairing n

Table 9.2: Schedule notation

9.2.1 Simulation

The simulation of a single hub airline is used without reserve crew to generate a set of disruption scenarios which contain information on the possible benefit of using reserve crew scheduled at specified times to mitigate the given disruption. These disruption scenarios form the input for the $MIPSSM$ formulation (Section 9.3.1).

The simulation of Figure 9.3 is the same as that introduced in Chapter 4 except for the added stages of collecting information for disruption scenarios in the event of uncovered crew absence or crew-related delay. The input crew and aircraft schedules were built using 'first in first out' scheduling (more details of the test instance can be found in Section 9.6.1).

The simulation has a dual purpose: disruption scenario generation and reserve crew schedule validation. For disruption scenario generation, no reserve crew are scheduled and none are therefore available for recovery (as the point of disruption scenario generation is to find information about when reserve crew are most needed). In contrast the validation simulation does include a reserve crew schedule and is used to compare the reserve crew schedules which were found using the $MIPSSM$ against reserve crew schedules obtained using alternative approaches.

resInstanceDiagram

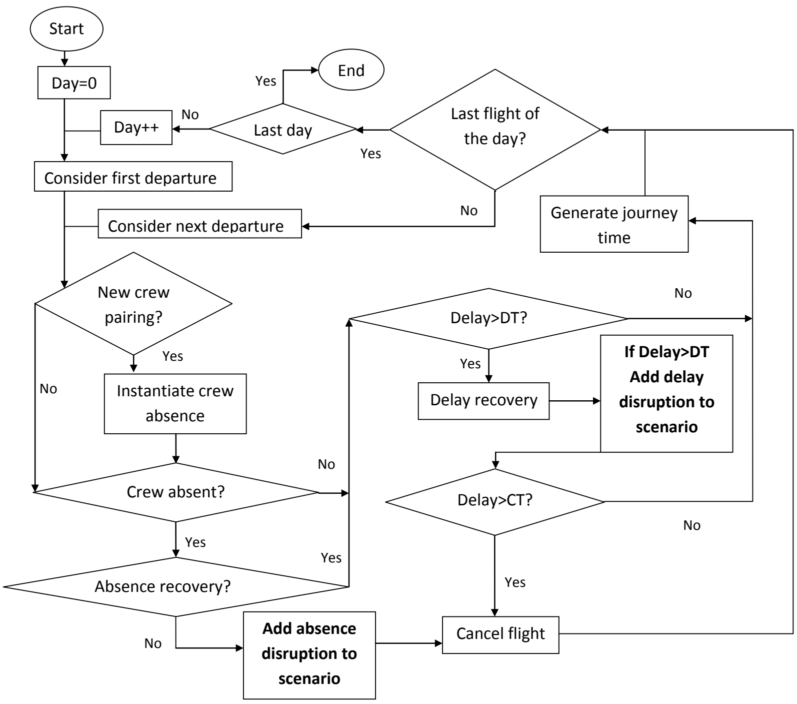

Figure 9.3: Flow chart of the simulation used to derive disruption scenarios

Figure 9.3 illustrates: how the simulation models the execution of an airline's schedule; how crew absence uncertainty and journey time uncertainty are included in the simulation; the process of airline recovery and the points in the simulation at which information is yielded about disruptions that are solvable by using reserve crew. This is then used to derive the disruption scenarios. 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 pairing then the number of absent crew is instantiated from the cumulative statistical distribution. If reserve crew are not available (as is always the case in the disruption scenario generating simulation) then the flight has to be cancelled. At this point in the simulation, information on the possible benefit of scheduling reserve crew at different start times is collected (Section 9.2.2). I.e. a crew absence disruption is added to the current disruption scenario. If reserve crew are available (as may be the case in the validation simulation used in Section 9.6) they are considered for use in earliest start time order and the reserve policy, which is introduced in Section 9.5, determines whether or not reserve crew should be used or held. If a departure is delayed by more than the delay threshold (DT =15 minutes) all combinations of single crew and aircraft swaps are considered in an attempt to recover from the delay (note that the approach would still work for other approaches to modelling swap recovery actions). If the delay is still above the delay threshold, even after the consideration of swap recovery actions, information is collected on the possible benefit of using reserve scheduled at different possible start times (Section 9.2.2). I.e. a delay disruption is added to the current disruption scenario. If the post recovery delay is still above the cancellation threshold (180 minutes), or crew absence cannot be covered, the affected flights are cancelled.

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 reserve crew are used 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} feasible reserve instance associated with disruption j in scenario i
$G_{i,j}$	:	Set of feasible reserve instances corresponding to feasible reserve instances 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} feasible reserve instance corresponding to a feasible reserve instance which could be used to absorb crew delay on a preceding flight that also has the knock-on effect of reducing delay disruption j in scenario i
$R_{i,l}$	:	Set of feasible reserve instances with start time index l in scenario i
$R_{i,l,k}$	:	k^{th} feasible reserve instance in the set of feasible reserve instances corresponding to corresponding to reserve crew with start time index l in scenario i
b	:	Feasible reserve instance (used in pseudocode to refer to a newly generated feasible reserve instance which is then added to the applicable sets of feasible reserve instances)
$V(b)$	:	Index of the reserve use variable corresponding to a given feasible reserve instance b
$U(b)$	:	Index of the reserve use variable corresponding to a feasible reserve instance generated for a knock-on disruption which, if feasible reserve instance b is used to cover the root delay (preceding flight), reduces the delay propagated to that follow-on flight
$CM(b)$	:	Cancellation measure contribution corresponding to a given feasible reserve instance b
$RD(b)$	:	Reserve delay corresponding to a given feasible reserve instance b

Table 9.3: Disruption scenario notation

9.2.2 Simulation derived scenarios

A given disruption scenario i corresponds to a single run of the simulation. This section explains how simulation is used to derive the information for disruption scenarios.

In disruption scenario i, a disruption j corresponds to the j^{th} crew disrupted flight for which reserve crew use could be a beneficial recovery action. Reserve crew use is beneficial when a flight is delayed due to crew (where the delay is greater a specified delay threshold DT), even 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 that can be absorbed by using reserve crew, data is collected regarding all of the possible feasible reserve start times that, if scheduled, could be used to reduce the disruption. For each such beneficial reserve start time, feasible reserve instances are generated. A feasible reserve instance (Section 9.1.4) corresponds to a feasible reserve crew duty start time index which can be used to cover a given crew disrupted flight in a given scenario (i.e. reserve start time/disruption pair). For each disruption the number of feasible reserve instances which are generated for each feasible reserve start time index is equal to the number of reserve crew required to cover the given disruption, which is either the number of crew absent (for crew absence disruption) or the size of the crew team assigned to flight h ($crewSize_h$) for a delay. Let $F_{i,j}$ denote the set of feasible reserve instances corresponding to possible reserve start times that could, if scheduled, be used to solve or reduce disruption j of disruption scenario i.

For the specific case of delay disruptions it is also possible that reserve crew use can have the effect of reducing or preventing knock-on delays. For this purpose the set $G_{i,j}$ is introduced which denotes the set of feasible reserve instances corresponding to the possible use of reserve crew originally used to absorb the root delay also being used to absorb the knock-on disruption. Note that crew-related delays occur when a flight has to wait for crew on a delayed connecting flight, so the reserve used for the root delay can only influence the delay of the following flight if other reserve crew are not used to absorb the delay of that following flight. Algorithms 18 and 19 outline the procedure of collecting information for the disruption scenarios from the single hub airline simulation.

The notation used in Algorithms 18 and 19 is that which was defined in Section 9.1.4, Tables 9.1, 9.2 and 9.3.

Algorithm 18 is used in the disruption scenario generating simulation when a crew absence occurs, the algorithm considers all of the possible ways the absence disruption can be covered using reserve crew (reserve crew with different start time indices l) and generates $N_{i,j}$ feasible reserve instances for each. The number of reserve crew required to cover a disruption equals the number of absent crew (line 6). The cancellation measure of the absence disruption is the number of hub departures in the disrupted crew pairing that would have to be cancelled if reserve crew were unavailable to cover the absent crew (line 7), with no delay contribution to the cancellation measure.

The algorithm then considers each possible reserve start time (line 8) used to cover absent crew at each hub departure in the disrupted crew pairing (line 9). If a reserve is feasible, $N_{i,j}$ new feasible reserve instances are generated with unique reserve use variable indices. For each of these the associated cancellation measures are 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 reserve-induced delay (lines 13-20). The newly generated feasible reserve instances use are stored in sets according to which disruption and scenario they are applicable to (line 17) and to which reserve start time

Algorithm 18 Pseudocode for deriving disruption scenario information for a crew absence disruption occurring at simulation run i departure k

1: Inputs: Crew-related disruption affecting departure k of simulation run i (number of absent crew)
2: Outputs: Disruption j of scenario i ($cm_{i,j}, N_{i,j}, F_{i,j}$...)
3: $RUVI$ =number of reserve use variable indices used so far
4: **if** crew absence disruption **then**
5: $\quad W_i = W_i + 1$
6: $\quad N_{i,j}$ = number of crew absent
7: $\quad cm_{i,j} = |P_{C_k}|$ (all hub departures cancelled if absence is not covered)
8: $\quad$ **for** $l = 1$ to total hub departures **do**
9: $\quad\quad$ **for** $m = 1$ to $|P_{C_k}|$ **do**
10: $\quad\quad\quad$ **if** reserve crew with start time D_l are feasible to cover crew absence at the m^{th} hub departure of the crew pairing assigned to crew team number C_k **then**
11: $\quad\quad\quad\quad f = P_{C_k,m}$ (m^{th} flight of the crew pairing assigned to flight k)
12: $\quad\quad\quad\quad cm' = m - 1 + \left(\frac{rd_{f,l}}{CT}\right)^n$ (number of cancellations before reserve with start time index l can be used plus cancellation measure due to reserve-induced delay when reserve is first used)
13: $\quad\quad\quad\quad$ **for** $n = 1$ to $N_{i,j}$ **do**
14: $\quad\quad\quad\quad\quad b$ = new feasible reserve instance
15: $\quad\quad\quad\quad\quad CM(b) = cm'$
16: $\quad\quad\quad\quad\quad V(b) = RUVI$ (index of new feasible reserve instance)
17: $\quad\quad\quad\quad\quad F_{i,j} = F_{i,j} \cup b$
18: $\quad\quad\quad\quad\quad R_{i,l} = R_{i,l} \cup b$
19: $\quad\quad\quad\quad\quad RUVI = RUVI + 1$
20: $\quad\quad\quad\quad$ **end for**
21: $\quad\quad\quad$ **end if**
22: $\quad\quad$ **end for**
23: $\quad$ **end for**
24: $\quad j = j + 1$
25: **end if**

index and scenario they are applicable to (line 18). These sets are useful later on when creating constraints for feasible reserve use in the $MIPSSM$ formulation.

Algorithm 19 is used in the disruption scenario generating simulation when a crew-related delay occurs. The algorithm stores the size of the disruption and then considers all of the possible reserve crew recovery actions and generates feasible reserve instances for each. Algorithm 19 differs from Algorithm 18 because of the type of disruption (delay rather than absence) and because of the possibility that, if they were used, feasible reserve instances generated for previous crew delay disruptions in the same simulation run could have reduced the current delay. If the current crew-related delay is a delay propagated from a previous crew-related delay, feasible reserve instances are generated corresponding to the reserve crew which could have been used to absorb the root crew-related delay also being used to cover the knock-on delay. These feasible reserve instances are stored in the set $G_{i,j}$. The number of reserve crew required to cover the given disruption in Algorithm 19 is the number of crew in the delayed crew team (line 6). The cancellation measure of the delay disruption when reserve crew are not available to cover the delayed crew is computed (line 7). The algorithm

Algorithm 19 Pseudocode for deriving disruption scenario information for a crew delay disruption occurring at simulation run i departure k

1: Inputs: Crew-related disruption affecting departure k of simulation run i (number of absent crew)
2: Outputs: Disruption j of scenario i $(cm_{i,j}, N_{i,j}, F_{i,j}, G_{i,j} ...)$
3: $RUVI$ = number of reserve use variable indices used so far
4: **if** crew delay disruption **then**
5: $\quad W_i = W_i + 1$
6: $\quad N_{i,j} = crewSize_k$
7: $\quad cm_{i,j} = \left(\frac{td_k}{CT}\right)^n$
8: $\quad$ **for** $l = 1$ to total hub departures **do**
9: $\quad\quad$ **if** reserve crew with start time D_l are feasible to absorb crew-related delay of departure k **then**
10: $\quad\quad\quad cm^{'} = \left(\frac{rd_{k,l}}{CT}\right)^n$
11: $\quad\quad\quad$ **for** $n = 1$ to $N_{i,j}$ **do**
12: $\quad\quad\quad\quad b$ = new feasible reserve instance
13: $\quad\quad\quad\quad CM\,(b) = cm^{'}$
14: $\quad\quad\quad\quad V\,(b) = RUVI$
15: $\quad\quad\quad\quad F_{i,j} = F_{i,j} \cup b$
16: $\quad\quad\quad\quad R_{i,l} = R_{i,l} \cup b$
17: $\quad\quad\quad\quad RUVI = RUVI + 1$
18: $\quad\quad\quad$ **end for**
19: $\quad\quad$ **end if**
20: $\quad$ **end for**
21: $\quad$ **if** current crew delay is crew delay propagated from the crew's previous flight **then**
22: $\quad\quad q$ =crew's previous flight
23: $\quad\quad o$ =disruption number of flight q
24: $\quad\quad$ **for** $l = 1$ to $|F_{i,o}|$ **do**
25: $\quad\quad\quad cm^{'} = \left(\frac{\max(0,\ td_k - (td_q - RD(F_{i,o,l})))}{CT}\right)^n$ (cancellation measure of the propagated delay if feasible reserve instance $F_{i,o,l}$ is used to absorb the root delay)
26: $\quad\quad\quad b$ = new feasible reserve instance
27: $\quad\quad\quad CM\,(b) = cm^{'}$
28: $\quad\quad\quad V\,(b) = RUVI$
29: $\quad\quad\quad U\,(F_{i,o,l}) = RUVI$ (reference to the knock-on effect reserve use variable)
30: $\quad\quad\quad G_{i,j} = G_{i,j} \cup b$
31: $\quad\quad\quad RUVI = RUVI + 1$
32: $\quad\quad$ **end for**
33: $\quad$ **end if**
34: $\quad j = j + 1$
35: **end if**

then considers each possible reserve start time (line 8) used to cover the delay. If the reserve start time is feasible (line 9) and can absorb the delay, $N_{i,j}$ $(= crewSize_k)$ new feasible reserve instances are generated (line 11) with unique reserve use variable indices and cancellation measures as calculated on line 10.

Lines 21 to 33 of Algorithm 19 apply if feasible reserve instances generated for the previous flight prevent or reduce the delay propagated to the current flight. For such feasible reserve use instances (line 24) $U\,(F_{i,o,l})$ (line 29) stores a new unique reserve use variable index corresponding to the same reserve being used to absorb the delay of the following flight. The reason why an extra reserve variable index is generated for the same reserve used on a following flight is that it is possible that other reserve crew might instead be used to cover the knock-on delay if the reserves used for the root crew-related delay do not absorb all of the delay and some delay can still propagate. The set G stores feasible reserve instances corresponding to feasible reserve instances which were generated for the root crew-related delay. Line 25 calculates the corresponding cancellation measures for these feasible reserve instances. The cancellation measure depend on the amount of delay that would have propagated if the feasible reserve instance corresponding to the root crew delay was utilised. The $MIPSSM$ has constraints that ensure that the beneficial knock-on effects can only apply if the reserve is actually used to absorb the root crew delay. After the disruption scenarios have been created they can be used to create the constraints and objective of the $MIPSSM$ formulation.

9.3 The MIPSSM's Mixed Integer Programming formulation

This section explains the mixed integer linear programming formulation. Table 9.4 defines the notation used, Section 9.3.1 presents and explains the objective and constraints.

x_l	:	Integer decision variable describing the number of reserve crew with start time index l (reserve crew schedule)
y_m	:	Binary decision variable describing whether (1) or not (0) a reserve crew with a particular duty start time is used for a particular disruption. This variable is referred to as a (the m^{th}) reserve use variable. There is one reserve use variable for each feasible reserve instance generated
$\delta_{i,j}$	:	Binary (output) variable describing whether or not disruption j in scenario i is left uncovered (1) or covered (0) by reserve crew in the $MIPSSM$ formulation
$\gamma_{i,j}$	:	Real valued (output) 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 the schedule

Table 9.4: $MIPSSM$ formulation notation

9.3.1 MIPSSM formulation

Once a set of disruption scenarios has been generated, they are used to form the objective and constraints of the $MIPSSM$ formulation. The $MIPSSM$ formulation finds the reserve crew schedule (x) that minimises the total cancellation measure over all disruption scenarios which were added to the formulation. The reduced cancellation measures that replace the original cancellation measures, that occurred in the disruption scenario generating simulations, depend on which reserve use variables (y) are selected to cover each disruption. Furthermore, which reserves can be used (Y) depends on which are scheduled (x).
Minimise:

$$\sum_{i=1}^{W}\sum_{j=1}^{W_i}\gamma_{i,j} \tag{9.5}$$

s.t.

$$\sum_{k=1}^{|F_{i,j}|} y_{V\left(F_{i,j,k}\right)} + \sum_{k=1}^{|G_{i,j}|} y_{V\left(G_{i,j,k}\right)} + \delta_{i,j} N_{i,j} = N_{i,j},\ \forall i \in 1..W,\ \forall j \in 1..W_i \quad (9.6)$$

$$\sum_{l=1}^{ND} x_l = TR \quad (9.7)$$

$$\sum_{k=1}^{|R_{i,l}|} y_{V\left(R_{i,l,k}\right)} \leq x_l,\ \forall l \in 1..ND,\ \forall i \in 1..W \quad (9.8)$$

$$y_{U\left(R_{i,l,k}\right)} \leq y_{V\left(R_{i,l,k}\right)},\ \forall k \in R_{i,l} | \exists y_{U\left(R_{i,l,k}\right)},\ \forall i \in 1..W,\ \forall l \in 1..ND \quad (9.9)$$

$$\delta_{i,j} cm_{i,j} \leq \gamma_{i,j}, \forall i \in 1..W,\ \forall j \in 1..W_i \quad (9.10)$$

$$y_{V\left(F_{i,j,k}\right)} CM\left(F_{i,j,k}\right) \leq \gamma_{i,j}, \forall i \in 1..W,\ \forall j \in 1..W_i,\ \forall k \in F_{i,j} \quad (9.11)$$

$$y_{V\left(G_{i,j,k}\right)} CM\left(G_{i,j,k}\right) \leq \gamma_{i,j}, \forall i \in 1..W,\ \forall j \in 1..W_i,\ \forall k \in G_{i,j} \quad (9.12)$$

$$y_m \in \{0,1\},\ \forall m \in Y \quad (9.13)$$

$$\delta_{i,j} \in \{0,1\},\ \forall i \in 1..W,\ \forall j \in 1..W_i \quad (9.14)$$

$$x_l \in \{0, 1 ... maxCA_l - 1, maxCA_l\},\ \forall l \in 1..ND \quad (9.15)$$

Objective 9.5 minimises the sum of all cancellation measures over all disruptions in all of the scenarios included in the model. The cancellation measures that apply for each disruption in each scenario depend on the decision variable vectors X and Y. Constraint 9.6 ensures that disruptions are only considered covered if the required number of reserve crew are used for the given disruption. Constraint 9.6 forces $\delta_{i,j}$ to 1 when no reserve recovery can be applied to disruption j in scenario i and to 0 otherwise. I.e. disruptions are only considered covered if each and every one of the disrupted crew affecting a flight are replaced. Constraint 9.7 ensures that no more than the total number of reserve crew available (TR) are scheduled. As the total number of reserve crew available for scheduling is assumed to be fixed. Constraint 9.8 ensures that in each disruption scenario the number of reserve crew used with the same start time index does not exceed the number of reserve crew which are scheduled to that start time index. This constraint prevents the same reserve crews being used for multiple disruptions in the same scenario, which is not allowed. Constraint 9.9 ensures that knock-on delays can only be absorbed by reserve crew if those reserve crew are actually used to cover the root delay. This constraint provides the mechanism through which this scenario-based approach to reserve crew scheduling allows for the possibility of reserve crew reducing the potential for downstream crew-related delays (which was the aim of the model presented in Chapter 7). Constraints 9.10 to 9.12 ensure that if reserve recovery actions are applied, the cancellation measure for the disruption is that which is associated with the latest available reserve crew used (as the flight can't take off before all of the crew are present). If no reserve crew are used for a given disruption, that disruption gets the cancellation measure $cm_{i,j}$

that occurred in the simulation run in which the disruption occurred. The indices of the y variables in Constraints 9.9, 9.11 and 9.12 make use of the functions U and V that return reserve use variable indices for any feasible reserve use instance passed as the argument, see Section 9.1.4. Constraints 9.13 to 9.15 are the integrality constraints for the indicator and decision variables of this model.

Section 3.1 cited the result that two-stage stochastic integer programming with discretely distribution parameters has #P-hard complexity. The $MIPSSM$ formulation is a two-stage stochastic integer program. The uncertainty of the second stage is modelled using scenarios sampled from Monte Carlo simulation which itself used discretely distributed random variables (for crew absence and journey times). The complexity of the $MIPSSM$ formulation under all possible disruption scenarios is therefore #P-hard. However, the act of including only a sample of all disruption scenarios means that the complexity is controlled by the number of scenarios that are included and the number of variables that are required to encode these scenarios in the formulation. The $MIPSSM$ approach is therefore to solve an approximation of the full problem.

9.4 MIPSSM modifications

This section firstly considers 2 alternative objective functions for the basic $MIPSSM$ formulation given by Equations 9.5 to 9.15. Then a scenario selection heuristic is presented, which has been 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.

9.4.1 Alternative objectives for the MIPSSM

$MiniMax1$

The objective of minimising the sum of cancellation measures over all disruption scenarios included in the model (Objective 9.5) could be replaced with the alternative objective $MiniMax1$ of minimising the largest sum of cancellation measures for any scenario. This is a minimax objective function, discussed in [103], and can be implemented by replacing Objective 9.5 with Objective 9.16 and adding Constraint 9.17. 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.

$$\text{min: } Z \tag{9.16}$$

$$\sum_{j=1}^{W_i} \gamma_{i,j} \leq Z, \forall i \in 1..W \tag{9.17}$$

$MiniMax2$

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 9.17 with Constraint set 9.18.

In the results (Table 9.6) 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.

$$\gamma_{i,j} \leq Z,\ \forall i \in 1..W,\ \forall j \in 1..W_i \tag{9.18}$$

9.4.2 Scenario Selection Heuristic

The basic $MIPSSM$ formulation and the two alternative formulations $MiniMax1$ and $MiniMax2$ are solved over a randomly generated 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 which should be included in the $MIPSSM$ formulation. 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 plenty of disruption scenarios are included in the model to increase the probability of including the important ones.

Algorithm 20 Pseudocode for the scenario selection heuristic

```
newScenarioFound = true
its = 0
while newScenarioFound ∧ its ≤ itLim do
   newScenarioFound = false
   rpts = 0
   while ¬newScenarioFound ∧ 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
```

The scenario selection heuristic given in Algorithm 20 is based on adding one disruption scenario to the model at a time and stopping when a new acceptable disruption scenario cannot be found within the iteration limit (*itLim*) (line 3), for which the sub-problem objective value (*subObj*) is larger than the objective contribution of the scenario already in the master problem with the largest objective contribution ($\max_j(masterObj_j)$). The sub-problem objective value of a new scenario is calculated (line 8) from the $MIPSSM$ formulation with the new scenario as the only input disruption scenario and with the incumbent reserve crew schedule (X) fixed. This heuristic is analogous to column generation in which the master problem and pricing problem are solved iteratively (described in Section 2.6.1). 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.

9.5 Optimal reserve use policy derivation

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 at flight q
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

Table 9.5: Notation for the $MIPSSM$ derived policy

The default reserve policy of the simulation 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. In the simulation, knowledge of future disruptions is not available, and so, if a scenario included in the $MIPSSM$ formulation repeats in the validation simulation, reserves might not necessarily be used in the same optimal way.

In this section, an algorithm for deriving a look up table reserve use policy, corresponding to a given reserve crew schedule, 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 numbers of reserve crew remaining for each departure for which using reserve teams to absorb crew-related delay is deemed globally beneficial. The threshold values are learned from repeat simulations in which the $MIPSSM$ is solved for the scenario generated by each single run of the simulation, with the given reserve crew schedule fixed. The aim is to learn the conditions under which reserve crew should be used to replace delayed crew.

The policy which is used for reserve crew use in response to crew absence is the default policy. The reasoning is that the penalty for not replacing absent crew with reserves is far too high (cancellation) to consider a crew absence reserve holding policy, and that the penalty of using teams of reserve crew 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 for delayed connecting crew is expensive as it solves a smaller disruption (delay compared to a cancellation) using more reserves than are usually required to cover absent crew. However, in certain circumstances using teams of reserve crew to cover delayed connecting crew can be globally beneficial in the long run as well as beneficial immediately.

This policy is optimal in a very limited sense, in that it is learned from optimal decisions derived from full knowledge of future outcomes, even though such information will not be available when the policy is used in the simulation. The algorithm used to learn the threshold numbers of reserve crew that allow reserves to be used to cover for delayed connecting crew is given in Algorithm 21.

Algorithm 21 Pseudocode for optimal reserve crew use policy derivation

Read in reserve schedule X
$n = 0$
$Rt_{0:ND} = R + 1$
$obs_{0:ND} = 0$
while $n < simRpts$ **do**
    Run simulation to generate disruption scenario $newScenario$
    Solve $MIPSSM$ for $newScenario$ with fixed reserve schedule
    $m = 0$
    while $m <$ disruptions in $newScenario$ **do**
        if disruption m is a delay and reserves were used to absorb it **then**
            deduce how many reserves were remaining at the time of the decision rr
            update policy
            $j =$ delayed departure number
            $Rt_j = ((Rt_j * obs_j) + rr)/(obs_j + 1)$
            $obs_j = obs_j + 1$
        end if
    end while
    $n = n + 1$
end while

In Algorithm 21 the policy's threshold values are calculated by repeatedly solving the reserve use variables of the $MIPPSM$ for different disruption scenarios with the given reserve crew schedule variables fixed (for the reserve schedule the policy is being derived for). The threshold value for a given flight is the average of the number of reserve crew remaining immediately before that crew delayed flight, averaged over instances where using a team of reserve crew was the recommended decision for that flight in the solution of the $MIPSSM$ formulation.

To implement the look up table reserve policy in the simulation, every time it is possible to use a team of reserve crew to absorb crew-related delay, check that the number of remaining reserves is greater than or equal to the corresponding element of Rt.

9.6 Experimental results

The $MIPSSM$ (Section 9.3.1), $MiniMax1$ and $MiniMax2$ (Section 9.4.1) and SSH (Section 9.4.2) approaches 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 and 6Gb of RAM.

These methods are also compared to alternative methods for reserve crew scheduling including:

Prob: this probabilistic approach is the SDM of Section 6.1.3. The SDM is used as the evaluator in a greedy algorithm (see Section 3.5.4).

Area: the area under the graph approach is the same as first introduced in Section 4.7.1.

USR: the uniform start rate heuristic approach was described in Section 3.5.3.

Zeros: this method was described in Section 3.5.3.

9.6.1 Experiment design

The reserve crew scheduling approaches listed above are used to derive reserve crew schedules for a generated test instance problem, which is described in this section. The input airline schedule features fully detailed crew connections and aircraft routings. Journey time uncertainty is modelled using 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 cockpit crew). The schedule is based on a 3 day single hub airline schedule with 243 flight legs a day with half of these being from the hub station and the other half back to the hub. The schedule uses 148 teams of crew 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 section following investigates the effect of the number of reserve crew available for scheduling for each solution approach.

9.6.2 Investigating the effect of varying the number of reserve crew available for scheduling

The results in Figure 9.4 show the effect on the average cancellation measure of varying the number of reserve crew available for scheduling, using 20000 repeat validation simulations for the reserve crew schedules from 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 9.4 shows how the various reserve crew scheduling approaches compare for different numbers of reserve crew available for scheduling. The

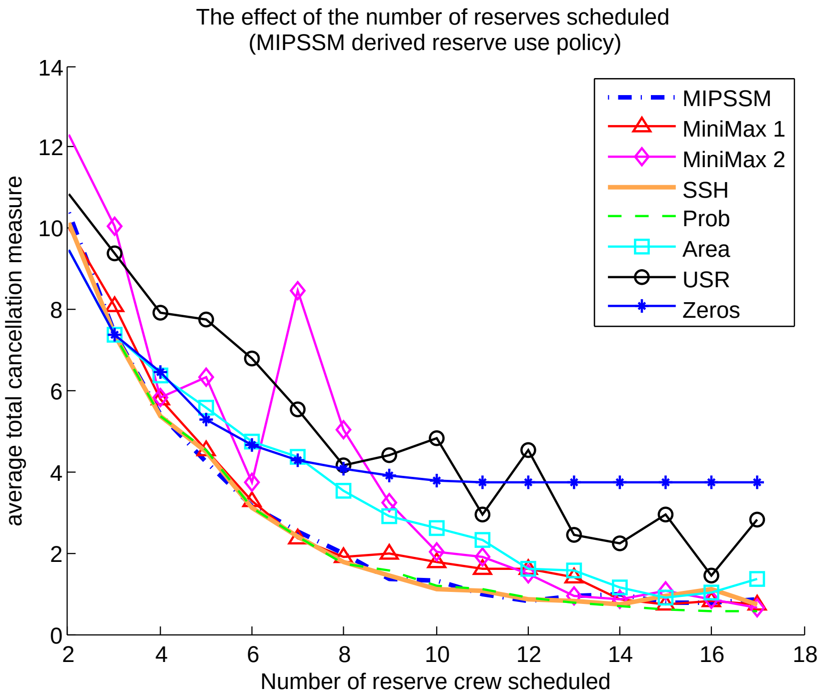

Figure 9.4: The effect of the number of reserve crew which are scheduled on the solution quality of different solution approaches

SSH, $MIPSSM$ and $Prob$ approaches obtain the lowest average cancellation measures of those tested for all numbers of available reserve crew. 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, especially in relation to the number repeat simulations the resultant schedules are tested in. The $MiniMax1$ modification generally leads 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 behaviour can be explained by the fact that the objective of the $MiniMax2$ modification is designed to suppress the single largest delay or cancellation disruption that can occur and is not to minimise the average cancellation measure. The $Area$ under the graph approach led to average cancellation measures similar to those from the $MiniMax2$ modification but without the fluctuations. The USR approach led 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.

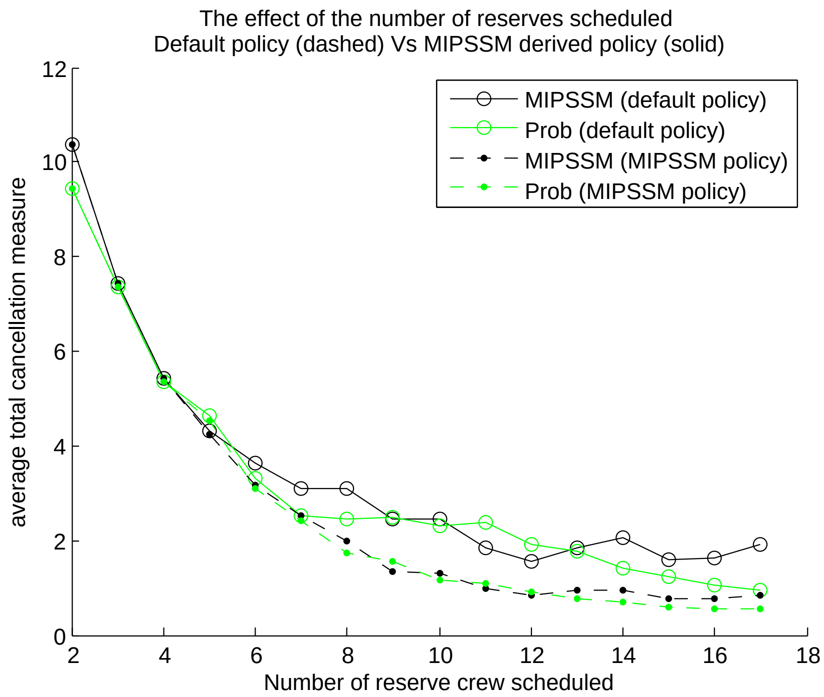

Figure 9.5: The effect of the $MIPSSM$ derived reserve use policy

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). In this range, Figure 9.4 shows that the best performing solution approach was the SSH. 10 to 12 reserve crew for the given problem instance is approximately proportionate to the number of reserve crew scheduled in reality.

Figure 9.5 shows the effect of using the $MIPSSM$ derived reserve use policy described in Section 9.5 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.

9.6.3 Other performance measures and solution reliability

Table 9.6 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 of Table 9.6 correspond to the case where 11 reserve crew are available for scheduling. The first column gives the methods which are being compared, the second column gives the average cancellation measures attained by each method. The third column gives the average delay calculated over the flights which experienced delays. The fourth column gives the probability that a flight is delayed by more than 30 minutes. The fifth column gives the probability a flight is cancelled. The sixth column gives the average reserve utilisation rate. The last column gives the average solution times.

Method name	Average cancellation measure	Average delay /mins	Probability of delay > 30mins	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
$MiniMax1$	1.246	12.393	0.00938	0.00154	0.666	7.060
$MiniMax2$	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 9.6: Performance measure averages from 20 repeats

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 28 minutes. The average cancellation measure can be interpreted as the number of cancellations expected in each of the simulations, but this also includes delays which have been converted to a cancellation measure using Equation 9.1. On the whole the SSH is a highly efficient approach with the lowest cancellation measure, a low average delay and a low solution time in comparison with the $MIPSSM$ approach. The low solution time of the SSH in comparison to the that of the $MIPSSM$ is a result of the termination criteria being satisfied, on average, before more than 10 disruption scenarios are added to the master problem. This result suggests that the SSH outperforms the $MIPSSM$ approach because it is possible to find a better reserve crew schedule with fewer input disruption scenarios, provided that some effort is made to find such a set of scenarios. The $Prob$ approach has the second lowest average cancellation measure, good average delay performance and a solution time much quicker than those of the $MIPSSM$ based approaches.

The results in Table 9.6 suggest that there is merit in both the probabilistic and $MIPSSM$ based approaches (SSH in particular) for scheduling airline reserve crew under uncertainty. Table 9.6 also includes performance measures when no reserve crew are scheduled at all as a point of reference. Contrary 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.

Figure 9.6 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. The percentile axis has exponential scale (cubed) for clarity, as this increases the linearity of the data. Figure 9.6 also displays the 100^{th} percentile (worst case) 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, so it appears that this modification does not achieve its goal. The reason for this is that $MiniMax2$ schedules reserve crew with respect to the worst case scenarios in a limited set of scenarios, so when a worst case scenario occurs in the validation simulation which is different from the worst case scenarios used to derive the reserve crew schedule, the reserve crew schedule performs worse than a reserve crew schedule aimed at the average case scenario.

Figure 9.6 demonstrates that for each given percentile the ordering of the methods supports the results given in Table 9.6 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. The $MiniMax2$ approach will only achieve it's goal if such worst case scenarios happen to be in the limited sets of scenarios. The other methods have relatively high worst case cancellation measures because they are aimed at the average case scenario.

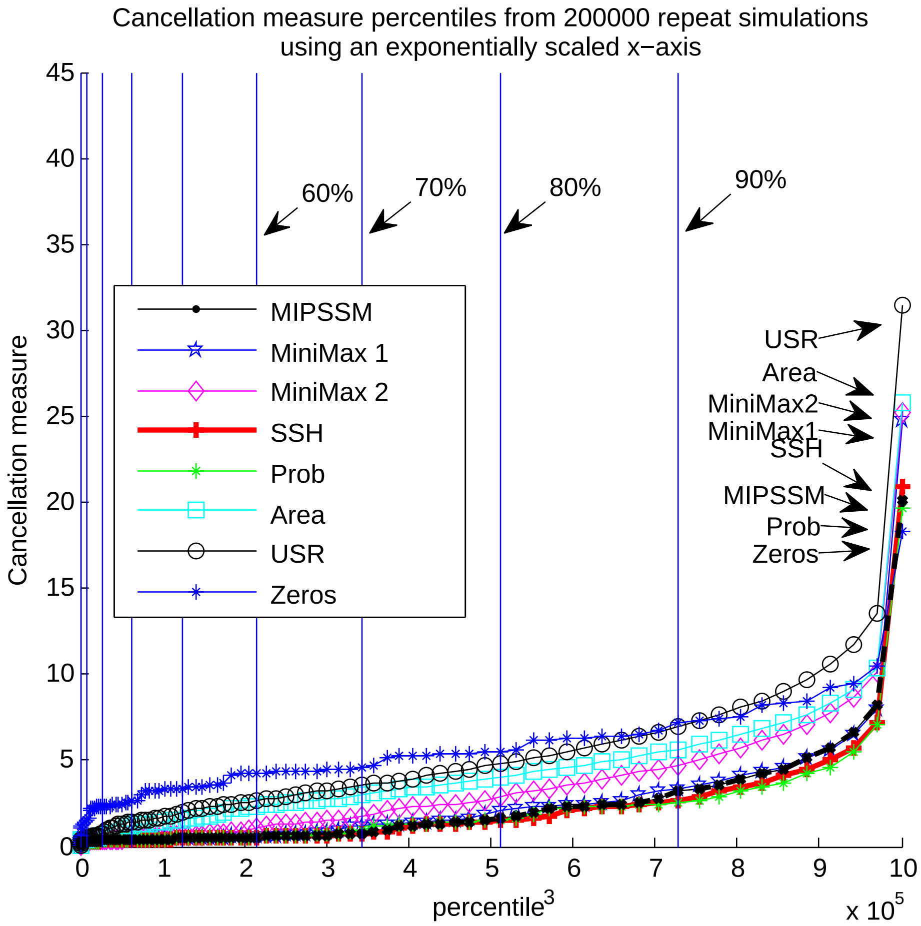

Figure 9.6: Percentile cancellation measures

Table 9.6 and Figure 9.6 show that the $MiniMax1$ and $MiniMax2$ approaches which were aimed at minimising the effects of worst case scenarios do not appear to have been effective in achieving this goal when considering the relatively high probabilities of delay over 30 minutes (Ta-

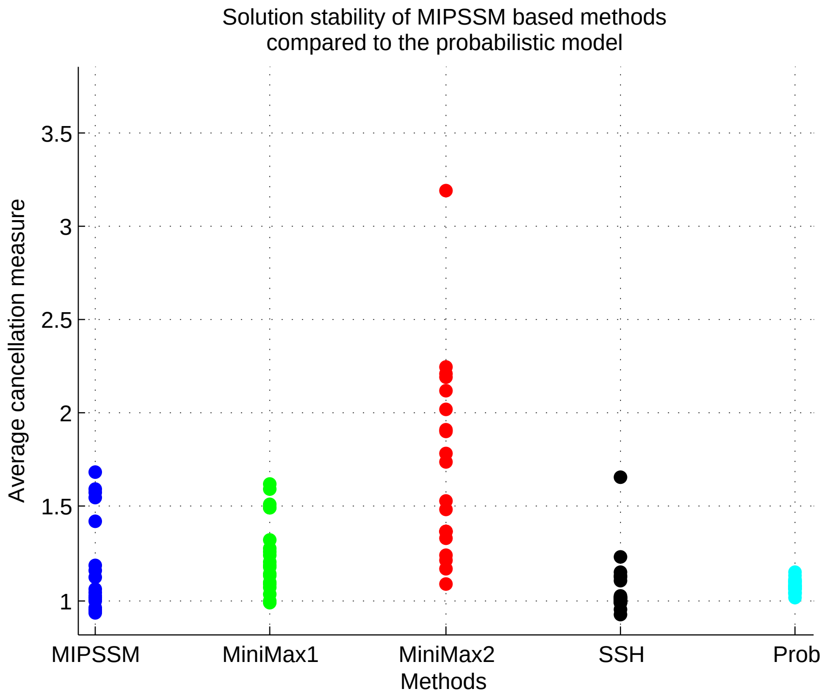

Figure 9.7: Solution reliability of $MIPSSM$ based methods compared to $Prob$

ble 9.6) and the 100^{th} percentile (worst case) cancellation measures (Figure 9.6) associated with these approaches. The possible explanation is that the best reserve crew schedule for one worst case is not the best reserve crew schedule for a different worst case scenario.

Each point on Figure 9.7 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 9.7 shows that the $MIPSSM$ based methods have a solution reliability issue. Figure 9.7 also shows that the $MIPSSM$ based methods have the potential to give solutions of higher quality than the probabilistic method ($Prob$), but this depends on the selection of disruption scenarios which are used as input for the given $MIPSSM$ based method. For this reason further research was performed to investigate the scenario selection mechanism, see Section 9.7.

9.7 The effect of scenario sets on reserve crew schedule quality

The basic $MIPSSM$ formulation requires a set of input disruption scenarios. This section attempts to address the issue of solution reliability illustrated in Figure 9.7, through careful selection of the scenarios added to the $MIPSSM$ formulation of Section 9.3.1. Disruption scenarios were generated randomly in the previous sections. In the case of the SSH, scenarios are selected if the cancellation measure for the new scenario is worse than the cancellation measure in any of the already selected scenarios, with the incumbent reserve crew schedule. This section investigates what makes a good set of scenarios. To answer this question attributes of sets of scenarios are defined. These are defined by the pool of scenarios that scenarios in the set belong to and the number of scenarios in the set. Three pools of scenarios are considered, and these are generated using the procedure outlined in Figure 9.8. Section 9.7.1 presents an investigation into the effect of the number of scenarios selected and the different types of pools from which they are selected of scenarios on the quality of reserve crew schedules derived from those sets of scenarios using the $MIPSSM$ formulation.

9.7.1 Attributes of sets of scenarios

As previously mentioned, the attributes of a set of scenarios are defined as the number of scenarios and the pool from which the scenarios are selected. Each pool of scenarios has a defining criterion for accepting scenarios into the pool.

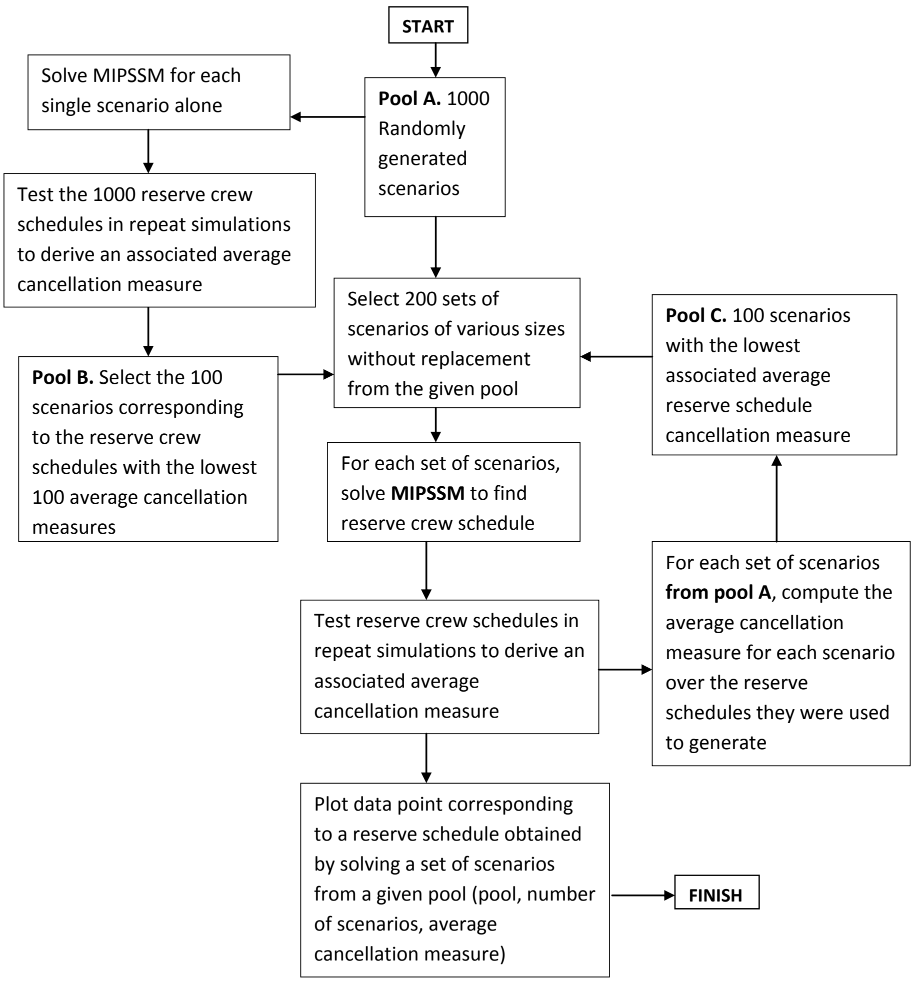

Figure 9.8: Flowchart of the population of three pools of scenarios

Pool A: 1000 random scenarios

Pool A consists of 1000 randomly generated scenarios.

Pool B: Good individual scenarios

Figure 9.8 shows how the two pools of scenarios B and C are derived from pool A. To create pool B, the first step is to solve the $MIPSSM$ formulation for each scenario in pool A on its own to obtain a reserve crew schedule corresponding to each scenario in pool A. Each reserve crew schedule corresponding to each scenario in pool A is then tested in the validation simulation to obtain an associated average cancellation measure. Pool B is then populated with the 100 scenarios from pool A which have the lowest associated average cancellation measures. Pool B represents scenarios, that when solved alone in the $MIPSSM$ formulation, give good reserve crew schedules.

Pool C: Good scenarios for sets

To create pool C, 200 sets of scenarios of various sizes are randomly sampled from pool A and solved in the $MIPSSM$ formulation. The reserve crew schedules corresponding to each set of scenarios are tested in the validation simulation to obtain associated average cancellation measures. Pool C is then populated with the 100 scenarios from pool A with the lowest average cancellation measures, where the average cancellation measure is calculated from the cancellation measures corresponding to the sample sets of scenarios they are a member of. Pool C represents scenarios that improve the quality of reserve crew schedules when added to a set of scenarios to be solved in the $MIPSSM$ formulation. Figure 9.8 outlines the process of populating pools B and C from pool A. Figure 9.8 also illustrates the process of deriving data points for Figure 9.9, which is designed to show the quality and variance of the quality of reserve crew schedules derived from sets of scenarios selected from each pool of scenarios.

9.7.2 Testing pools of scenarios

A total of 200 sets of scenarios were each selected, without replacement, from each pool (A, B and C), where the number of scenarios selected for each set is distributed uniformly between 5 and 45. The results in Figure 9.9 show the cancellation measures of the reserve schedules which were obtained by solving these scenario sets in the $MIPSSM$ formulation. Each data point in Figure 9.9 gives the number of scenarios in a set of scenarios used as input for the $MIPSSM$ (x-axis) and the cancellation measure of the resultant reserve crew schedule (y-axis), as derived from the validation simulation. The colour of the data point indicates which pool of scenarios the scenario set was selected from. The results displayed in Figure 9.9 show that the number of scenarios in a set is weakly negatively correlated with the average cancellation measure associated with the reserve crew schedule derived from that set of scenarios for all pools. I.e. Increasing the number

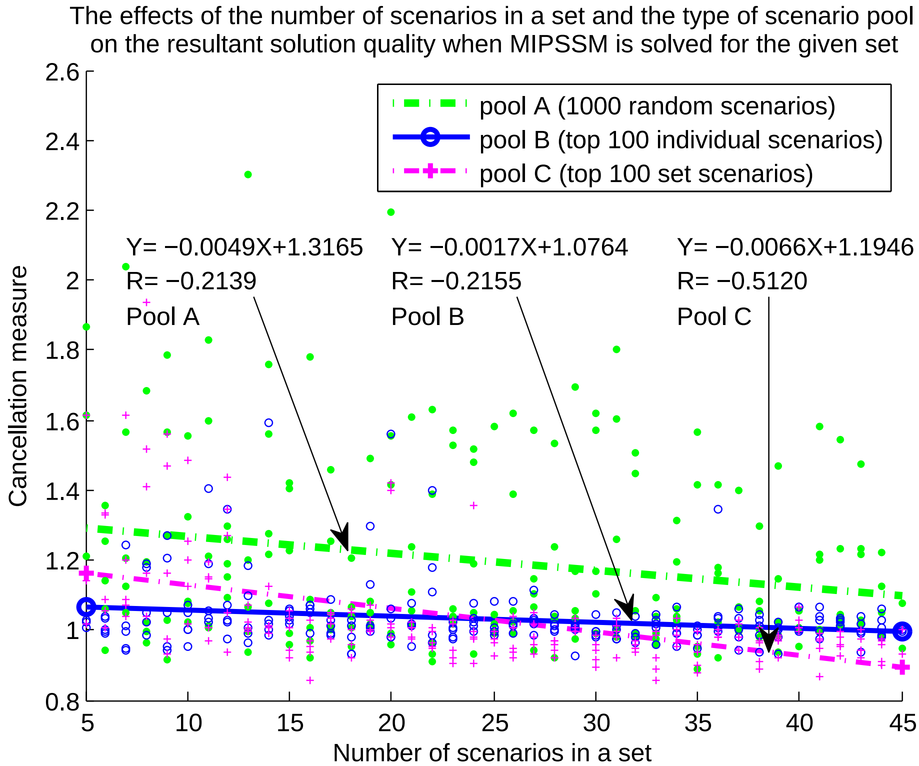

Figure 9.9: The effect of the pool from which scenarios are selected and the number of scenarios selected on the average cancellation measure associated with the reserve crew schedule derived from the given set of scenarios

of scenarios will decrease cancellations/delays. Figure 9.9 also shows that the quality of reserve crew schedules derived from sets of scenarios selected from pools B and C is on average greater than reserve crew schedules derived from sets of scenarios from pool A. Furthermore the quality of reserve crew schedules corresponding to sets of scenarios derived from Pool B is much less sensitive to the number of scenarios in those sets. This is intuitive as scenario pool B consists of scenarios that give good solution quality when solved alone. This also suggests that scenarios that work well as the single input for $MIPSSM$ do not necessarily lead to improved solutions when used together as a set of input scenarios for the $MIPSSM$. Figure 9.9 shows that the average cancellation measure of reserve crew schedules derived from sets of scenarios selected from pool C has the most convincing negative correlation (highest negative gradient and magnitude of correlation coefficient R) with the number of scenarios in those sets. This is also intuitive as pool C represents scenarios that improve reserve crew schedule quality when included in a set of scenarios.

The conclusion is that scenarios which are used as input for the $MIPSSM$ can be divided according to whether they work best as the sole input scenario (pool B) or whether they are scenarios that complement a pre-existing set of scenarios (pool C). The difference in the gradients of the regression lines corresponding to pools B and C in Figure 9.9 shows that pools B and C contain different scenarios. It is also interesting to note that the best result in Figure 9.9 occurred for a set of scenarios derived from pool C that only contained 16 scenarios. Increasing the number of scenarios beyond around 15 leads to an improvement in solution reliability for sets selected from pools B and C, however the same does not occur for the random scenarios of pool A. This is a positive result as solution reliability is one of the $MIPSSM$'s biggest problems (Figure 9.7).

9.7.3 Algorithms based on the results of the scenario set investigation

Best single scenario algorithm

To exploit these findings, one possible algorithm would involve finding the single scenario that leads to the highest solution quality. This could be a tractable approach as solution time is proportion to the number of scenarios in a set, with one scenario being solved very rapidly. Such a scenario can be said to have coincidental coverage.

Best set scenarios algorithm

Another algorithm would search for scenarios that work well as part of a set, however such an algorithm may be less scalable than the first suggested algorithm. The reason being that the measure used to populate pool C involves solving lots of sets of scenarios and testing the resultant reserve crew schedules, which can be very time consuming.

Probabilistic and $MIPSSM$ hybrid approach

The algorithms described above work by sampling disruption scenarios, and then using simulation testing to ascertain which disruption scenarios are associated with the best reserve crew schedules. However, in Chapter 8 it was shown that the $SDPM$ is an accurate evaluator of reserve crew schedules because it effectively replaces an infinite number of simulations with a single evaluation of the $SDPM$. So, replacing the simulation testing step with evaluations of the $SDPM$ should result in much more efficient algorithms, and in doing so, draws together the two main alternative approaches considered in the project, in algorithms which utilise the main strengths of each approach. I.e. reserve crew schedule generation in the case of the $MIPSSM$, and reserve crew schedule evaluation in the case of the $SDPM$. These hybrid algorithms were implemented and are tested in Chapter 10 on a range of schedule instances which include multiple fleets types and the ranks and qualifications of crew. But first, Section 9.8 presents the required fleets, ranks and qualifications extended $MIPSSM$ formulation.

9.8 Extending the MIPSSM approach for the case of fleets, crew ranks and qualifications

This section is analogous to Section 6.3 where the improved probabilistic crew absence model was extended to the case of multiple fleet types and the ranks and qualifications of crew. The extended formulation presented in this section is used in Chapter 10 when all approaches to reserve crew scheduling and policies are compared with one another in multiple fleet, crew rank and qualification test instances. Up until this point the $MIPSSM$ approach has been based on the simplified case of a single fleet and single crew rank. This meant that crew qualifications could be ignored as all crew considered were qualified for the single fleet. To recap, when considering multiple aircraft fleet types, crew qualifications have to be taken into account because crew will be qualified for a subset of the airline's fleets. Additionally different fleets have different crewing requirements. Cabin crew also come in a number of ranks, according to their training and experience. The highest cabin crew rank is purser and all flights require at least one purser. The consideration of crew ranks also gives rise to the possibility of "flying below rank", this means that higher rank crew rank are qualified to perform the duties of lower ranked crew.

The simulation disruption scenario generation procedure of Section 9.2 and the $MIPSSM$ formulation of Section 9.3 both require modifications in order to be applicable to the case of multiple fleets, crew ranks and qualifications. To incorporate fleets, crew ranks and qualifications, the notational changes given in Table 9.7 are required. FRQ is used as short hand for fleet, rank and qualification in the heading of Table 9.7.

9.8.1 Example problem

In the following and in Chapter 10, the extended $MIPSSM$ formulation is based on the case where there are 3 fleets, 3 reserve crew qualifications and 2 reserve crew ranks. The fleets that a reserveis qualified for are the same as those used in Table 6.3 of Section 6.3, which were captured by the simple expression (qualified if) $qualification \neq fleet$. The fleet crew requirements of the example problem are the same as those given in Table 6.4 of Section 6.3. $FCR_{fl,r}$, denotes the number of crew of each rank (r) each fleet (fl) legally requires.

9.8.2 Extended simulation generation of disruption scenarios

To extend the disruption scenario generation algorithms of Section 9.2.2 for the case of multiple fleets, crew ranks and qualifications the following changes need to be taken into account.

1. Reserve crew have distinct ranks and qualifications

FRQ formulation		
Single	Multiple	Description
x_l	$x_{l,q,r}$	Number of reserve crew with start time index l, qualification q and rank r
TR	$TR_{q,r}$	Number of reserve crew of each qualification (q) and rank (r) to be scheduled
$N_{i,j}$	$NL_{i,j}$	Number of low rank disrupted crew in disruption j of scenario i
$N_{i,j}$	$NH_{i,j}$	Number of high rank disrupted crew in disruption j of scenario i
$F_{i,j,k}$	$F_{i,j,q,r,k}$	The k^{th} feasible reserve instance of rank r, qualification q generated for disruption j of scenario i
$G_{i,j,k}$	$G_{i,j,q,r,k}$	The k^{th} feasible reserve instance of rank r, qualification q generated for disruption j of scenario i corresponding to a reserve crew used to absorb a secondary delay
$R_{i,l,k}$	$R_{i,l,q,r,k}$	The k^{th} feasible reserve instance of rank r, qualification q, with start time index l generated for scenario i

Table 9.7: Notational changes required for extending the $MIPSSM$ approach for the case of multiple fleets, crew ranks and qualifications

2. The number of disrupted crew for a given disruption is influenced by the fleet type
3. Reserve crew can only be used for the fleets they are qualified for
4. Reserve crew of high rank can if necessary by used to replace low rank disrupted crew

To account for reserve crew ranks and qualifications (Change 1) feasible reserve instances also have specific ranks and qualifications.

The effect of Change 2 is that the number of disrupted crew for a delay disruption depends on the crew requirements of the fleet. For crew absence disruptions there is also the possibility of different numbers of absent crew of each rank, for this purpose the notation $NH_{i,j}$ and $NL_{i,j}$ is introduced to denote the number of high rank and low rank crew disrupted crew respectively in disruption j of scenario i.

The effect of Change 3 is that feasible reserve instances have to be generated corresponding to each applicable qualification for a given disruption. No feasible reserve instances are generated corresponding to reserve crew with the incorrect qualification for a given disruption, this means that constraints are not required which state that reserves can only be used if qualified.

The effect of Change 4 is that the total number of feasible reserve instance generated for each applicable qualification is equal to the number of disrupted crew plus extra feasible reserve instances corresponding to high rank reserve crew to allow for the possibility of high rank qualified reserve crew flying below rank.

The sets used to store the feasible reserve instances F, G and R are extended with extra dimensions for the ranks and qualifications of feasible reserve instances (see Table 9.7).

9.8.3 Extended MIPSSM formulation

For the case of multiple fleets, crew ranks and qualifications the $MIPSSM$ formulation (Section 9.3) has to be modified. The following gives the modified constraints, stating which constraint it replaces in the original formulation.

$$\sum_{q=1}^{3}\left(\sum_{k=1}^{|F_{i,j,q,2}|} y_{V\left(F_{i,j,q,2,k}\right)} + \sum_{k=1}^{|G_{i,j,q,2}|} y_{V\left(G_{i,j,q,2,k}\right)}\right) + \delta_{i,j} NH_{i,j} = NH_{i,j} \quad \forall i \in 1..W,\ \forall j \in 1..W_i \tag{9.19}$$

$$\begin{gathered}\sum_{q=1}^{3}\sum_{r=1}^{2}\left(\sum_{k=1}^{|F_{i,j,q,r}|} y_{V\left(F_{i,j,q,r,k}\right)} + \sum_{k=1}^{|G_{i,j,q,r}|} y_{V\left(G_{i,j,q,r,k}\right)}\right)\\ + \\ \delta_{i,j}\left(NH_{i,j} + HL_{i,j}\right)\\ = \\ NH_{i,j} + HL_{i,j}\\ \forall i \in 1..W,\ \forall j \in 1..W_i\end{gathered} \tag{9.20}$$

Constraint 9.6 of the original formulation is replaced by Constraints 9.19 and 9.20. Constraints 9.19 and 9.20 ensure that the crew-related disruption is only considered covered if all disrupted crew are replaced with reserve crew. Additionally, these constraints allow for the possibility of reserve crew flying below rank because constraint 9.19 ensures that all high rank disrupted crew are replaced with high rank reserve crew, whilst Constraint 9.20 allows the low rank disrupted crew to be replaced with reserve crew of any rank.

$$\sum_{l=1}^{ND} x_{l,q,r} = TR_{q,r},\ \forall q \in 1..3,\ \forall r \in 1..2 \tag{9.21}$$

Constraint 9.7 of the original formulation is replaced with Constraint 9.21, which states that the total number of reserve crew of each rank and qualification scheduled must equal the total number of reserve crew of each rank and qualification available for scheduling. In Constraint 9.7 the equality could be replaced with $\leq$ if scheduling all available reserve crew is not a strict requirement.

$$\sum_{k=1}^{|R_{i,l,q,r}|} y_{V\left(R_{i,l,q,r,k}\right)} \leq x_{l,q,r},\ \forall q \in 1..3,\ \forall r \in 1..2,\ \forall l \in 1..ND,\ \forall i \in 1..W \tag{9.22}$$

Constraint 9.8 of the original formulation is replaced with Constraint 9.22, which states that the total number of reserve crew of each rank (r) and qualification (q) scheduled to each start time index (l) used to cover crew disruptions in each scenario (i) cannot exceed the number of reserve crew scheduled with the same start time index (l), rank (r) and qualification (q).

$$y_{U\left(R_{i,l,q,r,k}\right)} \leq y_{V\left(R_{i,l,q,r,k}\right)}$$
$$\forall k \in R_{i,l,q,r} | \exists y_{U\left(R_{i,l,q,r,k}\right)}, \forall q \in 1..3, \forall r \in 1..2, \forall i \in 1..W, \forall l \in 1..ND \tag{9.23}$$

Constraint 9.9 of the original formulation is replaced with Constraint 9.23 which states that reserve crew of each rank and qualification can only absorb knock-on crew-related delays if those reserve crew are used to absorb the root delay. Constraints 9.11 and 9.12 of the original formulation are replaced with equivalent constraints where the notation changes specified in Table 9.7 are applied and whilst also adding $\forall q \in 1..3$ and $\forall r \in 1..2$ to the combinations of indices those constraint sets are generated for.

Constraint 9.15 of the original formulation stated that the number of reserve crew scheduled to each start time index had to be an integer between 0 and the maximum number of disrupted crew affecting the departure associated with the given start time index. In the extended formulation this constraint is removed on the grounds that reserve crew of different qualifications can be scheduled to each start time index, but reserve crew of different qualifications are feasible for different fleets which have different crew requirements. Therefore such a constraint may preclude finding the best reserve crew schedule for a given set of disruption scenarios. In the extended $MIPSSM$ formulation, the reserve schedule x is constrained by Constraint 9.21 only.

Constraints 9.10, 9.13 and 9.14 are unchanged for the case of multiple fleets, crew ranks and qualifications.

9.8.4 Results

Chapter 10 compares the extended $MIPSSM$ formulation and each of its variants with the alternative approaches to reserve crew scheduling explored in this project, over a number of realistic problem instances.

9.9 Future work

9.9.1 Iterative solution approach to $MIPSSM$

One of the weaknesses of the $MIPSSM$ approach is that solution times increase dramatically as the number of input disruption scenarios increases. The reason for this is the large number of binary variables in the resultant $MIPSSM$ formulation. One possible approach to tackling this issue is to use an iterative solution approach in which some variables are fixed in each iteration. The variables of the $MIPSSM$ formulation can be divided according to reserve schedule variables (X) and reserve use variables (Y), therefore one possible approach is to alternative between fixing the reserve crew schedule (namely X) and fixing which disruptions are covered using reserve crew in each iteration (fixing certain Y variables to zero). The desired outcome of such an approach is that the reserve crew schedule converges to the optimum corresponding to the set of input disruption scenarios.

9.10 Chapter summary

This chapter has introduced an alternative approach to reserve crew scheduling, to that of the the probabilistic approaches of Chapters 5 to 8, in the form of a scenario-based approach to airline reserve crew scheduling. The main idea of which 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.

A range of alternative objective functions for the $MIPSSM$ formulation were tested, it was found that minimising the sum of all cancellation measures over all disruption scenarios was the most effective. A minimax objective function even failed to minimise the worst case scenario that occurred in the simulation testing of solutions. The reason for this was because the optimal solution for the worst case scenario in a limited sample was typically not a good solution for a different worst case scenario.

The SSH approach showed that the choice of 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 fewer input disruption scenarios. 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. Further investigation of the effect of selecting scenarios from pools of scenarios with particular characteristics revealed the existence of scenarios that lead to good quality reserve crew schedules when used as the single input scenario for the $MIPSSM$ formulation. Such scenarios were said to have a high level of coincidental coverage. In contrast evidence was found for the existence of scenarios that lead to good quality reserve crew schedules when used as one of a set of input scenarios from the same pool.

The $MIPSSM$ formulation was also extended to the case of multiple fleet types, crew ranks and qualifications with relatively few modifications to the initial model.

Chapter 10

Comparison of all reserve crew scheduling and policy approaches

In this chapter the approaches to reserve crew scheduling and reserve policies introduced in previous chapters are applied to the same test instances and compared with one another.

Chapter structure

The test instances are described in Section 10.1. Section 10.2 defines the approaches to reserve crew scheduling and the configurations of those approaches that are to be tested in this chapter. Section 10.3 describes the two phase experiment design. The first phase (Section 10.4) tests all approaches to reserve crew scheduling considered in this thesis, in a variety of configurations, with repetitions of each. The aim is to find representative reserve crew schedules for comparing all methods in the second phase. The second phase (Section 10.5) takes the best reserve crew schedule from each general reserve scheduling approach found in phase 1 and tests each in conjunction with each of the reserve policies considered in this thesis. This will allow an overall comparison of all of the reserve crew scheduling approaches and reserve policies considered in this thesis. Section 10.6 then summarises the main findings from this chapter.

10.1 The test instances

The test instances are based on data provided by KLM. The aircraft routings are the same as those in the data provided by KLM, whilst the crew schedules were generated using a set partitioning model solved in CPLEX. The test instances are based on a 3 fleet example. The fleet crewing requirements are those given in Table 6.4. There are crew/reserve crew with 2 ranks and 3 qualification types, making 6 different grades of crew. The 3 qualification types of crew were defined in Table 6.3. Table 10.1 summarises the six test instances considered in this chapter. The first three (columns) test instances are classed as real schedules, which means that the

Type	Real			Tightened real		
Schedule	1	2	3	4	5	6
Days	1	3	7	1	3	7
Hub deps	139	424	983	139	424	983
Total deps	280	849	1970	280	849	1970
Crew	123	195	272	120	190	262
Crew s.t. abs.	71	136	180	68	131	170
Aircraft	71	75	78	71	75	78
Reserve crew	10	14	16	10	14	16
Delay risk	0.04262	0.01051	0.01190	0.4997	0.4990	0.4960
Crew CR	0.4331	0.4196	0.4216	0.375	0.4124	0.4103

Table 10.1: Test instance properties

aircraft routings and the scheduled departure and arrival times are exactly the same as in the source data. The real schedules have a low risk of delay. The tightened real schedules are based on the same data but with minimum ground times scheduled, and the allocated journey times are equal to the average journey times. This explains the increased, approximately 0.5, probability of delay for the tightened schedules. The purpose of the tightened schedules is to show that the methods introduced in this thesis do not rely on the input airline schedule being very conservative in terms of the risk of delay propagation. The low risk of delay in the real schedules is a result of the use of hub banks, where an airline schedules sequences of arrivals together to allow passengers to connect to different aircraft before a subsequent sequence of departures. The low risk of delay ensures that passengers have a good chance of successfully catching their connecting flights. Figure 4.2 of Chapter 4 gives an example of a real airline schedule where hub banking is visible.

For both the real and the tightened schedules there are three schedules ranging in length from 1 day to 7 days. The length of the schedule represents the time horizon over which the set of available reserve crew can be scheduled. In each test instance reserve crew pairings span 3 days, this constraint only becomes important in the 7 day schedules. Table 10.1 also gives the number of departures, crew, crew which may be subject to crew absence, aircraft, reserve crew and the crew connection rate (CR) or mid shift aircraft change rate. The crew who are subject to crew absence correspond to the crew who are stationed at the hub station, crew absence affecting crew stationed at spoke station are covered by reserve crew at spoke stations (see Assumption RP6 of Section 4.2, the *deadheading is not a viable option for solving delay and unexpected crew absence disruptions assumption*). The number of reserve crew available for scheduling are chosen so that there is at least one of each rank/qualification combination available and the number of reserve crew of each type is 1.5 times the expected number of crew absence of each type. The expected number of absent crew is based on the assumed 1% chance (also used in Section 6.2.1) that any given member of crew is absent for their assigned crew pairing. The main focus in this thesis is on how the set of available reserve crew can best be scheduled and used,

Evaluator	SPCAM	SDM	CDM	SDPM
Thesis chapter	5	6	7	8
Absence model				
Simplified	✓			
Detailed		✓	✓	✓
Delay model				
None	✓			
static		✓	0.5	
Dynamic			0.5	✓
Reserve delays		✓	✓	✓
All delays				✓
Delay propagation			0.5	✓
Assumed reserve policy				
Absence only	✓	✓		✓
default			✓	
GRP	✓	✓	✓	✓

Table 10.2: Details of the probabilistic models used as evaluators in various search methodologies

see Section 2.3 for information about work regarding reserve crew sizing.

10.2 Description of the reserve crew scheduling approaches being compared

10.2.1 Probabilistic approaches

The reserve crew scheduling approaches based on the probabilistic models of Chapters 5 to 8 are each defined by a combination of a search methodology and a probabilistic reserve crew schedule evaluator. This section describes the search methodologies, and Table 10.2 details the features of the probabilistic evaluators. Table 10.2 shows that each probabilistic evaluation method corresponds to a different chapter of this thesis. The acronyms for each method were all defined in the acronyms section at the beginning of this thesis. The SPCAM (simplified probabilistic crew absence model) evaluator uses a simplified model of crew absence uncertainty, in which it is assumed that a maximum of one crew member can be absent from each crew pairing. The SPCAM evaluator does not take delay into account and assumes the absence only reserve policy.

The SDM (static delay model) evaluator uses the fully detailed model of crew absence uncertainty, with a static estimate of delays, which it uses to estimate the delays which are introduced when using reserve crew to cover for absent crew. The SDM assumes the absence only policy.

The CDM (crew delay model) evaluator is very similar to the SDM evaluator except that the model of delay takes the propagation of crew related delays into account and the effect that teams of reserve crew can have when used to replace such delayed connecting crew. The CDM evaluator is partially dynamic, in that the model of delay propagation changes in response to the probabilities that reserve crew have been used for particular crew related disruptions. The CDM of Chapter 7 was designed purely for reserve crew used for crew related delay disruptions, so the CDM evaluator is actually a combination of Chapters 6 and 7. The CDM evaluator assumes the default reserve policy, if it assumed the absence only policy CDM would reduce to SDM.

The SDPM evaluator uses the fully detailed model of crew absence uncertainty and a fully detailed model of delay propagation uncertainty. The SDPM is fully dynamic which means that the predicted delays of flights are responsive to the probabilities that reserve crew are used for each flight, in terms of any direct or indirect knock-on delay that may result from reserve crew use. The SDPM evaluator takes reserve-induced delay into account as well as propagated delays. The SDPM evaluator is also capable of modelling delay recovery using swaps and teams of reserve crew (see Section 8.1.4), however as was shown in Section 8.2.5 doing so resulted in little benefit in terms of reserve crew schedule quality and also vastly increases the time required to evaluate a single reserve crew schedule, which in turn inhibits the efficiency of any search algorithm it is used with. As a result of this, these features of the SDPM will not be utilised in this chapter. The SDPM evaluator assumes the absence only policy (because the reserve team use model will be switched off).

Search heuristics for reserve crew scheduling using a probabilistic model as the evaluator

In this section the probabilistic models of Chapters 5 to 8 are tested and compared with one another in terms of the quality of reserve crew schedules which can be derived whilst using those models as the evaluation function in a variety of search heuristics. The search heuristics tested in this section include:

- **GH: Greedy heuristic**. Implemented as described in Section 3.5.4.
- **LS: Local search**. Implemented as described in Section 3.5.4, with a cut and insert neighbourhood.
- **SA: Simulated annealing**. Implemented as described in Section 3.5.4.
- **GA: Genetic algorithm**. Implemented as described in Section 3.5.4, with a population size of 50, 4 competitor tournament selection, a crossover rate of 1 and simulated annealing based approach to mutation (see Section 3.5.4).

These frequently used search methodologies are all discussed in more detail in Sections 3.5.4 and 5.4.1. All of the approaches, except for the greedy heuristic, were limited to a maximum solution time of 10 minutes.

The *SDPM* experiments use the optimal trade-off interval sizes derived in Section 8.3.2. The GRP (generalised reserve policy) parameters used are those which were derived in Section 6.4.4.

The following experiments were implemented on a laptop with a 2.4GHz dual core Intel Core i7-5500U CPU, with 8Gb of RAM. All models, algorithms and the simulation were implemented in Java as single threaded applications.

10.2.2 MIPSSM based approaches

The reserve crew scheduling approaches based on the MIPSSM of Chapter 9 that will be tested and compared with alternative approaches, in this chapter, are as follows.

- **MIPSSM:** The method descibed in Section 9.3 with 15 randomly generated input scenarios.
- **SSH:** The method described in Section 9.4.2 with a limit of 200 repeat simulations to find each new scenario.
- **SingleScen:** Best single scenario algorithm, which was described in Section 9.7.3, with 1000 scenarios are generated and evaluated for each repeat.
- **SetScen:** Best set scenarios algorithm, which was described in Section 9.7.3, with 1000, 300 and 100 samples of 10, 5 and 3 scenarios respectively for 1, 3 and 7 day schedules (for tractability reasons). For tractability reasons sampling phase uses the iterative solution approach alluded to in Section 9.9.1.

Each repetition of each method (applied to the six test instances defined in Table 10.1) is limited to a maximum of 1 hour to find a solution. The experiments were carried out using IBM CPLEX Optimization Studio version 12.5 as the MIP solver, on a desktop computer with a 2.79GHz Core i7 processor and 6Gb of RAM.

10.2.3 Other approaches

This thesis also considers a number of rule of thumb heuristic and simulation based approaches to reserve crew scheduling. In Section 10.5 the best reserve crew schedule from the probabilistic and the MIPSSM based approaches are compared with the uniform start rate heuristic (USR) of Section 3.5.3, and the area under the graph approach (Area) of Section 4.7.1.

10.3 Experimental design

The following experiment is divided into two phases. Phase 1 will perform repeat experiments for each of the probabilistic (Section 10.4.1) and scenario-based (Section 10.4.2) approaches. The aim in phase 1 is to find reserve crew schedules that represent the best possible solutions that can be derived from these approaches. This approach vastly reduces the number of reserve crew schedules which have to be tested in phase 2, in conjunction which the very computationally intensive SDPM and SIM reserve polices.

Phase 2 will allow for a comparison of the best approaches to reserve crew scheduling and reserve policies considered in this thesis.

In phase 1 the probabilistic and MIPSSM approaches defined in Section 10.2 are applied to the test instances defined in Table 10.1 with 20 repetitions in each case. The repetitions are required because most of the approaches involve stochastic inputs at one point or another. 20 repetitions gives 20 chances to find a representative solution from each approach whilst allowing the experiments to be performed in a timely manner. Furthermore, a method that requires more than 20 repetitions to find a representative solution has the disadvantage of being unreliable. The search heuristics used in conjunction with probabilistic model evaluators, such as simulated annealing, genetic algorithms and the initial solutions for local search involve stochastic inputs. The scenarios of the MIPSSM based approaches are derived from simulations which use random inputs. The resultant reserve crew schedules are then each tested in 20000 repeat simulations in conjunction with a subset of all of the reserve policies considered in this thesis. Appendix Section G.1 shows that 20000 repeat simulations gives reliable average cancellation measures using a 10-fold cross-validation analysis. The default policy, absence only policy and the look up table (LUT, as described in Section 4.7.3) policy will be used, whilst the SIM and the SDPM policies will not be applied until phase 2 (as described above).

10.4 Phase 1: Reserve crew scheduling results

This section gives results for the probabilistic and MIPSSM approaches applied to test instances 1 to 6. The best reserve crew schedules from each approach will be used to represent the corresponding approach in Section 10.5 when all approaches to reserve crew scheduling and reserve policies are compared.

10.4.1 Probabilistic model results

When analysing the results for the probabilistic approaches the main questions are: 'what is the most effective method of evaluation?' and 'what is the most effective solution methodology?' To answer these questions the average cancellation measures (cancellations plus cancellation measures of delays) derived from simulation testing were plotted for each test instance, reserve policy, evaluator and solution methodology. Figure 10.1 shows the dot plots corresponding to test instance 1 with the absence only policy. Equivalent results were obtained for the other policies. i.e. the choice of online policy between the default policy, absence only policy and the LUT policy does not change the ordering of the evaluators and search methodologies used for reserve crew scheduling.

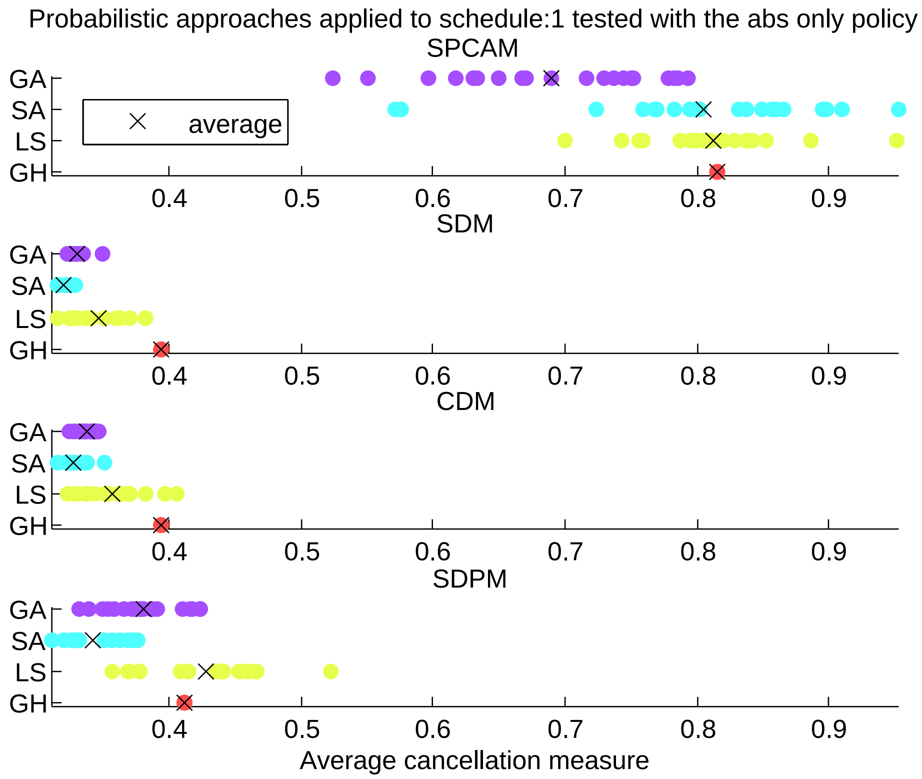

Figure 10.1: The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 1 with the absence only policy

Figure 10.1 shows that for test instance 1, reserve crew schedules derived from the SPCAM evaluator used with the absence only reserve policy result in the highest cancellation measures. In fact, this pattern is repeated for each test instance and each reserve policy. The equivalent plots for test instances 2 to 5 are given in appendix Section G.2. Furthermore, excluding the SPCAM results, the local search approach is on average dominated by simulated annealing and the genetic algorithm in all cases. Again, this is a result which occurred for each test instance and reserve policy tested (see appendix Section G.2). Figure 10.1 also shows that the variance of the cancellation measures is relatively high for the SDPM evaluator compared to the cases when the SDM and CDM evaluators are used. This can be attributed to the increased evaluation times for the SDPM model, which for test instance 1 were roughly 50 times larger than the other evaluators (0.125 seconds as opposed to 0.06 seconds), meaning that fewer evaluations could be performed within the 10 minute time limit. Despite this some good reserve crew schedules are obtained from the SDPM evaluator.

Dominated approaches: local search and the SPCAM evaluator

The next two figures sequentially eliminate the local search methodology and the SPCAM evaluator from consideration, which will allow for a closer examination of the differences between the remaining approaches.

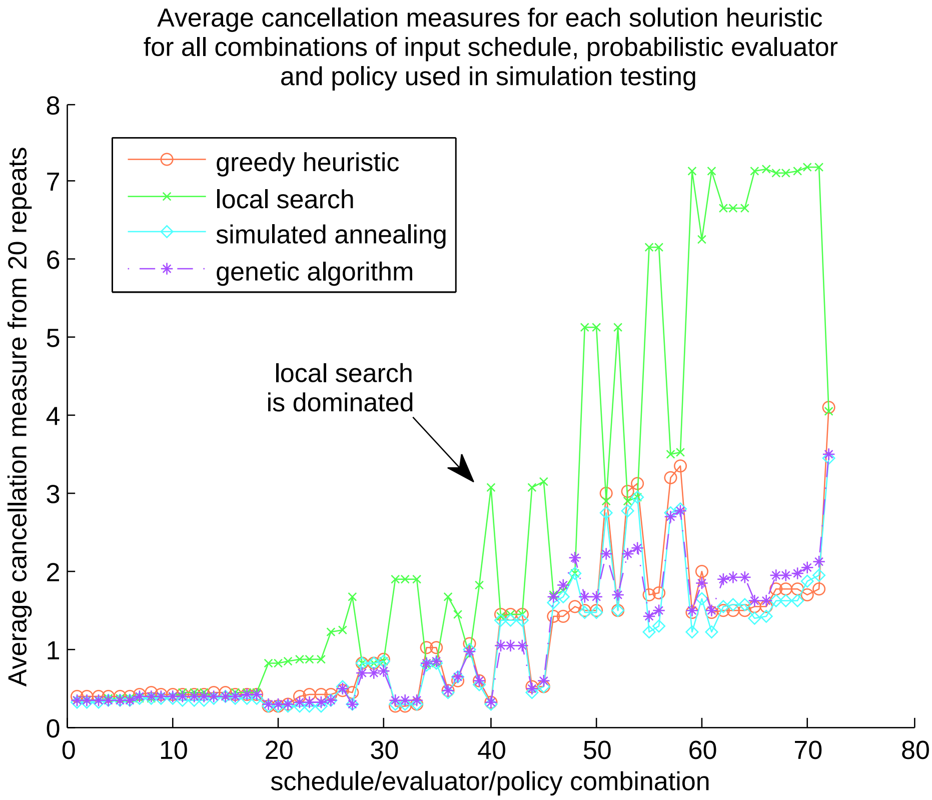

Figure 10.2: Cancellation measures for each solution methodology for all combinations of test schedule, reserve policy and evaluator, in descending order of the average cancellation measure for each combination of test schedule, reserve policy and evaluator

In Figure 10.2 each location on the x-axis corresponds to a combination of a test schedule, a probabilistic evaluator and reserve policy used in simulation testing, the information on which combination is which has been anonymised because the main aim is to demonstrate that the local search method is dominated by the other solution methodologies in all cases and can therefore be eliminated from consideration at this stage. Figure 10.2 shows that the local search approach never found the best reserve crew schedule in any case, and that in fact they gave the worst solutions in the vast majority of cases. The reason why the local search approach was much less effective compared to simulated annealing, genetic algorithm and the greedy heuristic is due to the size of the search neighbourhood which has to be enumerated at each stage before a new solution can be accepted and the process repeated. The probabilistic model is a very detailed evaluator and this is reflected in the time required per evaluation. For a local search based approach to be an effective search methodology the size of the local neighbourhood needs to be small enough given, the cost/time requirements of evaluations and a sufficient amount time in which to reach a local optimum. These criteria are not met for the case of the probabilistic model, especially for the larger test schedule instances and especially as a 10 minute solution time limit had been imposed for each repeat experiment.

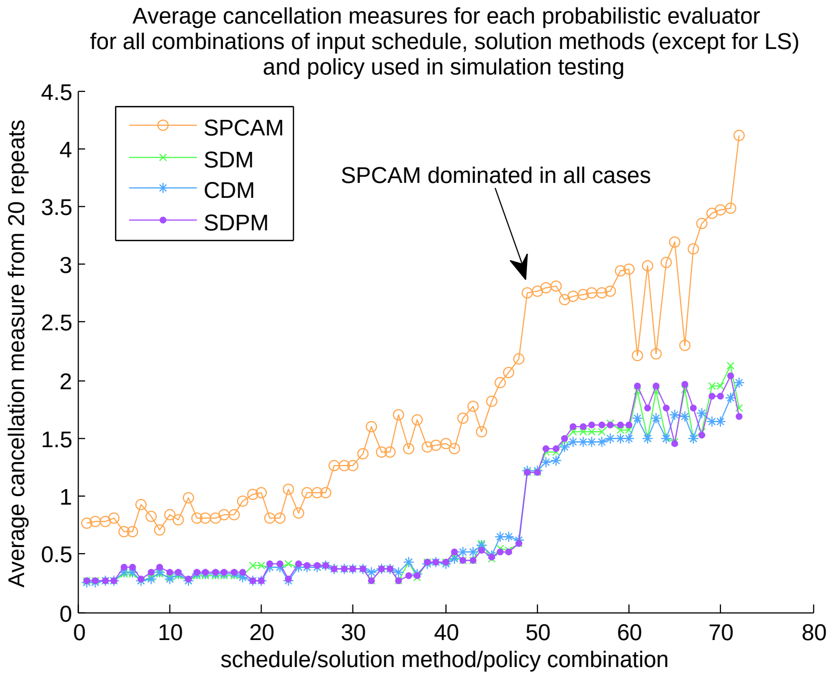

Figure 10.3: Cancellation measures for each probabilistic evaluator for all combinations of test schedule, reserve policy and solution methodology, in descending order of the average cancellation measure for each combination of test schedule, reserve policy and solution methodology

Figure 10.3 shows that after the elimination of the dominated solution methodology (local search) it is apparent that the SPCAM evaluator is on average dominated by each of the other probabilistic evaluators for all combinations of test schedule, reserve policy and solution methodology. This result was expected because the SPCAM evaluator corresponds to the initial simplified probabilistic model which used a simplified model of crew absence uncertainty and made no attempt to model the effect that reserve crew can have on expected delays.

Delay and cancellation performance of the remaining approaches

Having eliminated the local search methodology and the SPCAM on the grounds of consistently low reserve crew schedule quality, the remaining approaches are examined in terms of their associated delay and cancellation performance measures. Figure 10.4 shows the average delay and cancellation rates corresponding to the reserve crew schedules derived using the SDM, CDM and SDPM evaluators for test instances 1 and 4. The results for test instances 2, 3, 5 and 6 are not shown because test instance 1 captures the results of test instances 2 and 3, whilst test instance 4 captures the results of test instances 5 and 6. I.e. a pattern emerges with regard to the performance of different evaluators which depends on the type of test instance being solved: real or tighter generated schedules.

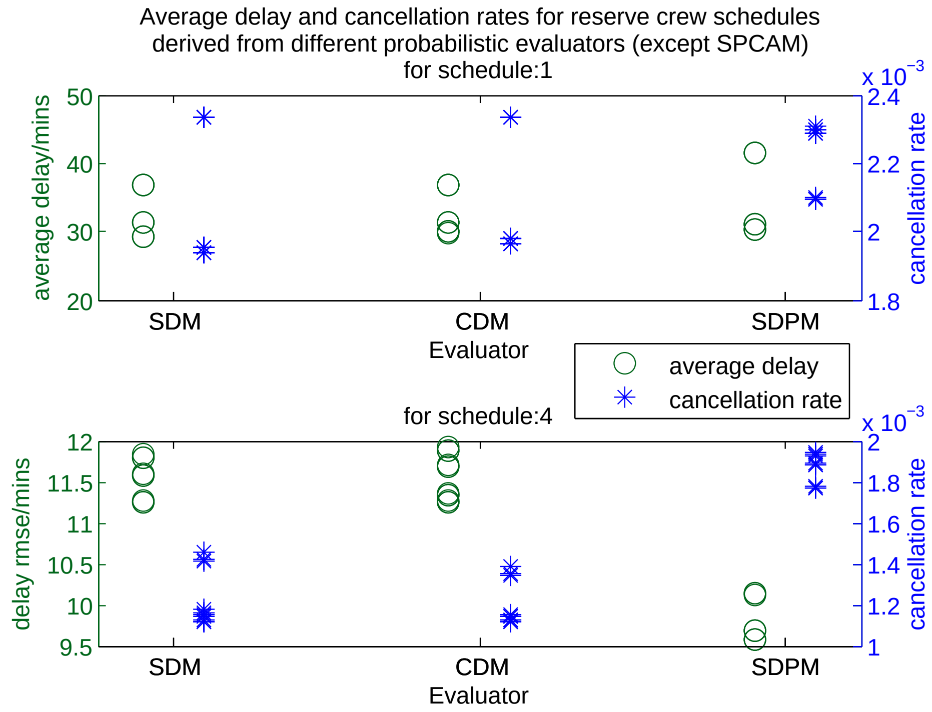

Figure 10.4: Average delay and cancellation performance of the reserve crew schedules derived using the SDM, CDM and SDPM probabilistic evaluators

In Figure 10.4 the data has been anonymised in terms of the solution methodology and reserve policy that each data point corresponds to, as the purpose here is to highlight a trend that occurs regardless of these features. Figure 10.4 shows that the average delay and cancellation rate performance of reserve crew schedules derived using the SDM, CDM and SDPM evaluators are closely matched for test instance 1. However for test instance 4 (the tightened version of test instance 1) the reserve crew schedules derived using the SDPM evaluator attain a minimised average delay at the expense of an elevated cancellation rate, without a significant increase in the cancellation measure. This pattern is repeated in pairwise fashion for test instances 2 and 5, and instances 3 and 6 (see appendix Section G.3). The explanation for why the SDPM evaluator leads to reserve crew schedules which minimise average delays for the tightened real schedules is that the SDPM evaluator uses the SDPM, a model which allows for delays from all causes including propagated delays. The tightened schedules have an elevated risk of delay propagation and so the reserve crew schedules derived from the SDPM evaluator are primed for delay minimisation. Note that the CDM evaluator only allows for crew related delays and their propagation, the vast majority of delayed connecting flights in the tightened real schedules are due to delayed aircraft rather than crew, which is because aircraft operate all day whereas crew operate on a relatively small number of flights before new crew begin and typically have to wait for a delayed aircraft. The real strength of the SDPM evaluator lies in its ability to model the propagation of reserve-induced delays that occur when using reserve crew to cover for absent crew. Additionally, the SDM and CDM use static models of delay for reserve-induced delays, that occur when covering for absent crew, but do not allow for their propagation.

Figure 10.4 also shows that test instance 1 has much larger average delays than test instance 4, even though instance 4 has the elevated risk of delay. The explanation for this is that in test instance 1 all delays were reserve-induced delays, which can be very large, whereas in test instance 4 many small delays also occurred due to journey times exceeding the allocated block time and then the minimal ground time being unable to prevent these delays from propagating. So, although the average delay was lower in test instance 4, the total delay was higher in test instance 4 than in test instance 1. On a similar note the cancellation rates are typically higher in test instance 1 than in test instance 4, which can be explained by test instance 4 being a tightened schedule. This means that in test instance 4 reserve crew will typically be feasible for a larger number of flights that may be affected by crew absence, which reduces the expected number of cancellations due to crew absence.

Results for probabilistic model based approaches

Table 10.3 gives the average and minimum cancellation measures for the non-dominated probabilistic reserve crew scheduling approaches. In Table 10.3 each number is an average (minimum) cancellation measure over 60 reserve crew schedules derived using the same combination of evaluator and solution methodology. The 20 repeats for each combination of evaluator and solution methodology were each tested with 3 reserve policies. Table 10.3 does not give results for the local search methodology or the SPCAM evaluator as these have already been shown to be dominated approaches. Table 10.3 shows that in general the simulated annealing solution method provides the reserve crew schedules with the lowest average and minimum cancellation measures for all six test instances. However, there is no single probabilistic evaluator which always provides the reserve crew schedules with the lowest cancellation measures. In fact, across the six test instances each evaluator is associated with a lowest minimum or lowest average cancellation measure at one time or another.

The best (probabilistic model based) reserve crew schedules for each schedule are judged to be those which correspond to the lowest minimum cancellation measures for each schedule in Table 10.3. In Section 10.5 those best reserve crew schedules are used to test all reserve policies and also to compare the probabilistic approaches with the alternative approaches to reserve crew scheduling.

10.4.2 MIPSSM results

The results in Figure 10.5 show the average cancellation measures for the 20 repeats of each MIPSSM based method (see Section 10.2.2) tested in conjunction with 3 different reserve policies (60 data points per method) for each test instance.

	GH mean (min)	SA mean (min)	GA mean (min)
		schedule 1	
SDM	0.3944 (0.3944)	**0.3206** (0.3153)	0.3299 (0.3229)
CDM	0.3944 (0.3944)	0.3270 (0.3153)	0.3377 (0.3244)
SDPM	0.4110 (0.4110)	0.3419 (**0.3117**)	0.3809 (0.3317)
		schedule 2	
SDM	0.4038 (0.3945)	0.2643 (**0.2578**)	0.3044 (0.2703)
CDM	0.2746 (0.2729)	**0.2633** (0.2584)	0.2896 (0.2643)
SDPM	0.2759 (0.2742)	0.2778 (0.2621)	0.3402 (0.2901)
		schedule 3	
SDM	1.5028 (1.5017)	1.5589 (1.4831)	1.9104 (1.7627)
CDM	1.5028 (1.5017)	**1.4827** (**1.3977**)	1.6816 (1.5149)
SDPM	1.7568 (1.7567)	1.6106 (1.4635)	1.9540 (1.6294)
		schedule 4	
SDM	0.4282 (0.4259)	**0.3742** (**0.3478**)	0.3925 (0.3682)
CDM	0.4245 (0.4218)	0.3748 (0.3501)	0.3948 (0.3599)
SDPM	0.4247 (0.4241)	0.3745 (0.3492)	0.4019 (0.3596)
		schedule 5	
SDM	0.4963 (0.4264)	0.4203 (0.2592)	0.4502 (0.2554)
CDM	0.4866 (0.4251)	0.5008 (0.3068)	0.4831 (0.2801)
SDPM	0.4555 (0.3196)	**0.4139** (**0.2557**)	0.4634 (**0.2577**)
		schedule 6	
SDM	1.5928 (1.4638)	1.5155 (1.1602)	1.7480 (1.3005)
CDM	1.8039 (1.7012)	**1.3877** (1.1356)	1.5904 (1.3038)
SDPM	1.5581 (1.4600)	1.4957 (**1.1074**)	1.7198 (1.2307)

Table 10.3: Average cancellation measures for each solution methodology and probabilistic evaluator

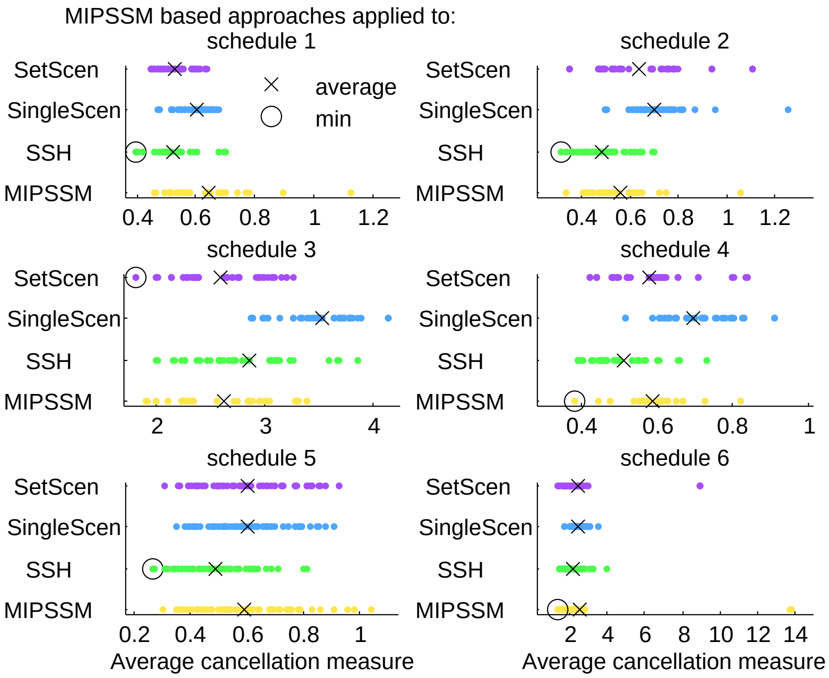

Figure 10.5: Average cancellation measures for MIPSSM based approaches

Figure 10.5 shows that the SetScen algorithm approach always outperforms the SingleScen algorithm in terms of both the average and minimum cancellation measure attained. The SSH has the lowest average cancellation measure in 5 out of 6 of the test instances and the lowest minimum average cancellation measure in half of the test instances. The MIPPSM attained the overall best solutions for test instances 4 and 6, whilst the SetScen algorithm attained the best overall solution for test instance 3.

Figure 10.5 also shows that the variance of the quality of solutions from each method across all schedules are roughly similar, none of the methods have an outstanding level of reliability. This is a general weakness of the scenario-based approach to reserve crew scheduling. Another contributing factor to the unreliability of MIPSSM based approaches is the imposed time limit for finding a solution, which meant that the MIP solutions were not always proven to be optimal, but instead an optimality gap still existed as the time limit ran out. The use of a time limit is in a practical context a reasonable constraint. In summary the MIPSSM based approaches have an aspect of unreliability due to the requirement that only a limited set of disruptions can be included in the model before it becomes computationally intractable and the time required to solve the resultant problems are effectively random variables in any given case. In future, to address these issues, the following issues need to be considered: more efficient solution methods; model simplification and further refinement of the scenario selection procedure. Section 12.1.2 discusses these issues in more detail.

The results 10.5 also demonstrate that when aircraft fleets, crew ranks and qualifications are modelled and larger schedule instances are considered, the increase in solution reliability of the SetScen and SingleScen algorithms that was demonstrated in Figure 9.9 is reduced.

The MIPSSM derived best reserve crew schedule for each test instance (circles in Figure 10.5) will be used to represent the MIPSSM based approach in Section 10.5 when the effects of different reserve policies are considered.

10.5 Phase 2: Comparing all approaches for reserve crew scheduling and reserve policies

In this section, the best reserve crew schedules for each test instance from the probabilistic and MIPSSM approaches are tested in conjunction with a larger variety of reserve policies. Figure 10.6 also gives results for the other two main types of reserve crew schedule, those of the simulation based and heuristic based approaches. The reserve policies being tested are:

- default: Default policy (see Section 3.5.2). Use reserve crew as soon as any demand occurs.
- abs only: Absence only policy (see Section 3.5.2). Use reserve crew to cover absence as demand occurs, but never cover for delayed crew.
- LUT: Look up table (see Section 4.7.3). Evaluate all alternative reserve decisions using a look up table.
- SIM: Simulation policy (see Section 4.7.2). Evaluate all alternative reserve decisions by running simulations starting from those decisions. With 1000 repeat simulations to evaluate each alternative.
- SDPM: Statistical delay propagation model policy (see Section 8.2.5). This implementation uses the $SDPM1$ version of the policy to evaluate all alternative reserve decisions.

Of these reserve policies, the default and abs only policies are rule of thumb policies. The LUT policy represents a learning based approach, the SIM policy represents a direct application of simulation as a reserve policy and the SDPM policy represents a theoretical probabilistic approach. The LUT, SIM and SDPM are policies are all based on evaluating the effect that each available reserve decision has on the expected cancellation measure that will be accumulated from the given time until the end of the considered time horizon and then choosing the decision with the lowest associated cancellation measure. The SIM and SDPM policies are more computationally intensive than the other policies. This is the reason they were not used in Section 10.4 when testing multiple repeat reserve crew schedules for each approach to reserve crew scheduling.

The results in this section are based on testing the best reserve crew schedules from each general approach to reserve crew scheduling, for each test instance, in conjunction with the reserve policies listed above. The general approaches to reserve crew scheduling are: probabilistic (Prob); MIPSSM; simulation based; and rule of thumb approaches. The area under the graph approach (Area) of Section 4.7.1 represents a simulation based approach. Whilst the uniform start rate heuristic (USR) (Section 3.5.3) represents a rule of thumb approach to reserve crew scheduling. For each test instance, each combination of a reserve crew schedule and a reserve policy will be tested in the same 5000 repeat simulations. This helps to make the comparison a fair one. The results of this experiment are given in Figure 10.6 and Tables 10.5, 10.6 and 10.7.

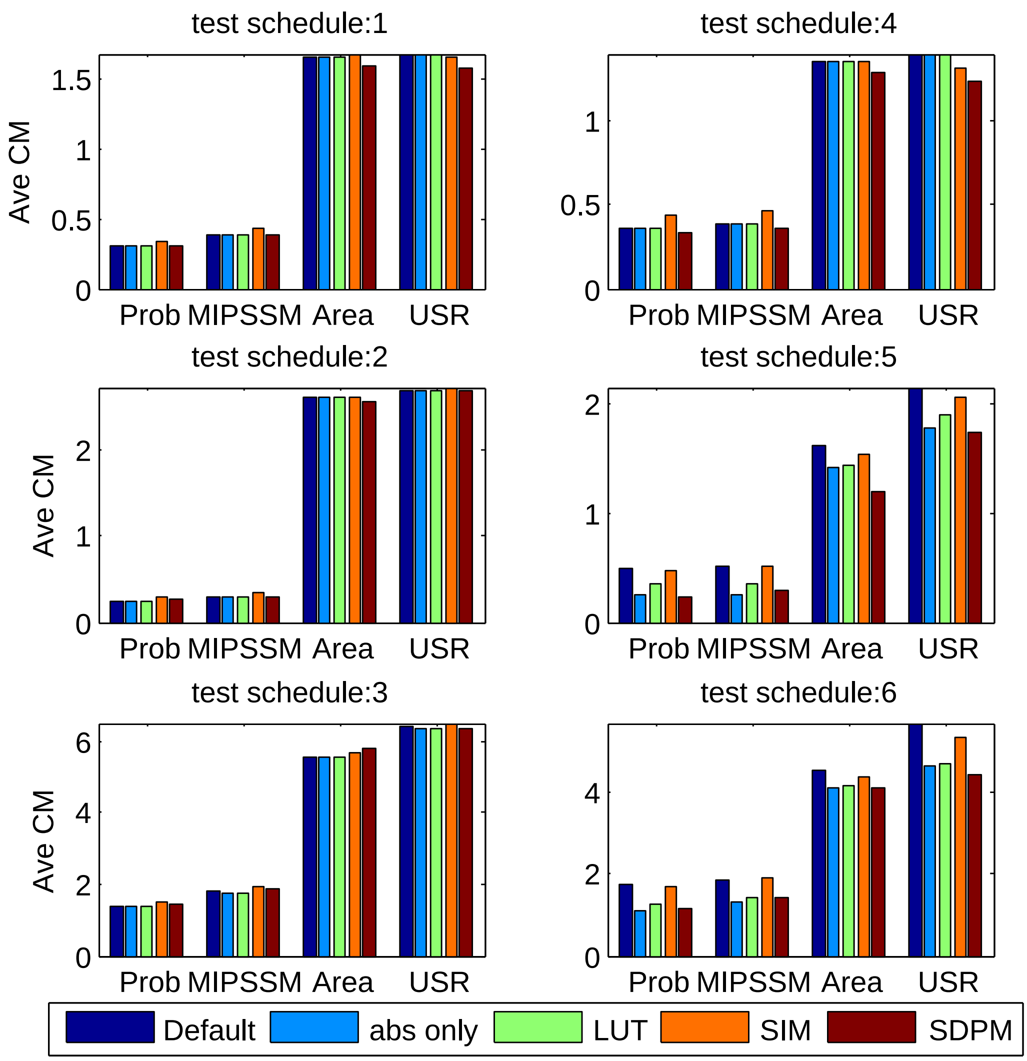

Figure 10.6: Average cancellation measures for all tested approaches to reserve crew scheduling and reserve policies, for each test instance

Figure 10.6 shows the average cancellation measures attained by each reserve crew schedule and reserve policy combination for each test instance. Colours correspond to reserve policies and groups of adjacent bars correspond to the same approach to reserve crew scheduling. Figure 10.6 shows that the probabilistic and MIPSSM approaches to reserve crew scheduling lead to average cancellation measures far lower than those of the Area approach and the USR heuristic regardless of the reserve policy used. The reason for this is that the probabilistic and MIPSSM approaches use all available information to schedule reserve crew whereas the Area and USR approaches do not. The difference between the probabilistic and MIPSSM approaches when used with different policies is not very clear in Figure 10.6, Tables 10.5, 10.6 and 10.7 show more clearly the differences between the these approaches to reserve crew scheduling when used in conjunction with different reserve policies.

Figure 10.6 is arranged into subplots with 3 rows and 2 columns. Rows 1, 2 and 3 correspond to test instances of length 1, 3 and 7 days respectively. The left hand column corresponds to the test instances which have a low risk of delay. The right hand column contains the corresponding test instances which have an increased risk of delay. This arrangement highlights the result that reserve policies become more important as the risk of delay increases, which is seen as an increase in the variance of the average cancellation measures across the range of policies tested. This is because there will be more instances where critical reserve use decisions are required, such as whether or not (and if so, which) reserve crew should be used to cover for delayed crew. The default policy suffers the largest increase of average cancellation measure for the test instances with a heightened delay risk. The reason for this is that using reserve crew for every disruption that occurs runs the risk of using reserve crew unnecessarily for small disruptions (such as small delays) when they could have been held for larger disruptions (such as replacing absent crew to prevent flight cancellations). In contrast, the difference between the performance of different reserve policies is minimal for the case of the low delay risk test instances. The reason for this is that fewer reserve crew are required in the low delay risk test instances and therefore reserve holding, as considered by the LUT, SIM and SDPM policies, will be rarely beneficial.

Another pattern that can be seen in Figure 10.6 is that the SDPM policy leads to significant reductions in average cancellation measure when used in conjunction with a relatively poor quality reserve crew schedule, such as those corresponding to the area under the graph and uniform start rate heuristic approaches. On the other hand, the total benefit of using the SDPM policy is reduced if the reserve crew schedule it is used with is of a relatively high quality. These observations can be attributed to the fact that the high quality reserve crew schedules are typically derived with knowledge of an assumed reserve policy as an input, i.e. they are typically designed to work well with simple reserve policies. This means that it could be possible to find further performance improvements by increasing the sophistication of the reserve policy that is assumed when scheduling reserve crew.

A result that is also visible in Figure 10.6 is that the SIM policy never achieves the best average cancellation measure. Additionally, the SIM policy attains particular high average cancellation measures for test instances 5 and 6, i.e. the longer schedules with an increased delay risk. An explanation for this is that in longer schedules where there are more disruptions, the number of simulations required to accurately evaluate each alternative decision also increases. So the SIM policy used an insufficient number of repeat simulations to evaluate alternative decisions in these cases.

This result reiterates an advantage of the SDPM policy over the SIM policy, it only requires one evaluation to obtain predictions with an accuracy equal to that obtained from an arbitrarily large number of repeat simulations.

Figure 10.6 also shows that the abs only policy and the LUT policy are very reliable policies, given their relative simplicity, with average cancellation measures that are below average in most cases. The good results for the abs only policy can be attributed to it being a risk averse policy. It never replaces delayed crew with reserve crew, which avoids future cancellations due to crew absence that were avoidable. In contrast, the LUT, SIM and SDPM policies allow for the possibility of taking risky decisions, such as using reserve crew to cover for delayed crew and holding reserve crew in the event of crew absence, if this is expected to minimise the average cancellation measure in the long term.
Due to low solution quality, the Area approach and the USR heuristic are discarded at this point and no further results for these approaches are given hereafter.

10.5.1 Applying policies in instances of crew absence

The results for the LUT, SIM and SDPM policies in Figure 10.6 and Tables 10.5, 10.6 and 10.7 are each taken from the best of two variant applications. Those where the policies are used for all reserve use decisions (to cover crew absence and delayed crew) and those where the policies are only applied in cases where reserve crew can be used to cover for delayed crew, whilst always covering for crew absence disruptions as they occur. Table 10.4 states the number of times these policy variants outperformed one another for each policy type in each test instance. The test instances are arranged in increasing length order. Table 10.4 shows that for the LUT policy,

		Not applied at absence (NA) Applied at absence (AA)		
days in schedule	test schedule	LUT NA:AA	SIM NA:AA	SDPM NA:AA
1	1	0 : 4	1 : 3	0 : 4
	4	0 : 4	2 : 2	2 : 2
3	2	0 : 4	4 : 0	0 : 4
	5	0 : 4	2 : 2	1 : 3
7	3	1 : 3	3 : 1	3 : 1
	6	0 : 4	3 : 1	2 : 2
	Total	1 : 23	15 : 9	8 : 16

Table 10.4: Number of times two policy variants outperform one another

applying the policy for all reserve decisions nearly always resulted in the lowest average cancellation measure, only one exception to this was found. For the SIM policy it was found that in the majority of cases it was beneficial to only apply the policy in cases where reserve crew could be used to cover for delayed crew. In contrast to the SIM policy, for the SDPM policy it was found that in the majority of cases using the policy to make all reserve decisions was beneficial. Table 10.4 also shows that for the SIM and SDPM policies, applying those policies to make all reserve decisions tended to become less beneficial as the schedule length increased. The explanation for this is that in the longer schedules, risky decisions such as holding reserve crew in the event of crew absence, has a higher penalty, that of a larger number of cancellations, because crew pairings are longer in those schedules. So in such schedules, a reserve policy with a risk averse approach to crew absence reserve decisions tended to work best. Additionally, it is rare for reserve holding to be a beneficial decision in the event of crew absence. Even if a policy determines that the expected value of holding is greater, in a particular scenario it may in hindsight still turn out to be the wrong decision. This is due to the variance of the outcomes of events on any given day. Giving the best results from two alternatives for each of the LUT, SIM and SDPM policies (as in Figure 10.6 and Tables 10.5, 10.6 and 10.7) helps to indicate the potential of these general approaches to reserve decision making in their best configurations.

10.5.2 Results tables

Tables 10.5, 10.6 and 10.7 display a variety of performance measures and statistics for the best probabilistic and MIPSSM derived reserve crew schedules for each test instance, tested in conjunction with each reserve policy. The performance measures given include: cancellation measure; reserve utilisation rate for absence and delay disruptions; average delay; probability of delay greater than 15 minutes; cancellation rate; and the maximum cancellation measure recorded over 5000 repeat simulations. Each of Tables 10.5, 10.6 and 10.7 gives the results for two test instances of the same length, one based on actual scheduled departure times and the corresponding schedule with a heightened risk of delay. In Tables 10.5, 10.6 and 10.7 an asterix after a policy name (LUT*, SIM* and SDPM*) indicates that the results for that policy correspond to the variant of that policy where reserve holding is never considered in the event of crew absence disruptions (see Section 10.5.1).

Test instance	Reserve schedule	Policy	Cancellation measure	Reserve use rate		Average delay	Probability of delay (> 15 mins)	Cancellation rate	Max CM
				absence	delay				
1	Probabilistic	default	0.3050	0.3126	4.000E-4	28.53	0.005715	**0.001881**	**6.172**
		abs only	0.3051	0.3126	0	28.55	0.005722	**0.001881**	**6.172**
		LUT	0.3050	0.3126	2.400E-4	28.54	0.005718	**0.001881**	**6.172**
		SIM	0.3404	0.3101	0	28.97	0.005822	0.002121	7.172
		SDPM	**0.3012**	0.3122	0	**24.71**	**0.005453**	0.001964	**6.172**
	MIPSSM	default	0.3911	0.3099	0.006560	41.77	0.006873	0.002071	7.665
		abs only	0.3892	0.3101	0	41.44	0.006983	**0.002053**	7.665
		LUT	0.3888	0.3101	0.002880	41.54	0.006932	**0.002053**	7.665
		SIM	0.4258	0.3079	0.002200	41.63	0.007118	0.002296	8.665
		SDPM	**0.3820**	0.3087	0.003620	**36.71**	**0.006381**	0.002222	**6.295**
4	Probabilistic	default	0.3611	0.3018	0.01944	10.96	0.1000	0.001134	9.619
		abs only	0.3585	0.3020	0	10.96	0.1003	**0.001114**	9.619
		LUT	0.3588	0.3019	0.00832	10.95	0.1002	0.001118	9.619
		SIM	0.4354	0.2946	0.00310	**10.64**	0.0992	0.001803	**7.261**
		SDPM*	**0.3328**	0.3011	0.00352	10.87	**0.0843**	0.001191	9.616
	MIPSSM	default	0.3906	0.3007	0.004480	10.85	0.0999	0.001364	8.989
		abs only	0.3885	0.3008	0	10.85	0.1000	**0.001348**	8.989
		LUT	0.3885	0.3008	0.000880	**10.85**	0.1000	0.001350	8.989
		SIM*	0.4587	0.2964	0.002480	11.06	0.1002	0.001755	8.989
		SDPM*	**0.3635**	0.3003	0.001920	11.00	**0.0849**	0.001380	**8.986**

Table 10.5: 1 day test instance performance measures for the best probabilistic and MIPSSM derived reserve crew schedules used in conjunction with each reserve policy

Table 10.5 shows the results for the 1 day test schedules. It shows that the SDPM policy leads to the lowest average cancellation measure and lowest probability of delay in each case. The SDPM policy is able to minimise the average cancellation measure by accepting a small increase in cancellation rate in order to achieve a significant reduction of delay. Compared to Figure 10.6 the results of Table 10.5 clearly demonstrate that the probabilistic model based approaches to reserve crew scheduling result in lower average cancellation measures compared to the MIPSSM based approaches, the reason for this is discussed in Section 10.6. The probabilistic approach also outperforms the MIPSSM based approach in each of the other performance measures. The reserve use rate statistics are not regarded as measures of performance, because it is possible to inefficiently use a high number of reserve crew. The reserve use rate statistics can however be used to infer the sorts of decisions made by policies which may have lead to better or worse performance compared to other policies. For test instance 1 the rate of using reserve crew to replace delayed crew is low, as expected. In the same test instance, the MIPSSM derived reserve crew schedule has a significantly higher rate of using reserve crew to replace delayed crew than that of the probabilistically derived reserve crew schedule. A possible reason for this is that the MIPSSM based approach explicitly models reserve crew being used to cover for delayed crew, and is therefore more likely to schedule reserve crew such that feasible combinations of reserve crew can be combined into teams to replace delayed connecting crew. On the other hand, the probabilistic approach worked best when assuming the absence only policy rather than the default policy during scheduling. The default policy always has the highest rate of using reserve crew to cover for delayed crew, which is the expected result.

Table 10.5 also shows that the abs only policy always minimises the cancellation rate for the 1 day test instances. This can be attributed to the fact that the abs only policy only uses reserve crew to cover for crew absence, the leading cause of cancellation. The SIM policy leads to the highest average cancellation measures, however it does minimise the maximum cancellation measure (worst case scenario) and average delay for the probabilistic reserve crew schedule of test instance 4. The poor performance of the SIM policy can be attributed to the low accuracy of the evaluations of alternative reserve decisions, which in turn can be attributed to the use of a finite number of repeat simulations for each evaluation. This means that the SIM policy has tractability issues that have to be overcome.

Test instance	Reserve schedule	Policy	Cancellation measure	Reserve use rate absence	Reserve use rate delay	Average delay	Probability of delay (> 15 mins)	Cancellation rate	Max CM
2	Probabilistic	default	0.2552	0.2867	0.003029	23.42	0.003744	0.0004406	**11.45**
		abs only	0.2551	0.2867	0	23.36	0.003763	**0.0004401**	**11.45**
		LUT	**0.2549**	0.2867	0.002057	23.38	0.003749	**0.0004401**	**11.45**
		SIM*	0.2987	0.2858	0.002457	24.18	0.003812	0.0005259	15.17
		SDPM	0.2567	0.2865	0.001157	**21.09**	**0.003586**	0.0004849	**11.45**
	MIPSSM	default	0.3035	0.2864	0.002000	30.16	0.003479	0.0004958	11.89
		abs only	0.3022	0.2864	0	30.06	0.003491	**0.0004929**	11.89
		LUT	0.3022	0.2864	0.001314	30.13	0.003483	**0.0004929**	11.89
		SIM*	0.3503	0.2858	0.001600	30.58	0.003510	0.0005958	15.05
		SDPM	**0.2906**	0.2861	0.001014	**25.80**	**0.003093**	0.0005377	**11.81**
5	Probabilistic	default	0.4898	0.2687	0.1847	8.000	0.09093	0.0005792	12.09
		abs only	0.2481	0.2733	0	**7.416**	0.09139	**0.0001981**	11.18
		LUT	0.3516	0.2708	0.1384	7.556	0.09072	0.0004000	12.02
		SIM*	0.4753	0.2682	0.1197	7.620	0.09152	0.0006703	12.10
		SDPM*	**0.2341**	0.2729	0.0527	7.599	**0.07472**	0.0002340	**11.17**
	MIPSSM	default	0.5196	0.2678	0.1844	7.963	0.09104	0.0006623	11.15
		abs only	**0.2619**	0.2733	0	7.571	0.09211	**0.0001995**	11.15
		LUT	0.3609	0.2713	0.1344	7.728	0.09127	0.0003679	11.15
		SIM*	0.5069	0.2674	0.1223	7.794	0.09221	0.0006995	11.15
		SDPM	0.2910	0.2729	0.0602	**7.570**	**0.07496**	0.0003429	**11.14**

Table 10.6: 3 day test instance performance measures for the best probabilistic and MIPSSM derived reserve crew schedules used in conjunction with each reserve policy

Table 10.6 shows that for the 3 day test instances, the probabilistic approach to reserve crew scheduling leads to lower average cancellation measures, average delays and probabilities of delay compared to the MIPSSM approach. However, the probabilistically derived reserve crew schedule only marginally outperforms the MIPSSM reserve crew schedule in test instance 5.

In terms of reserve policies Table 10.6 shows that the SDPM policy resulted in the lowest cancellation measure in two instances. The LUT policy minimised the average cancellation measure for the probabilistically derived reserve crew schedule of test instance 2 and the abs only policy minimised the average cancellation measure for the MIPSSM reserve crew schedule of test instance 5. The SDPM policy also minimised the probability of delay and the maximum cancellation measure for both 3 day test instances for each reserve crew schedule. However the maximum cancellation measures were very similar in all 3 day schedule cases except for the SIM policy in test instance 2. The average delay reductions associated with the SDPM policy applied to test instance 2 are greater than those of test instance 5. This can be explained by the fact that most of the delays in test instance 2 are reserve-induced delays and these types of delay are very sensitive to the reserve policy used.

The reserve use rates of Table 10.6 show that for the 3 day test instances, beside the abs only policy, the SDPM policy always had the lowest rate of using reserve crew to cover for delayed crew. Table 10.6 also shows that the SDPM approach of exploiting the trade-off between delays and cancellations and their effect on the average cancellation measure does not always pay off. In fact Table 10.6 gives evidence that sometimes the risk averse abs only policy can out perform more advanced policies. Additionally, the SDPM policy, which worked best with the probabilistically derived reserve crew schedule for test instance 5, was the more risk averse variant, where reserve crew holding is never considered in the event of crew absence (as indicated by the asterix).

Test instance	Reserve schedule	Policy	Cancellation measure	Reserve use rate absence	Reserve use rate delay	Average delay	Probability of delay (> 15 mins)	Cancellation rate	Max CM
3	Probabilistic	default	1.395	0.3186	3.500E-4	27.02	0.002636	0.001288	23.60
		abs only	**1.393**	0.3186	0	27.01	0.002637	**0.001286**	23.60
		LUT	**1.393**	0.3186	0	27.01	0.002637	**0.001286**	23.60
		SIM*	1.522	0.3165	1.000E-4	**26.97**	**0.002623**	0.001419	**23.01**
		SDPM*	1.458	0.3184	0	28.10	0.002627	0.001338	24.83
	MIPSSM	default	1.783	0.3254	0.001000	33.49	0.002814	0.001575	27.34
		abs only	1.778	0.3254	0	33.45	0.002815	**0.001570**	27.34
		LUT*	**1.751**	0.3199	0	28.70	0.002636	0.001616	**26.28**
		SIM	1.931	0.3236	0.000125	**28.11**	**0.002364**	0.001818	27.32
		SDPM*	1.894	0.3250	0	33.44	0.003025	0.001667	27.34
6	Probabilistic	default	1.742	0.3007	0.1472	7.61	0.08929	0.001329	30.57
		abs only	**1.126**	0.3062	0	**7.54**	0.08982	**0.000724**	**26.67**
		LUT	1.239	0.3044	0.0333	7.54	0.08969	0.000836	**26.67**
		SIM*	1.710	0.3001	0.1075	7.59	0.08974	0.001299	30.59
		SDPM*	1.148	0.3059	0.0010	8.05	**0.07422**	0.000737	26.73
	MIPSSM	default	1.861	0.3020	0.1145	7.32	0.08861	0.001493	38.65
		abs only	**1.322**	0.3047	0	**7.30**	0.08917	**0.000955**	26.27
		LUT	1.411	0.3040	0.0323	7.30	0.08905	0.001044	26.27
		SIM*	1.918	0.2992	0.0894	7.37	0.08905	0.001531	27.21
		SDPM	1.450	0.2988	0.0142	7.57	**0.07355**	0.001119	**24.21**

Table 10.7: 7 day test instance performance measures for the best probabilistic and MIPSSM derived reserve crew schedules used in conjunction with each reserve policy

Table 10.7 shows, just as Tables 10.5 and 10.6 do, that the reserve crew schedules derived using a probabilistic approach lead to average cancellation measures lower than those derived from a MIPSSM based approach. For the seven day test instances the abs only policy leads to the lowest average cancellation measures in three out of four cases. The LUT policy attains the lowest average cancellation measure for the MIPSSM derived reserve crew schedule of test instance 3. Table 10.7 also shows that the benefit of using the SDPM is much reduced compared to the cases of the one and three day test instances. An explanation for this is that, for the seven day schedules, the penalty for holding reserve crew in the event of crew absence is much higher, because crew pairings are longer (up to five days) in these schedules. So as a result, the potential benefit of using reserve crew to cover for delayed crew will typically also be reduced. It is very likely that the SDPM policy made some reserve holding decisions in the event of crew absence that in hindsight turned out to be very costly. The evidence for this is that the SDPM policy has a low rate of reserve crew used to cover for delayed crew, average rates of using reserve crew to cover for absent crew, but an elevated cancellation rate, which is the reason for the elevated average cancellation measures. Table 10.7 leads to the conclusion that risk averse reserve policies (such as the abs only policy) are recommended for long schedules where risky decisions have potentially very large penalties if things go wrong.

10.6 Results summary

In this section the overall meaning of the results in Tables 10.5, 10.6 and 10.7 is discussed. The results of Section 10.5 consistently showed that the probabilistic approach outperforms the MIPSSM based approach in all cases. The reason for this can be attributed to the fact that the MIPSSM based approach is limited by the number input scenarios that can be solved within a reasonable amount of time. This in turn means that the MIPSSM has an incomplete picture of reserve crew demand in all situations. So in fact, it is perhaps surprising how good the quality of the MIPSSM derived reserve crew schedules actually is. On the other hand, the probabilistic approaches use a continuous model of uncertainty and therefore do not suffer the same limitation as the MIPSSM approach. The surprisingly high quality of the MIPSSM derived solutions, given its limitations, may also be taken as indicating that the probabilistic approach has its own limitations, namely that a continuous model is unrepresentative of any individual possible outcome. In all, it may be possible that if the tractability of the MIPSSM approach, in terms of the number of input scenarios that can be solved simultaneously, can be addressed, the MIPSSM might have greater potential than a probabilistic approach. Another possibility to consider is that the best approach might involve a hybridisation of both distinct approaches, which is an interesting further research question.

Another general theme of the results given in Section 10.5 is the varying performance level of different reserve policies in different situations. In general it was found that the benefit of using the SDPM policy was greatest when the maximum penalty of making risky decisions, such as holding reserve crew in the event of crew absence, was low and when the potential reward of using reserve crew to cover delayed crew was high. Which was particularly the case for the 1 day schedule with a high delay risk (test instance 4). When there was a low risk of delay or when the schedule was long and the potential penalty of not covering for absent crew was highest, the benefits of the SDPM policy were reduced or removed entirely. In such cases the risk averse rule of thumb policy, abs only, gave comparable, and in some cases, better results. The SIM policy proved to be an unreliable evaluator of alternative reserve decisions because it requires a very large number of repeats to accurately evaluate the alternative decisions. The LUT policy was found to be a reliable policy, except for when it was applied to the longer test instances where there was also a high risk of delay. The main result in terms of policies is that in nearly all cases alternative reserve policies were found that outperformed the rule of thumb policies.

Based on these results, an airline considering using this research would be advised to use a probabilistic approach to reserve crew scheduling, unless they can solve the MIPSSM for vastly more scenarios than that considered in this research. In terms of reserve policies, a risk averse policy such as the abs only policy would be advised if there are few reserve crew available for a relatively long period of time. Otherwise the SDPM policy is recommended.

Chapter 11

Conclusion

In this thesis the problem of airline reserve crew scheduling under uncertainty has been tackled. This thesis introduced a general problem definition, for the problem of scheduling reserve crew and using reserve crew to mitigate unforeseen crew-related disruptions that occur on the day of operation. In contrast to previous work on reserve crew scheduling, this thesis focusses on the *fine detail modelling of the uncertainty of day of operation crew-related disruptions*. *Probabilistic* and *scenario* based approaches to modelling the uncertainty of crew-related disruptions which can be absorbed by using reserve crew were introduced. As a secondary objective, *consideration was given to the policy for reserve use on the day of operation*, as well as to how this can be improved and how the policy can be taken into account when scheduling reserve crew.

Chapter summary

Section 11.1 summarises the main findings from the work on the probabilistic models. Section 11.2 does the same for the work on the scenario-based approach. Section 11.4 summarises the main findings from the consideration of online reserve policies in this thesis. Section 11.5 lists the general insights gained from this research. Section 11.6 contains advice to KLM (or other airline's that operate in similar ways) on the most promising parts of this research that they could exploit. Section 11.7 summarises the main points from this conclusion chapter.

11.1 Probabilistic reserve crew scheduling

The first approaches considered were the probabilistic models of crew absence disruptions ($SPCAM$ of Chapter 5) and crew-related delay disruptions (CDM of Chapter 7). The $SPCAM$ was developed further (Chapter 6, resulting in the CAM and SDMs. Finally the $SDPM$ (Chapter 8) was developed, which *integrated the crew absence and delay models* within a probabilistic framework. The $SDPM$ yields cancellation and delay predictions which agree closely with those derived from repeat simulations and *attained some positive results when applied to the problems of reserve crew scheduling and online reserve crew decision making.*

The probabilistic models for crew absence disruptions (developed over Chapters 5 and 6) started from a set of simplifying assumptions which ignored many of the details which were judged to be obscuring the real underlying problem, i.e. an initial abstraction. The *initial model provided the basic modelling principle* on which all subsequent probabilistic models were built. This basic modelling principle included: the procedure used to calculate the effect a given reserve crew schedule has on absorbing crew-related disruptions, a procedure which meant that an *assumed reserve use order policy could be taken into account during the scheduling phase*; and the basic form of the objective function used when using the model to schedule reserve crew, that of minimising the expected number of cancellations. In Chapter 6 the $SPCAM$ was modified (to the CAM) to include the problem specific details which were ignored in Chapter 5. This meant modelling the possibility of different numbers of crew of different rank and qualification combinations being absent simultaneously from different crew pairings. This step significantly increased the complexity of the model and as a result made the model a significant computational bottleneck in any search methodology in which it was used. Despite this, no attempt was made to decrease the level of detail, or introduce the use of approximations into the model, the aim was always to make the model as accurate as possible. The underlying hypothesis was that: *a more accurate model used in conjunction with a simple search methodology gives higher quality solutions than an approximate model used in conjunction with an advanced (evaluation intensive) search methodology.* Evidence for and against this hypothesis was reported in Chapter 10, which indicated that for given a computational budget there is a *trade-off between model complexity and search methodology complexity in terms of the quality of the resultant solutions*. The exact topology of this trade-off is relatively unexplored. A scientific investigation of such a trade-off requires a quantification of the complexity of a model, on the other hand many search methodologies have well defined complexities (as assessed by the number of evaluations required for a given problem size). Furthermore, consideration should also be given to labour costs that are required for building more accurate models and the cost of increasing the raw CPU power that is used to solve the models.

Before the inclusion of aircraft fleet types, crew ranks and qualifications, the assumed reserve policy was the default policy of using reserve crew in earliest start time order. After the inclusion of aircraft fleet types, crew ranks and qualifications this policy was no longer adequate. The reason being that reserve crew of different types with the same start times may be feasible to cover the same disruption. The solution was a *generalised (default) reserve policy* (GRP) which orders reserve crew "on the fly" using a weighted sum of a number of different criteria, including: earliest start time; a measure of reluctance to using reserve crew who would have to fly below their assigned rank and the expected future demand of individual reserve crew. Experiments showed that *the GRP weights used in reserve crew scheduling and online each have a significant influence on the quality of the resultant reserve crew schedules.*

During the development of the CAM, the goal was to attain a model which gave cancellation rate predictions (associated with any given reserve crew schedule which it might be evaluating) which match those derived from simulation testing. It was found in Section 6.1.7 that the CAM fails to account for worst case crew absence scenarios, because it implicitly assumes that the total number of absent crew is always exactly the expected number, the result being an underestimation of cancellation rates. To circumvent the problem, *separate CAMs, corresponding to different points on the distribution of total absent crew, were evaluated simultaneously*. The overall objective value was then a weighted sum of the objective values from each of these evaluations of the CAM, with weights derived from the binomial distribution. Such an approach could also be applied in other problem domains where many events have associated probabilities and overall worst case analysis is of particular interest, for example portfolio optimisation.

The SDM accounted for (in a simplistic way) delays introduced when waiting for reserve crew to begin their standby duties before they can be used to cover for absent crew. For this purpose the *delay cancellation measure was developed* (Section 3.5.1) to map delays to a measure of cancellation. The delay cancellation measure function requires a subjective input parameter to determine the relationship between the equality of the level of disruption caused by delays of different sizes with respect to a cancellation. The delay exponent parameter of the delay cancellation measure function provides a means of pinpointing a Pareto optimal reserve crew schedule for a trade-off between cancellation and delay minimisation. In general, using a higher delay exponent pushes the reserve crew schedule towards cancellation minimisation and away from delay minimisation (see Section 3.5.1).

The probabilistic crew delay model of Chapter 7 was an analogous application of the approach developed in Chapter 5, adapted and applied to delay disruptions. It modelled the occurrence of crew-related delays (those caused by aircraft waiting for delayed crew on connecting flights) and how those delays can propagate in the form of subsequent crew-related delays. The approach was shown to be effective in schedules where there exists a high probability of delay propagation and a high rate of mid shift crew aircraft changes. A limitation of the CDM was that it only accounted for crew-related delays, which are, in reality, quite rare. An improved model based on this approach should allow for all types of delay, for example the improved model should allow for delays which propagate not only in the form of crew-related delays but propagated delays in general. Another limitation of the CDM is that it relies on a simulation learning phase, this means that if the approach is to be applied in an online context to evaluate alternative reserve decisions, it turns out to be just as efficient to just use simulation directly to evaluate the alternatives.

The $SDPM$ of Chapter 8 *addressed these limitations*, as it is a full theoretical model of delay propagation in general. The $SDPM$ is based on the idea of a delay propagation cycle and the idea that it is possible to calculate departure time distributions from arrival time distributions from earlier flights. This in itself is not an original idea, the main contribution of the $SDPM$ was applying this approach to a specific problem and how the associated problem specific events can be modelled in such a framework. The implementation of the *SDPM* uses three-dimensional matrices to store arrival time distributions for all airline resources as a result of previous flights at a given time. These ETA matrices are used to sequentially calculate the departure time distributions for all flights in a schedule which are then used along with journey time distributions to update the ETA matrices ready for subsequent departure distribution calculations. The three dimensional structure of ETA matrices *allows for the modelling of swap recovery actions and their effect on departure time uncertainty.* The *SDPM* provides *accurate delay predictions* as compared with the predictions derived from repeat simulations. Additionally, it was found in experiments that the fixed time interval size used by the *SDPM* influences both prediction accuracy and, as a result, the quality of reserve crew schedules. In general larger interval sizes result in decreased prediction accuracy. However, for a fixed computational budget, it was *beneficial to find a trade-off interval size* that resulted in the highest reserve crew schedule quality possible, because small interval sizes vastly increased the time required for reserve crew schedule evaluations. Future work could involve an investigation of the use of variable step lengths, which ensure that frequently occurring event times correspond to interval mid points.

The *SDPM* required input probabilities that each crew pairing has a full complement of crew, this information was provided by the *CAM*. The *SDPM* was then able to evaluate the expected levels of delay and cancellation for each flight in a schedule as a function of a given airline schedule, crew absence and journey time uncertainty, and a given combination of a reserve crew schedule and an assumed reserve policy. The *SDPM* was validated in terms of prediction accuracy, applied as an online reserve policy and used to schedule reserve crew. When the *SDPM* was applied to the problem of reserve crew scheduling, it was found that the average cancellation measures were roughly the same as those of reserve crew schedules derived from the *SDM* and the *CDM*. However, the reserve crew schedules derived from *the SDPM yielded qualitatively different reserve crew schedules*, especially for schedules with a high risk of delay: they reduced average delay, but at the cost of a small increase in cancellation rate. The explanation was that the *SDPM* gives a higher weight to delay minimisation in the objective function, because it accounts for all delays and not just reserve crew induced delays (as in the *SDM*), or crew-related delays (as in the *CDM*).

11.2 Scenario-based reserve crew scheduling

The alternative to the probabilistic approach considered in this thesis was the scenario-based MIPSSM. Instead of using probabilities to model the uncertainty of the occurrence of disruptions, the occurrence of disruptions was captured in the form of a set of actual disruption scenarios. The potential advantage of such an approach was that the scenario-based approach is mathematically simpler in terms of modelling how disruptions can be absorbed by using reserve crew, as scenarios correspond to actual sequences of outcomes. As a result of this, a scenario-based approach may be better able to capture the correlations that exists between disrupted flights. On the other hand, the probabilistic approach is based on a single scenario where all events have some probability of occurring, which does not correspond to any possible outcome. Which may have the effect of limiting how accurately the correlations that exist between disrupted flights.

Chapter 9 defines a framework for simulation scenario collection and a mixed integer programming formulation for finding a reserve crew schedule which *minimises the occurrence of delay and cancellation disruptions* that occur in a given set of scenarios. Additionally, an extended formulation is given for the reserve crew scheduling problem for the case where there are multiple fleet types and multiple crew ranks and qualifications. It found that large solution times prevented the MIPSSM from being solved with more than 50 scenarios (for the smallest problems considered). A result of this was that alternative objective functions which were aimed at minimising the worst case scenario did not perform well, as the worst case scenario in a limited sample of scenario is unlikely to be representative of a truly worst case scenario. The most effective objective function corresponded to minimising the average level of delay and cancellation disruptions (cancellation measure objective).

It was also found that *the particular set of scenarios included in the MIPSSM had a significant impact on the quality of the resultant reserve schedules*. This led to the introduction of a *scenario selection heuristic* (SSH) which gave improved solution quality using fewer input scenarios. An investigation into the effect of the types of scenarios included in a set of scenarios revealed that scenarios can be classified according to: 1) how well they complement an existing set of disruption scenarios or; 2) whether they are scenarios which lead to good solution quality when used as the only input scenario for the MIPSSM. Algorithms based on these ideas resulted in improved solution reliability for a small example problem. The main limitation for the MIPSSM approach is the large solution times that occur for a seemingly small set of input scenarios, due to the large number of binary variables in the resultant models. Given more time, the scenario-based approach could be explored in more detail to reveal its true potential, to do this solution approaches such as Dantzig-Wolfe decomposition could be applied, and a revised formulation of the MIPSSM might improve the situation.

11.3 Comparison of the probabilistic and scenario-based approaches

Chapter 10 *compared the probabilistic and MIPSSM based approaches* to reserve crew scheduling. It was shown that in comparison to other simpler heuristic approaches to reserve crew scheduling such the area under the graph and uniform start rate approaches, the probabilistic and MIPSSM approaches give reserve crew schedules of higher and to each other comparable quality. However, on closer inspection it was found that the *probabilistic approach almost always outperformed the MIPSSM based approach*. The reason for this, was because the MIPSSM based approach cannot be solved efficiently for a large number of input disruption scenarios. The probabilistic approach does not suffer the same problem because it is able to model all possible outcomes in what is effectively a single probabilistic scenario. It is still possible that with additional work a scenario-based approach or a hybrid of the MIPSSM and probabilistic approaches might have a greater potential than a purely probabilistic approach.

11.4 Online reserve policies

As depicted in Figure 1.1 online reserve policies formed a secondary thread of this thesis, which is interwoven with the work on reserve crew scheduling. In Chapter 5 the probabilistic crew absence model provided *a means of modelling an assumed reserve use order policy*. At that point a default policy was assumed, which for the simplified model was reasoned to be the optimal policy. In Chapter 6 the CAM required a revised default policy in the form of the GRP, which was investigated in terms of the effect of the policy weights used offline and online on the expected cancellation measure. It was found that *the policy parameters assumed both offline and online had a significant impact on the quality of the reserve crew schedules derived* from the CAM. In Chapter 8, another aspect of reserve policies was *investigated, that of the possibility of reserve holding* when reserve crew are the best recovery for a given disruption. The $SDPM$ model was applied to evaluate alternative decisions in such circumstances. It was found that such an approach has *a clear potential for minimising overall disruptions at the expense of short term penalties.*

The MIPSSM approach implicitly assumes an optimal policy because it has access to knowledge of future disruptions, which in an online context would not be available. Despite this, in Section 9.5 a look-up table (LUT) reserve policy was derived from the MIPSSM corresponding to a given reserve crew schedule. Overall disruptions were reduced by this approach, as demonstrated in Section 9.6. When the MIPSSM was extended to the case of multiple fleet types, crew ranks and qualifications, the previously described policy did not also extend. A more appropriate *LUT structure was proposed* in Section 4.7.3 for a LUT policy based on approximate dynamic programming. A possible area for future research is to incorporate approaches such as the MIPSSM in simulation based learning approaches for finding reserve policies.

The simulation introduced in Chapter 4 suggested two reserve policies, a LUT approach (referred to above) and a direct application of simulation to evaluate alternative reserve decisions. The LUT approach was based on approximate dynamic programming, in which the values of all states are learned and used to select optimal decisions by directing the system towards the state with the best value. A policy is optimal if the values of states are based on following an optimal policy thereafter, which makes the problem of policy learning a circular one. In section 4.7.3 the values of states are learned from a reliable risk averse rule of thumb policy which approximates the behaviour of the optimal policy. No exploratory learning phase was attempted for the LUT policy, this represents possible future work.

In Chapter 10 all of the considered reserve policies were compared over a range of realistic test schedules used in conjunction with the best reserve crew schedules from a variety of approaches. This comparison revealed the strengths and weaknesses of the considered reserve policies. It was found that the reserve policies based on evaluating alternative reserve decisions have a *clear potential for reducing overall disruptions* in comparison to rule of thumb policies, but can run into trouble in certain circumstances. For instance, *the SDPM policy can sometimes be caught out when basing decisions on expected future outcomes which do not materialise.* The direct simulation policy was found to be unreliable because evaluation accuracy is related to the number of repeat simulations used to evaluate alternative decisions. The risk averse absence only policy (Section 3.5.2) *provided a good benchmark performance level* for the more advanced approaches to aim for, and in the longer schedules appeared to be a good approximation of an optimal policy. The reason for this was that the penalty of not covering crew absence disruptions is much higher when crew pairings are longer. This meant that it was rare for the risky decisions, such as holding reserve crew in the event of crew absence and using reserve crew to cover for delayed crew, to be the globally optimal decisions. In these cases the maximum penalty associated with risky decisions (as opposed to following the risk averse abs only policy) need to be accounted for.

11.5 General insights gained

The general insights gained from this research were as follows:

- A trade-off exists between model fidelity and search algorithm complexity with respect to the quality of the reserve crew schedule that can be derived using a fixed computational budget.

- Approaches developed for offline reserve crew scheduling extend easily to online applications as reserve policies.

- Fully detailed models of delay propagation are most useful in an online context, as these models are able to exploit information on the outcomes of previous events (events such as arrival times and crew absence) as they occur. In an offline perspective similar results can be achieved with a much less detailed model than the SDPM.

- Meta heuristics such as simulated annealing and genetic algorithms can find good quality solutions from the probabilistic models, even when the available solution time is limited.

- The scenario-based approach became intractable for relatively few input disruption scenarios, but given careful selection of the input disruption scenarios, reserve crew schedules of surprisingly high quality can still be obtained.

11.6 Advice for KLM

Based on the findings from this research KLM may find the following useful:

- When scheduling reserve crew, exploit all available information about the structure of the airline schedule:
  - The aircraft routings and crew pairings determine how potentially damaging any given delay is and which subsequent flights will also be delayed.
  - The structures of crew pairings determine how potentially damaging crew absence disruptions can be if the absent crew are not replaced.
- The cancellation measure function (Section 3.5.1) allows for scheduling and recovery decisions which are based on finding a trade-off between cancellation minimisation and delay minimisation.
- If a mathematical model is used to evaluate a set of alternative reserve crew schedules it should be noted that the assumed recovery policy, including the reserve policy, influences the effectiveness of any given reserve crew schedule. So an investigation of possible recovery/reserve policies is required to determine the best policy to assume whilst scheduling reserve crew.
- Reserve order policies (such as the GRP, see Section 6.4) and reserve holding policies (such as the online application of the SDPM, see Section 8.2.1) can identify beneficial reserve crew based recovery actions which may not be locally optimal but are beneficial in the long run.
- Approaches such as the probabilistic models and the scenario-based approaches can be applied to investigate manpower planning in terms of the number of reserve crew required based on them being scheduled and used in an efficient way.

11.7 Thesis summary

In conclusion, this thesis has shown the potential of two distinct approaches to modelling reserve crew demand uncertainty and scheduling reserve crew. It has been shown for the test instances considered that the reserve crew schedule can have a big impact on the expected level of day of operation disruptions and that improving the reserve crew schedule can be very beneficial. It has also been shown that the reserve policy used on the day of operation similarly influences the expected level of disruptions on the day of operation, and that reserve decisions that may not be immediately beneficial can be beneficial in the long run. Several approaches were developed which successfully demonstrated this.

Chapter 12

Potential future extensions

A number of potential future extensions based on the work presented in this thesis are described in this chapter.

12.1 Reserve crew scheduling

This thesis has explored two distinct approaches to modelling reserve crew demand uncertainty, which are then used to schedule reserve crew. The MIPSSM scenario-based approach is an integer programming approach which uses simulation derived disruption scenarios as the input, whereas the probabilistic approach is deterministic, because it does not require stochastically generated inputs. Future work based on these approaches is described below.

12.1.1 Probabilistic based approaches

One of the main aims during the development of the probabilistic approaches to reserve crew scheduling (Chapter 5 to 8) was to make the model as detailed as possible. As a result of this, model evaluation became a significant computational bottleneck in the scheduling and policy applications. Future work could be to look into how different aspects of the model can be approximated in some way without a significant loss of accuracy or a drop in the quality of reserve crew schedules or reserve policy performance. For example, the CAM enumerates all feasible combinations of reserve crew for each possible disruption (Section 6.1.5). This step could be approximated by limiting the enumeration to, for example, the first 0.99 of the cumulative probability of the reserve combinations most likely to be used. The remaining combinations will have vanishingly small associated probabilities and may not have a noticeable effect on the reserve crew schedule quality or reserve policy performance. If the probabilistic models can be evaluated quicker without a significant loss in accuracy, then more advanced (evaluation intensive) methodologies can be used to find better reserve crew schedules. If the level of approximation/model accuracy can be parameterised then it would be possible to investigate the topology of the trade-off between modelling accuracy and solution methodology complexity for a given computational budget (maximum solution time).

12.1.2 Scenario-based approaches

One of the issues with the approaches to reserve crew scheduling which are based on the $MIPSSM$ is that the number of scenarios has a big impact on the time required to solve the resultant $MIPSSM$ formulation. Future work could be to develop specialised solution techniques other than solving the model directly in CPLEX. One possible alternative is to develop a hybrid meta-heuristic and integer programming approach where the meta-heuristic is used to search for a subset of the variables, which are then fixed for an iteration of the $MIPSSM$, whilst the others are optimised.

Another approach might involve further improvements of the SSH where scenarios are not only added but can also be removed, for example the best case scenario could be removed and replaced with a new worst case scenario.

Another possible improvement might involve an iterative solution approach of the $MIPSSM$ formulation where the reserve schedule variables (x) and the reserve use variables (y) are alternately held fixed, this would greatly reduce the number of integer variables in each iteration, the desired outcome is that the solution converges to the optimal solution of the full problem.

Another possible area for future research is in the use of the $MIPSSM$ formulation to aid recovery decisions in an online context. The solution time is small when considering a single disruption scenario with a fixed reserve crew schedule. This could be exploited to evaluate alternative reserve crew recovery decisions, by solving the $MIPSSM$ for each of a large sample of possible future disruption scenarios for each alternative recovery decision. Such an approach would require an airline to have the facility to run simulations of future events based on the current schedule and the expected departure and arrival times for all current flights.

12.1.3 Hybridised approaches

In Section 9.7.3 two reserve crew scheduling algorithms were proposed, which were based on the idea of utilising the best features of the scenario and probabilistic approaches within a single algorithm. I.e. using the scenario-based approach to generate reserve crew schedules which are then evaluated using a probabilistic approach. Another possible approach might involve using a probabilistic model for modelling the crew absence part of the problem and the MIPSSM for the delay part of the problem. The reserve crew available for scheduling could be divided between the the two models. The MIPSSM could schedule a number of reserve crew in anticipation of delayed crew whilst the probabilistic model schedules the remaining crew in anticipation of the expected crew absences. On the day of operation reserve crew need not necessarily be restricted to the types of disruptions they were scheduled for.

Another possible hybridisation of the probabilistic and scenario-based approaches would be to somehow formulate the MIPSSM model with the integer requirement of certain decision variables relaxed so that the solutions or inputs correspond to probabilities. For example a set of disruption scenarios could be modelled as a single probabilistic scenario. Then, the reserve use decision variables integer requirement could be relaxed, but the reserve crew schedule variables would remain integer. This approach would vastly decrease solution times due to a vast decrease in the number of integer decision variables and hence much reduced branching. However, a likely pitfall of such an approach is that the probabilistic model will not be linear, which would preclude the use of a linear programming solver. The reason for the non-linearity is that the probabilities that reserve crew are used for different possible disruptions depend on the probabilities that other reserve crew are not available for the same disruptions, and the calculations for this involve multiplying probabilities together, see Equation 5.1. The model would, at most, be non-linear to a degree equal to the number of reserve crew being scheduled. This occurs when all reserve crew standby duties overlap with each other for a number of flights, in these cases the probability that the last reserve crew is used depends on the probabilities that all other reserve crew are not available.

12.1.4 Extended formulations

The current work is based on a single hub model, the single hub model accounts only for disruptions that occur at the hub, whilst assuming that disruptions that occur at spoke stations are dealt with there. Therefore, to extend the current model to the case of multiple hubs, one option would be to solve the model separately from each hub's perspective. However, if there are often frequent flights between hubs, a partially integrated multiple hub model may be more appropriate. Another alternative would be to model the schedules for the hubs as a single combined schedule. In this approach reserve crew could be scheduled in the same way as described in this work, provided that the additional spatial constraints for reserve crew use are respected.

This thesis has focussed on reserve crew demands which occur on the day of operation and which only become known close to the scheduled departure time of the disrupted flight. These include crew absence and delay disruptions. The approaches developed during this thesis could also be extended to account for other types of crew related disruptions, such as delayed crew who become infeasible for their next flight because they are expected to exceed maximum working hour constraints. These types of disruptions are the same as delays but they can also be considered to have a cancellation threshold which depends on the particular crew assigned to a flight. To extend the current work to account for disruptions due to delayed infeasible crew the cancellation threshold (CT) has to be updated to be function of the particular crew assigned to the given flight. In this thesis CT was modelled as a constant, but incorporating a variable CT in both the probabilistic and scenario approaches would be relatively straight forward. Appendix H considers a variable cancellation threshold formulation.

As described above, this thesis has focussed on scheduling reserve crew at the hub of a single hub network, whilst assuming that disruptions that occur at spoke stations are dealt with there. The justification for this assumption is that hubs are by their nature heavily congested and therefore the minimisation of disruptions at the hub is of vital importance. Disruptions that occur at spokes on the other hand are both less frequent and more difficult to deal with using recovery actions available at the hub. For example, if crew absence occurs at a spoke station and the disruption can not be covered by reserve crew stationed at that spoke station, reserve crew would have to deadheaded to that spoke from the hub station, but since flights from the hub to any given station are relatively infrequent it is unlikely that reserve crew will be deadheaded in time to avoid the cancellation of the disrupted flight. However, deadheading is a viable option for disruptions that can be anticipated in enough time, such as a crew absence disruption with plenty of notice, so that a flight is available to the spoke station on which to deadhead the required replacement crew. Future work could include allowing for deadheading in the approaches proposed in this thesis. Doing so would also pave the way for a possible application of the proposed approaches to reserve crew scheduling in point to point network structures. However, it is very unlikely that airlines that operate point to point networks will schedule standby reserve crew at stations in their network, because there will be a very low utilisation rate. Instead, temporary or agency replacement crew provide a viable alternative.

12.1.5 Integrated crew scheduling and reserve crew scheduling

During this thesis it has been demonstrated that exploiting the structure of the crew schedule is very important when scheduling reserve crew, because the structure of the crew schedule determines which flights are disrupted when different crew are disrupted. Therefore, future work could be to schedule crew and reserve crew in an integrated manner. One possibility is to use an iterative approach (analogous to that of Weide's, see Section 2.2.3) which alternates between optimising the reserve crew schedule for a fixed crew schedule and optimising the crew schedule for a fixed reserve crew schedule. Such an approach could use the $SDPM$ of Chapter 8 to evaluate the robustness of a given combination of a crew schedule and a reserve crew schedule. The solution space of such an approach is enormous, but the solution space could be limited to allowing only small changes to an initially cost optimal crew schedule.

12.1.6 Other applications

This thesis proposes several approaches to scheduling reserve crew under uncertainty. These approaches could also be extended to apply to other problem domains. One possible application is in public transport, such as scheduling reserve coach drivers or train crew. The public transport systems of large cities could provide an ideal application for some of the proposed approaches, because the central stations are analogous to the hub stations of hub and spoke networks.

Another potential application for the proposed approaches is in portfolio optimisation. In particular, the *CAM* of Chapter 6 could possibly be applied to provide a probabilistic model of investment risk. The model refinement of Section 6.1.7 helped to increase the accuracy of the cancellation rate predictions by explicitly modelling a range of crew absence scenarios including that where many crew are absent and the available reserve crew cannot possibly cover for all of them. The model was able to quantitatively model the best case, average case and worst case scenarios simultaneously as opposed to just the average case scenario.

This thesis has considered reserve crew scheduling and day of operation reserve policies, another application of the proposed approaches is in manpower planning. This was briefly considered in Section 9.6.2 where reserve crew scheduling models were solved for different numbers of available reserve crew, to determine an appropriate number of reserve crew to schedule.

12.2 Reserve Policies

The main focus in this thesis was on modelling reserve crew demand uncertainty and scheduling reserve crew. The study of reserve policies provides another approach to minimising day of operation disruptions. The problem of optimising reserve crew use on the day of operation can be cast as a multistage decision making problem. Such problems can be tackled using the family of techniques referred to as approximate dynamic programming (see Section 2.7.4).

12.2.1 Approximate dynamic programming

In Section 4.7.3 a look-up table (LUT) reserve policy was described. A state in the LUT corresponded to the number of reserve crew remaining of each crew rank and qualification combination and the departure number. The value of a state was the expected cancellation measure contribution that would be accumulated after reaching that state. The LUT policy of Section 4.7.3 was designed as a policy that was to be applied in instances when reserve crew could be used to cover for delayed crew in order to determine whether such an action was appropriate. The LUT table values were learned from repeat simulations in which the absence only reserve policy was used. The reasoning was that if LUT values suggest that using reserve crew to cover for delayed crew is beneficial even though the policy values were derived using the absence only policy, then using reserve crew to cover for delayed crew cannot be a very risky action and is likely to be the best recovery action in terms of overall disruption minimisation.

In approximate dynamic programming the values of states have to correspond to those that would occur when following an optimal policy. The absence only policy only approximates the behaviour of the optimal policy. In future work the techniques of approximate dynamic programming could be applied to the reserve policy problem to try to find a truly optimal policy. Such a task is a worthy of a research project in itself. There are many hurdles to overcome when applying approximate dynamic programming. The main difficulties include: the very large state spaces of problems usually requires some kind of coarse approximation; and the problem of learning optimal state values requires knowledge of optimal state values, which makes it a circular learning problem. The MIPSSM implicitly assumes an optimal reserve policy, such an approach could be used in the learning phase or initialisation of an approximate dynamic programming approach. The probabilistic models could be used in a similar way.

12.2.2 Hybridised reserve policies

In Section 10.5 it was found that different reserve policies work best in different situations. Therefore, future work would be to devise a framework for selecting reserve policies that are to be used in each possible situation. Such a framework could be rule based. Another possibility is to derive a LUT policy for the values of using different policies in different states. In this context a state may refer to characteristics of an airline schedule such as its average delay risk or the worst case scenario associated with making certain risky decisions.

12.3 Integrated reserve crew scheduling and reserve policy optimisation

During this thesis it has been demonstrated that using knowledge of the reserve policy during reserve crew scheduling has a significant impact on the quality of the derived reserve crew schedules. Additionally, it has been shown that the reserve policy used on the day of operation can also significantly improve the performance of a reserve crew schedule. In Section 6.4.4 the interaction between the reserve order policy assumed offline and that used online was explored, and very little evidence was found to suggest a complex interaction between assumed offline policies and online policies. This conclusion is applicable to the GRP only. Future work could investigate integrated reserve crew scheduling and reserve policy optimisation for more complex reserve policies such as those alluded to in Section 12.2.

Another future work possibility for the approximation dynamic programming approach to reserve policies is to make the reserve crew schedule an element of the policy, thus integrating reserve crew scheduling and reserve policy optimisation. In such an integrated framework the policy learning phase would have to learn the optimal policy and reserve crew schedule simultaneously. One possibility is that the learning rates could fluctuate between both aspects of the overall policy to mimic an iterative approach to integrated reserve crew scheduling and reserve policy optimisation. It may also be possible to calculate or accurately approximate how the values of states change as the reserve schedule is slightly modified, for example, scheduling one reserve crew at a slightly different time will automatically change the potential for reserve crew induced delay of some of the disruptions they are feasible to cover.

Bibliography

[1] Ahmed Abdelghany, Goutham Ekollu, Ram Narasimhan, and Khaled Abdelghany. A proactive crew recovery decision support tool for commercial airlines during irregular operations. *Annals of Operations Research*, 127:pp309–331, 2004.

[2] Yanina V. Ageeva and John-Paul Clarke. Approaches to incorporating robustness into airline scheduling. Master's thesis, MIT, 2000.

[3] Shervin AhmadBeygi, Amy Cohn, and Yihan Guan. Analysis of the potential for delay propagation in passenger airline networks. *Journal of Air Transport Management*, 14(5):221–236, September 2008.

[4] Shervin AhmadBeygi, Amy Cohn, and Marcial Lapp. Decreasing airline delay propagation by re-allocating scheduled slack. Industry Studies Conference, 2008.

[5] Ranga Anbil, Eric Gelman, Bruce Patty, and Rajan Tanga. Recent advances in crew-pairing optimization at american airlines. *Interfaces*, (21):62–74, Jan/Feb 1991.

[6] Poornima Balakrishnan. *Approximate dynamic programming*. Verlag Dr Muller, 2010.

[7] Micheal Ball, Cynthia Barnhart, George Nemhauser, and Amedeo Odoni. Air transportation: Irregular operations and control, 2007.

[8] Micheal Ball and Anito Roberts. A graph partitioning approach to airline crew scheduling. *Transportation Science*, 19(2), May 1985.

[9] Joy O. Banks, Katrina E. Avers, Thomas E. Nesthus, and Erica L. Hauck. A comparative study of international flight attendant fatigue regulations and collective bargaining agreements. *Journal of Air Transport Management*, (19):21–24, 2012.

[10] Cynthia Barnhart, Peter Belobaba, and Amadeo R Odoni. Applications of operations research in the airline industry. *Transportation Science*, 2003.

[11] Cynthia Barnhart, Natasha L. Boland, Lloyd W. Clarke, Ellis L. Johson, George L. Nemhauser, and Rajesh G. Shenoi. Flight string models for aircraft fleeting and routing. *Transportation Science*, 32(3):208–220, August 1998.

[12] Cynthia Barnhart, Amy M. Cohn, Ellis L. Johnson, Diego Klabjan, George L. Nemhauser, and Pamela H. Vance. Airline crew scheduling. *Handbook of Transportation Science*, pages 517–560, 2003.

[13] Cynthia Barnhart, Ellis L. Johnson, George L. Nemhauser, Martin W. P. Savelbergh, and Pamela H. Vance. Branch-and-bound: Column generation for solving huge integer programs. *Operations Research*, 1998.

[14] Cynthia Barnhart, Ellis L. Johnson, George L. Nemhauser, Martin W.P. Savelsbegh, and Pamela H. Vance. Branch-and-price: Column generation for solving huge integer programs. *Operations research*, 46(3):316–329, May-June 1998.

[15] C. Bayliss, G. De Maere, J.A.D. Atkin, and M. Paelinck. Scheduling airline reserve crew to minimise crew related delay using simulated airline recovery and a probabilistic optimisation model. *in proceedings IEEE International Conference on Systems, Man and Cybernetics*, 2013.

[16] C. Bayliss, G. De Maere, J.A.D. Atkin, and M. Paelinck. A simulation scenario based mixed integer programming approach to airline reserve crew scheduling under uncertainty. In *Proceedings of the 10th International Conference of the Practice and Theory of Automated Timetabling*, pages p62–81, 2014.

[17] 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.

[18] Massoud Bazargan. *Airline operations and scheduling.* Ashgate, 2010.

[19] J.E. Beasley and B. Cao. A tree search algorithm for the crew scheduling problem. *European Journal of Operations Research*, (94):517–526, 1996.

[20] Richard Bellman. The theory of dynamic programming. 1954.

[21] Annabell Berger, Andreas Gebhardt, Matthias Mller-Hannemann, and Martin Ostrowski. Stochastic delay prediction in large train networks. In *11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, OpenAccess Series in Informatics (OASIcs), pages 100–111, 2011.

[22] Dmitris Bertsimas, David B. Brown, and Constantine Caramanis. Theory and applications of robust optimisation. *Society for Industrial and Applied Mathematics*, 2011.

[23] John R. Birge and Francois Louveaux. *Introduction to stochastic programming*. Springer, 1997.

[24] Alexandre Boissy. Crew reserve sizing. In *AGIFORS Symposium*, 2006.

[25] Edmund Burke and Graham Kendall, editors. *Search Methodologies*. Springer, 2005.

[26] Shaw-Ching Chang. A duty based approach in solving the aircrew recovery problem. *Journal of Air Transport Management*, 19:16–20, 2012.

[27] S. Cicerone, G DAngelo, G. Di Stefano, D Frigioni, and A Navarra. Robust algorithms and price of robustness in shunting problems. In *7th Workshop on Algorithmic Approaches for Transportation Modelling, Optimisation and Systems*.

[28] S. Cicerone, G. Di Stefano, M Schachtebeck, and A. Schobel. Dynamic algorithms for recoverable robustness problems. In *8th Workshop on Algorithmic Approaches for Transportation Modelling, Optimisation and Systems*.

[29] John Paul Clarke. Virtual spares. AGIFORS, 2008.

[30] John-Paul Clarke, Terran Melconian, Elizabeth Bly, and Fabio Rabbani. Means-mit extensible air network simulation. *Simulation*, 83, 2008.

[31] R. G. Coyle. *System dynamic modelling*. Chapman and Hall, 1996.

[32] Jeffrey E. Dillon and Spyros Kontogiorgis. Us airways optimises the scheduling of reserve flight crews. *Interfaces*, 29(5):pp123, September/October 1999.

[33] Viktor Duck, Lucian Ionescu, Natalia Kliewer, and Leena Suhl. Increasing stability of airline crew and aircraft schedules. *Transportation Research*, 2012.

[34] Micheal Dudley and Delano Clarke. Irregular airline operations: a review of the state-of-the-practice in airline operations control centers. *Journal of Air Transport Management*, 1998.

[35] Michelle Dunbar, Gary Froyland, and Cheng-Lung Wu. Robust airline scheduling planning: Minimizing propagated delay in an integrated routing and crewing framework. *Transportation Science, Articles in Advance*, pages pp1–13, February 2012.

[36] Martin Dyer and Leen Stougie. Computational complexity of stochastic programming. *Mathematical Programming, Series A*, 106:423–432, 2005.

[37] Dilwyn Edwards and Mike Hamson. *Guide 2 mathematical modelling.* Palgrave, second edition, 2001.

[38] Matteo Fischetti and Michele Monaci. Light robustness. In *Arrival-TR-0119.*

[39] Rainer Frick. Simulation of transportation networks. Technical report, V-Research BmbH, Austria.

[40] Adel Gaballa. Planning callout reserves for aircraft delays. *Interfaces*, (2), February 1979.

[41] Virginie Gabreal, Ceceile Murat, and Aurelie Thiele. Recent advances in robust optimisation and robusteness: An overview. July 2012.

[42] F Glover and M Laguna. *Tabu search.* Kluwer acedemic publishers, 1997.

[43] David E. Goldberg. *Genetic algorithms in search optimization and machine learning.* Addison-Wesley, 1989.

[44] Ram Gopalan and Kalyan T. Talluri. Mathematical models in airline schedule planning. *Annals Of Operations Research*, 1998.

[45] Tore Grunert and Hans-Jurgen Sebastian. A hybrid tabu search/branch and bound algorithm for the direct flight network design problem. *Transportation Science*, 34(4):364–380, November 2000.

[46] William E Hart, N Krasnogor, and J E Smith. *Recent advances in memetic algorithms.* Springer, 2005.

[47] Hastie and Tibshirani. Cross validation and bootstrap. In *http://www-stat.stanford.edu/ tibs/sta306b/cvwrong.pdf*, 2009.

[48] James J. Higgins and Sally Keller McNulty. *Concepts in probability and stochastic modelling.* Duxbury press, 1995.

[49] Wallace J. Hopp and Mark L. Spearman. *Factory physics.* McGraw Hill, 2000.

[50] Ahmad Hossny, Saeid Nahavandi, and Douglas Creighton. Minimizing impact of bounded uncertainty on mcnaughtons algorithm via interval programming. In *IEEE SMC 2013*, 2013.

[51] Peter Kall and Stein W. Wallace. *Stochastic programming.* John Wiley and Sons, 1994.

[52] Josef Kallrath and John M. Wilson. *Business optimisation using mathematical programming.* Macmillan, 1997.

[53] Mohammad H. Keyhani, Mathias Schnee, Karsten Weihe, and Hans-Peter Zorn. Reliability and delay distributions of train connections. In *12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, OpenAccess Series in Informatics (OASIcs), pages 35–46.

[54] S Kirkpatrick, C D. Gelatt, Jr, and M P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, May 1983.

[55] Diego Klabjan. Large-scale models in the airline industry. *Business and Economics Journal*, 2005.

[56] Diego Klabjan, Ellis Johnson, George L. Nemhauser, Eric Gelman, and Srini Ramaswamy. Airline crew scheduling with regularity. *Transportation Science*, 35(4):359–374, November 2001.

[57] KLM. Private communication.

[58] Niklas Kohl and Stefan E. Karish. Airline crew rostering: Problem types, modelling and optimization. *Annals of Operational Research*, 2004.

[59] Nilas Kohl, Allan Larsen, Jesper Larsen, Alex Ross, and Sergey Tiourine. Airline disruption management-perspectives, experiences and outlook. *Journal of Air Transport Management*, (13):149–162, 2007.

[60] Marcial Lapp and Florian Wilkenhauser. Incorporating aircraft efficiency measures into the tail assignment problem. *Journal of Air Transport Management*, (19):25–30, 2012.

[61] Ladislav Lettovsky, Ellis Johnson, and George L. Nemhauser. Airline crew recovery. *Journal of Air Transport Management*, 34(4):337–348, 2000.

[62] Wenkan Ken Li, Chikage Miyoshi, and Romano Pagliari. Dual hub-connectivity: An analysis of all nippon airways use of narita and haneda airports. *Journal of Air Transportation Management*, 23:12–16, 2012.

[63] Christian Liebchen, Marco Lubbecke, Rolf H. Mohring, and Sebastian Stiller. Recoverable robustness. Technical report, August 2007.

[64] Marco E. Lubbecke and Jacques Desrosiers. Selected topics in column generation. *Operations Research*, 2005.

[65] Terry Lucey. *Quantitative techniques*. Thomson, sixth edition, 2002.

[66] Cheng lung Wu. *Airline operations and delay management*. Ashgate, 2010.

[67] Geert De Maere. *Multi-objective approaches to investigate airline schedule robustness*. PhD thesis, University of Nottingham, 2010.

[68] Dennis F. X. Mathaisel. Decision support for airline system operations control and irregular operations. *Computers Ops Res*, 23(11):1083–1098, 1996.

[69] Claude P. Medard and Nidhi Sawney. Airline crew scheduling from planning to operations. *European Journal of Operational Research*, (183):1013–1027, 2007.

[70] Zbigniew Michalewics and David B. Fogel. *How to Solve it: Modern Heuristics*. Springer, 2004.

[71] N. Mladenovic and P Hansen. Variable neighborhood search. *Computers and Operations Research*, 24(11):1097–1100, 1997.

[72] Juliana M. Nascimento and Warren B. Powell. An optimal approximate dynamic programming algorithm for the dispatch problem with grid level storage. Technical report, Princeton University, 2009.

[73] Lars Kjaer Nielson, Leo Kroon, and Gabor Maroti. Absorption robustness of railway resource schedules. *ARRIVAL project*, 2007.

[74] Marc Paelinck. KLM cabin crew reserve duty optimisation. In *AGIFORS proceedings*, 2001.

[75] Marc Paelinck. Crew processes. *Personal communication KLM*, 2011.

[76] Jon D. Peterson, Gustaf Solveling, Ellis Johnson, John-Paul Clarke, and Sergey Shebalov. An optimisation approach to airline integrated recovery. *Transportation Science*, 2012.

[77] Micheal L. Pinedo. *Scheduling, tools and algorithms*. Springer, 2008.

[78] Joao P. Pita, Cynthia Barnhart, and Antonio P. Antunes. Integrated flight scheduling and fleet assignment under airport congestion. *Transportation Science, articles in advance*, pages pp1–16, November 2012.

[79] Warren B. Powell. *Approximate dynamic programming*. Wiley, 2007.

[80] Warren B. Powell. What you should know about dynamic programming. *Wiley InterScience*, 2009.

[81] Warren B. Powell, Hugo P. Simao, and Belgacem Bouzaiene-Ayari. Approximate dynamic programming in transportation and logistics: A unified framework. Technical report, Princeton University, 2012.

[82] Fabio Faizi Rahnemay Rabbani. *Implementation of an airline recovery model in an event based simulation*. PhD thesis, MIT, 2004.

[83] Colin R. Reeves, editor. *Modern heuristic techniques for combinatorial problems*. Orient longman, 1992.

[84] Micheal Romer and Taieb mellouli. Pairing optimisation for hierarchical crews. 2012.

[85] Jay M. Rosenberger, Ellis L. Johnson, and George L. Nemhauser. Rerouting aircraft for airline recovery. *Transportation Science*, 37(4):pp408–421, November 2003.

[86] Jay M. Rosenberger, Ellis L. Johnson, and George L. Nemhauser. A robust fleet-assignment model with hub isolation and short cycles. *Transportation Science*, 38(3):pp357–368, August 2004.

[87] Jay M. Rosenberger, Andrew J. Schaefer, David Goldsman, Ellis L. Johnson, Anton J. Kleywegt, and George L. Nemhauser. A stochastic model of airline operations. *Transportation Science*, 36(4):357–377, November 2002.

[88] Gerardo Rubino and Bruno Tuffin. *Rare event simulation using monte carlo methods*. Wiley, 2009.

[89] Russell A. Rushmeier, Karla L. Hoffman, and Manfred Padberg. Recent advances in exact optimization of airline scheduling problems. Technical report, USair, 1995.

[90] Masatoshi sakawa, Ichiro Nishizaki, and Hideki Katagiri. *Fuzzy stochastic multiobjective programming*. Springer, 2011.

[91] Andrew J. Schaefer, Ellis L. Johnson, Anton J. Kleywegt, and George L. Nemhauser. Airline crew scheduling under uncertainty. *Transportation Science*, 39(3):340–348, August 2005.

[92] Lan Shan. *Planning for robust airline operations: Optimizing aircraft routings and flight departure time to achieve minimum passenger disruption*. PhD thesis, MIT, 2003.

[93] Alexander Shapiro, Darinka Dentcheva, and Andrezej Ruszczynski. *Lectures on stochasic programming, Modelling and theory*. SIAM and MPS, 2009.

[94] Sergey Shebalov and Diego Klabjan. Robust airline crew pairing: Moveup crews. *Transportation Science*, 2006.

[95] Jennie Si, Andy Barto, Warren Powell, and Donald Wunsch. *Handbook of learning and approximate dynamic programming*. IEEE press, 2004.

[96] Barry Craig Smith. *Robust Fleet Assignment*. PhD thesis, Georgia Institute of Technology, 2004.

[97] Milind Sohoni, Yu-Ching Lee, and Diego Klabjan. Robust airline scheduling under block-time uncertainty. *Transportation Science*, 45(4):451–464, 2011.

[98] Milind G. Sohoni, Ellis L. Johnson, and T. Glennn Bailey. Operational airlines reserve crew scheduling. *Journal of Scheduling*, 9(3):203–221, June 2006.

[99] Dusan Teodorovic and Goran Stojkovic. Model to reduce airline schedule distrurbances. *Journal of Transportation Engineering*, pages 324–331, July/August 1995.

[100] Donald Waters. *A practical intoduction to management science.* Addison-Wesley, second edition, 1998.

[101] Oliver Weide. *Robust and integrated airline scheduling.* PhD thesis, University of Aukland, 2009.

[102] Oliver Weide, David Ryan, and Matthias Ehrgott. Iterative approach to robust aircraft routing and crew scheduling. *Computers and Operations Research*, 37:pp833–844, 2010.

[103] H. Paul. Williams. *Model building in mathematical programming.* Wiley, 2002.

[104] Wayne L. Winston. *Operation Research, applications and algorithms.* Duxbury, fourth edition, 2004.

[105] Jinn-Tsai Wong and Shy-Chang Tsai. A survival mode for flight delay propagation. *Journal of Air Transportation Management*, 23:5–11, 2012.

[106] Joyce W. Yen and John R. Birge. A stochastic programming approach to the airline crew scheduling problem. *Transportation Science*, 2006.

Appendix A

Extra results for the probabilistic crew delay model

A.1 Experimental results representative of the three main types of methods of reserve crew scheduling for each of the 25 data instances

This appendix gives the experiment results for each performance measure and each of the 25 test instances which were averaged in Chapter 7. The results are given for the three main approaches to reserve crew scheduling considered in Chapter 7. These results show that the averaged results given in Chapter 7 are representative of the results achieved in each of the 25 test instances.

A.1.1 *Prob 1* results

The results in this section correspond to the Prob 1 method of Section 7.4.4.

Table A.1: Prob 1 cancellation rate results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	5.67E-6	7.50E-6	7.33E-6	9.83E-6	9.50E-6
60	1.33E-6	3.83E-6	2.17E-6	2.50E-6	2.17E-6
65	0.00E+0	1.67E-7	8.33E-7	5.00E-7	3.33E-7
70	3.33E-7	1.67E-7	8.33E-7	0.00E+0	0.00E+0
75	0.00E+0	0.00E+0	0.00E+0	0.00E+0	0.00E+0

Table A.2: Prob 1 reserve utilisation rate results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	0.9085	0.9652	0.9615	0.9591	0.9663
60	0.8801	0.9135	0.9157	0.9318	0.9121
65	0.8197	0.8155	0.8676	0.8535	0.8762
70	0.6698	0.7032	0.7234	0.7453	0.7266
75	0.4887	0.4380	0.5865	0.5510	0.5468

Table A.3: Prob 1 average crew delay results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	0.2655	0.4473	0.4240	0.5109	0.4677
60	0.1377	0.1846	0.1312	0.2019	0.1448
65	0.0684	0.0510	0.0990	0.0643	0.1347
70	0.0164	0.0173	0.0217	0.0304	0.0246
75	0.0062	0.0050	0.0117	0.0097	0.0097

Table A.4: Prob 1 average total delay results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	3.1606	3.9289	3.8914	4.1917	4.0285
60	1.7284	2.0006	1.8662	2.0804	1.9148
65	1.0244	1.0012	1.2273	1.1154	1.3822
70	0.5186	0.5689	0.5645	0.5929	0.5768
75	0.2690	0.2624	0.2963	0.2932	0.2935

Table A.5: Prob 1 probability of delay over 30 minutes results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	1.53E-3	2.83E-3	2.61E-3	2.98E-3	3.03E-3
60	6.58E-4	7.24E-4	7.83E-4	1.01E-3	7.90E-4
65	3.54E-4	2.71E-4	5.69E-4	3.28E-4	7.12E-4
70	7.27E-5	8.38E-5	1.05E-4	1.44E-4	9.02E-5
75	2.90E-5	1.98E-5	5.13E-5	4.20E-5	3.85E-5

A.1.2 ***Area 1*** results

The results in this section correspond to the Area 1 method of Section 7.4.4.

Table A.6: Area 1 cancellation rate results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	7.17E-6	9.67E-6	1.12E-5	5.83E-6	7.83E-6
60	2.17E-6	2.00E-6	2.33E-6	5.17E-6	2.17E-6
65	6.67E-7	8.33E-7	1.17E-6	1.67E-7	6.67E-7
70	6.67E-7	0.00E+0	1.67E-7	0.00E+0	0.00E+0
75	0.00E+0	3.33E-7	0.00E+0	0.00E+0	1.67E-7

Table A.7: Area 1 reserve utilisation rate results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	0.8849	0.9579	0.9580	0.9738	0.9722
60	0.8274	0.9158	0.8940	0.9288	0.8824
65	0.5930	0.6735	0.7630	0.7218	0.8374
70	0.6078	0.6653	0.6683	0.6884	0.6783
75	0.4182	0.3770	0.5006	0.4895	0.4536

Table A.8: Area 1 average crew delay results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	0.2977	0.4713	0.4201	0.5606	0.5215
60	0.1596	0.1927	0.1396	0.2120	0.1604
65	0.0984	0.0758	0.1216	0.0915	0.1581
70	0.0202	0.0172	0.0229	0.0251	0.0260
75	0.0082	0.0060	0.0191	0.0121	0.0182

Table A.9: Area 1 average total delay results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	3.3800	4.0415	3.8705	4.3012	4.1980
60	1.8951	2.0084	1.9932	2.1515	2.0531
65	1.2517	1.1504	1.3978	1.2900	1.5120
70	0.5500	0.5792	0.5961	0.6149	0.6118
75	0.2772	0.2659	0.3342	0.3094	0.3248

Table A.10: Area 1 probability of delay over 30 minutes results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	1.77E-3	2.91E-3	2.56E-3	3.37E-3	3.35E-3
60	6.47E-4	7.19E-4	6.94E-4	1.04E-3	8.12E-4
65	4.80E-4	2.12E-4	5.36E-4	2.99E-4	7.25E-4
70	7.00E-5	4.30E-5	8.05E-5	4.95E-5	4.22E-5
75	1.35E-5	7.17E-6	5.42E-5	2.78E-5	2.87E-5

A.1.3 *Results for uniform distribution*

The results in this section correspond to the Uniform method of Section 7.4.4.

Table A.11: Uniform cancellation rate results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	8.67E-6	7.33E-6	8.50E-6	5.67E-6	8.50E-6
60	2.17E-6	1.17E-6	3.17E-6	3.50E-6	4.00E-6
65	6.67E-7	1.67E-6	8.33E-7	1.17E-6	1.17E-6
70	0.00E+0	1.67E-7	0.00E+0	1.67E-7	1.67E-7
75	0.00E+0	0.00E+0	0.00E+0	0.00E+0	0.00E+0

Table A.12: Uniform reserve utilisation rate results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	0.7502	0.8303	0.8378	0.8475	0.8472
60	0.7020	0.7900	0.7877	0.7846	0.7994
65	0.6563	0.6847	0.7252	0.7284	0.7414
70	0.5608	0.6241	0.6371	0.6491	0.6402
75	0.4366	0.3918	0.5078	0.4857	0.4807

Table A.13: Uniform average crew delay results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	0.3611	0.5143	0.5257	0.5887	0.5644
60	0.2299	0.2403	0.2002	0.2640	0.1972
65	0.1021	0.0886	0.1423	0.1054	0.1870
70	0.0298	0.0270	0.0374	0.0411	0.0371
75	0.0102	0.0090	0.0291	0.0173	0.0171

Table A.14: Uniform average total delay results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	3.1351	3.6842	3.6217	3.8921	3.8962
60	1.8655	1.8982	1.9784	2.0921	1.9824
65	1.0665	1.0602	1.3404	1.2124	1.4012
70	0.5414	0.5833	0.5919	0.6100	0.6102
75	0.2752	0.2668	0.3352	0.3062	0.3023

Table A.15: Uniform probability of delay over 30 minutes results from 25 schedule instances

	Probability of aircraft change				
On time %	0	0.1	0.2	0.3	0.4
55	2.69E-3	3.94E-3	4.00E-3	4.30E-3	4.25E-3
60	1.76E-3	1.38E-3	1.54E-3	1.87E-3	1.47E-3
65	5.79E-4	5.51E-4	1.00E-3	7.94E-4	1.20E-3
70	1.19E-4	1.22E-4	1.53E-4	1.69E-4	1.18E-4
75	3.68E-5	3.10E-5	1.62E-4	5.25E-5	5.50E-5

Appendix B

GRP parameter experiment results: Average cancellation measure sensitivity to policy parameter sets used in reserve crew scheduling and online

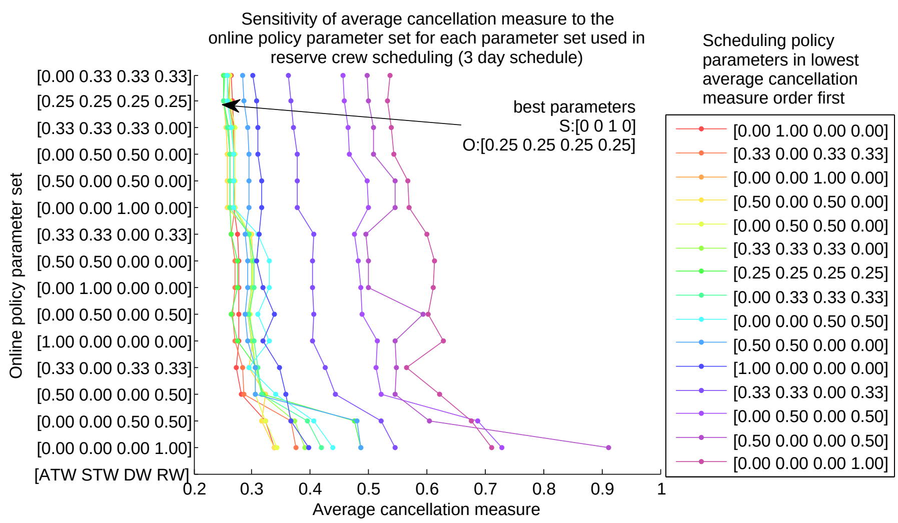

Figure B.1: Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 2

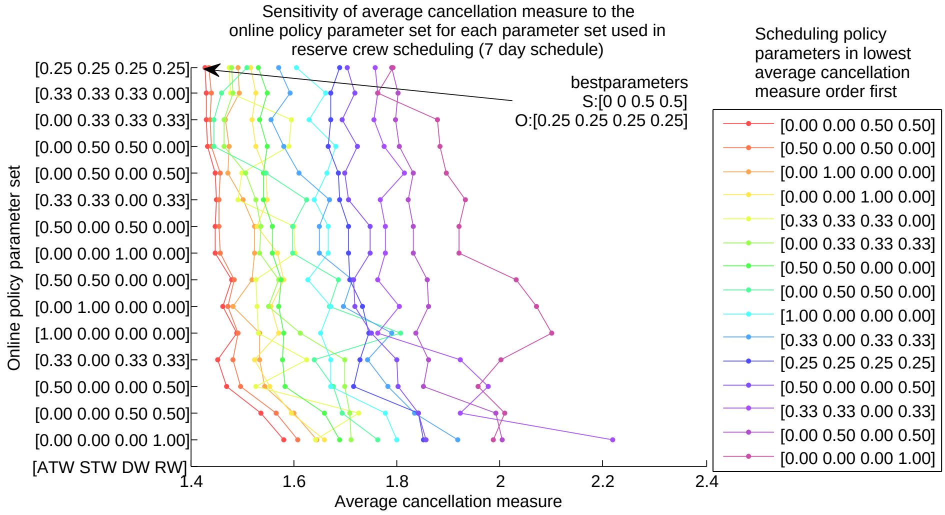

Figure B.2: Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 3

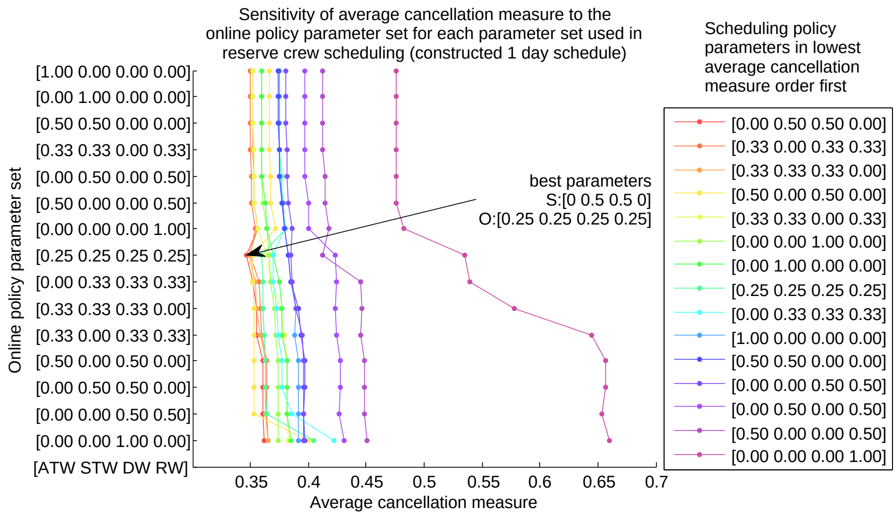

Figure B.3: Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 4

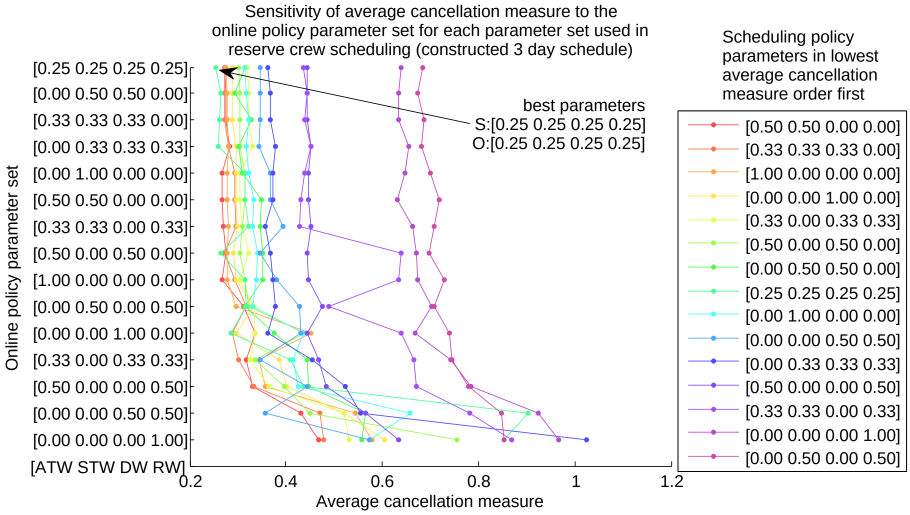

Figure B.4: Average cancellation measures corresponding to each of the reserve crew schedules generated using each parameter combination tested using each of the test parameter combinations online for schedule 5

Appendix C

Conditions for a delay-reducing resource swap

In the single hub airline simulation (Chapter 4), if a flight (d) is delayed beyond a specified delay threshold (DT), the airline resource (crew or aircraft) individually responsible for the delay can, if feasible, be swapped with a resource that is available and is awaiting a later scheduled departure than departure d (because flights are recovered in earliest departure time order first, see assumptions RP1 (*sequential recovery assumption*) and L4c (*later flights are delayed less by delayed resources assumption*) of Section 4.2). For a swap to be feasible the resources must be able to legally complete each others remaining flights on the same day (Assumption C9b, the assertion that crew swaps must be mutually duty length feasible), and it must be possible to undo the swap before the next day's duties begin (Assumption L4b, the *same overnight station swap assumption*). For a swap to be beneficial the following conditions are required:

1. The available resource on the other line of flight must be available before the delayed resource.

2. The overall delay must be reduced by the swap.

For the case where only one of the resources assigned to a flight is delayed the following theorem holds true.

Theorem 1

A resource swap can reduce overall delay if the alternative resource is not delayed for its own next scheduled departure.

Proof

Let $d1$ be the delayed flight for which a resource swap is required, let $d2$ be the next flight of the alternative/swappable resource for flight $d1$. Let rt_{d1} and rt_{d2} be the earliest times the resources initially assigned to flights $d1$ and $d2$ are available. The two conditions for feasible resource swaps can be expressed as follows.

$$rt_{d2} < rt_{d1} \tag{C.1}$$

$$\begin{gathered}(rt_{d1} - St_{d1} - DT) + \max(0, rt_{d2} - St_{d2} - DT) \\ > \\ \max(0, rt_{d2} - St_{d1} - DT) + \max(0, rt_{d1} - St_{d2} - DT)\end{gathered} \tag{C.2}$$

Theorem 1 is proven by considering the alternative case where the alternative resource is delayed for its own next flight before the swap, i.e. $rt_{d2} > St_{d2} + DT$. Then given that $St_{d1} \leq St_{d2}$ (see assumptions RP1 (*sequential recovery assumption*) and L4c (*later flights are delayed less by delayed resources assumption*)) and the first condition (Equation C.1), the second condition (Equation C.2) then reads as follows.

$$\begin{gathered}(rt_{d1} - St_{d1} - DT) + (rt_{d2} - St_{d2} - DT) \\ > \\ (rt_{d2} - St_{d1} - DT) + (rt_{d1} - St_{d2} - DT)\end{gathered} \tag{C.3}$$

Which is a contradiction (rearranges to $0 > 0$). This means that a resource swap can only reduce overall delay if the alternative resource is not delayed for its own next scheduled departure. The intuitive meaning of Theorem 1 is that a delayed resource should only be replaced with available resources which are assigned to later flights if they are not delayed for their assigned flight, otherwise both flights will actually be delayed by a greater amount than they would have been without the swap. The second condition for a beneficial resource swap with the precondition given in Equation C.1 now reads as follows.

$$\begin{gathered}(rt_{d1} - St_{d1} - DT) \\ > \\ \max(0, rt_{d2} - St_{d1} - DT) + \max(0, rt_{d1} - St_{d2} - DT)\end{gathered} \tag{C.4}$$

Appendix D

Additional experimental results for the $SDPM$

Schedule	1	2	3	4
Type	Generated	Generated	real	real
Days	1	3	1	3
Hub departures	130	373	116	354
Crew	74	91	77	89
Aircraft	37	37	42	43
Reserve crew	6	10	6	10
Crew subject to absence	74	91	50	62
Delay risk	0.1620	0.1554	0.0374	0.0413
Crew connection rate	0.2308	0.2106	0.3032	0.3208

Table D.1: Test instance properties

The additional experimental results for the $SDPM$ are based on the test instances which are summarised in Table D.1. The difference between real and artificial schedule instances is that the real schedules are based actual aircraft routings with the actual scheduled departure and arrival times. The artificial schedules are the same but have adjusted scheduled departure and arrival times to elevate the risk of delay propagating from one flight to the next. The crew were scheduled using a set partitioning formulation solved in CPLEX.

D.1 Modelling accuracy of the $SDPM$

The following results are based on generated 1 day test instance of Table D.1, using a $SDPM$ interval size of $W = 5$.

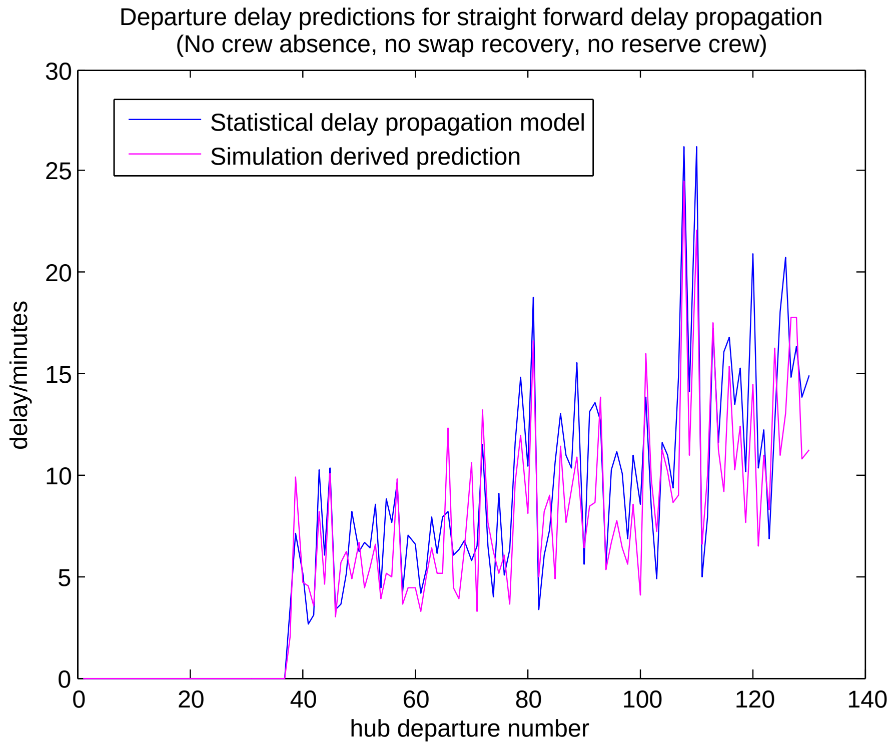

Figure D.1: Straight forward delay propagation

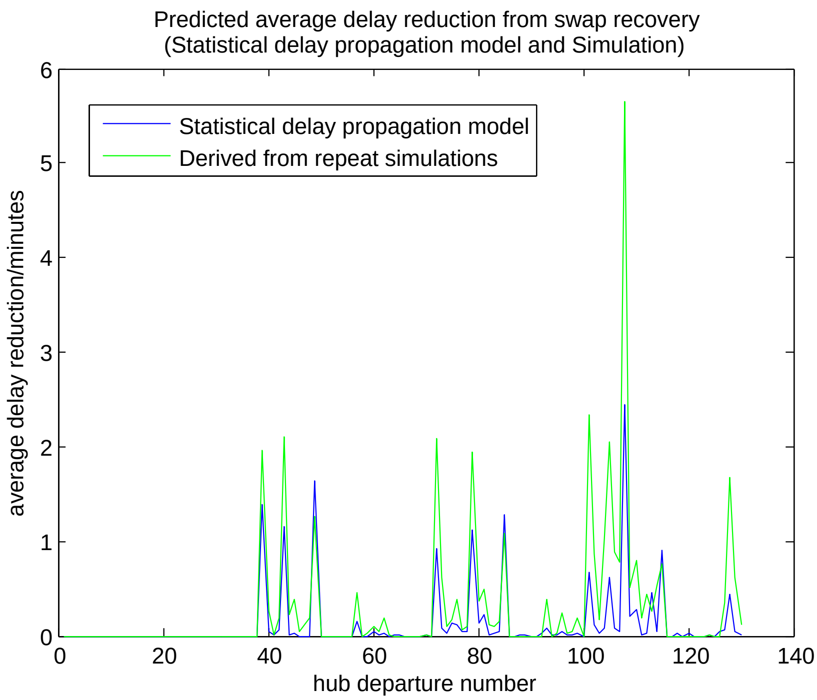

Figure D.2: The predicted delay reduction due to swap recovery actions from simulation and the $SDPM$

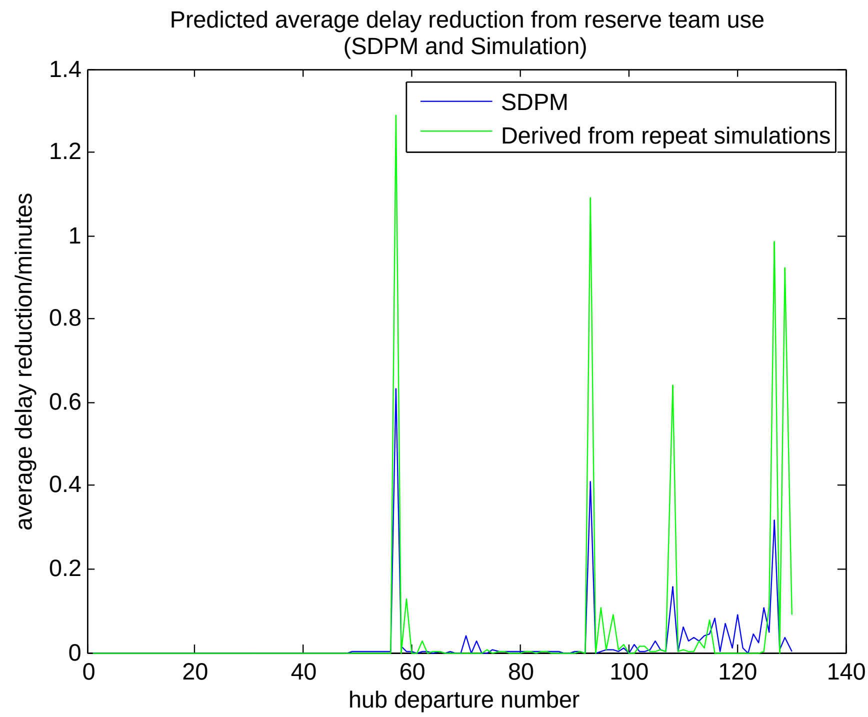

Figure D.3: Predicted decrease in average delays when reserve crew can be used to cover for delayed crew as well as covering for absent crew

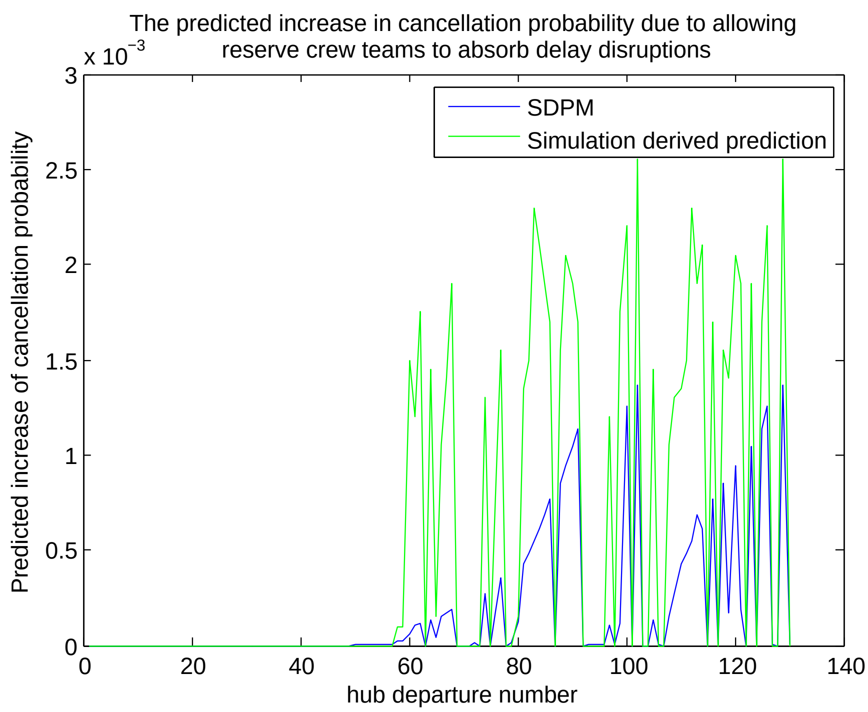

Figure D.4: Predicted increase in cancellation probabilities when reserve crew can be used to cover for delayed crew as well as covering for absent crew

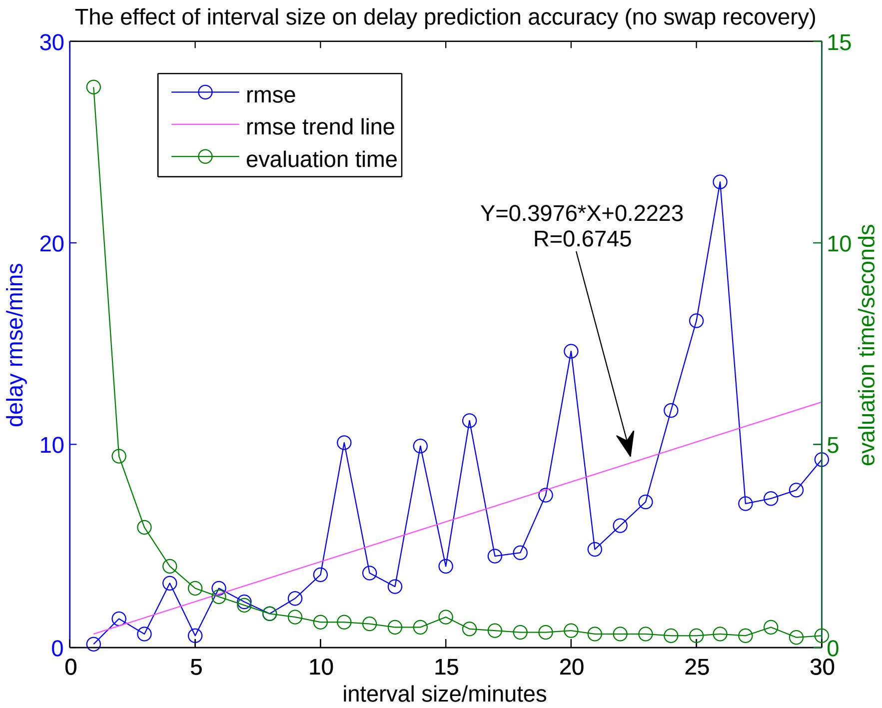

Figure D.5: The effect of interval size on prediction accuracy and evaluation time

D.2 Scheduling and policy applications of the *SDPM*

Schedule	Policy								Average
	CAM	SDM	SDPM1	SDPM2	SDPM3	SDPM4	default	abs only	
CAM	2.4241	2.7426	1.3056	1.2956	1.3056	1.3083	2.5569	2.0028	1.8677
SDM	0.3175	0.6236	0.2291	0.2293	0.2291	0.2526	0.7804	0.2845	0.3683
SDPM1	0.3246	0.5841	0.2464	0.2462	0.2464	0.2739	0.7790	0.2765	0.3721
SDPM2	0.2826	0.5350	0.2463	0.2457	0.2463	0.2739	0.7765	0.2710	0.3597
SDPM3	0.4552	0.6554	0.4039	0.4044	0.4039	0.4068	0.6878	0.5144	0.4915
SDPM4	0.4218	0.6217	0.3575	0.3576	0.3575	0.3620	0.6908	0.4388	0.4510
Average	0.7043	0.9604	0.4648	0.4631	0.4648	0.4796	1.0452	0.6313	

Table D.2: Average cancellation measures for different combinations of configurations of the $SDPM$ used for reserve crew scheduling and as a reserve policy averaged over 10 repeats for each configuration used for reserve crew scheduling

Schedule	Policy								Average
	CAM	SDM	SDPM1	SDPM2	SDPM3	SDPM4	default	abs only	
CAM	0.9717	1.2860	0.8500	0.8497	0.8500	0.8514	0.8272	0.7347	0.9026
SDM	0.1182	0.2706	0.0617	0.0617	0.0617	0.0620	0.0718	0.0718	0.0975
SDPM1	0.1195	0.2547	0.0611	0.0612	0.0611	0.0623	0.0719	0.0719	0.0955
SDPM2	0.1242	0.2634	0.0608	0.0609	0.0608	0.0622	0.0720	0.0720	0.0970
SDPM3	0.1018	0.2297	0.0575	0.0575	0.0575	0.0590	0.0719	0.0719	0.0884
SDPM4	0.1289	0.2703	0.0605	0.0604	0.0605	0.0618	0.0750	0.0750	0.0990
Average	0.2607	0.4291	0.1920	0.1919	0.1920	0.1931	0.1983	0.1829	

Table D.3: Actual event times version of the 2 day test instance. Average cancellation measures for different combinations of configurations of the $SDPM$ used for reserve crew scheduling and as a reserve policy averaged over 10 repeats for each configuration used for reserve crew scheduling

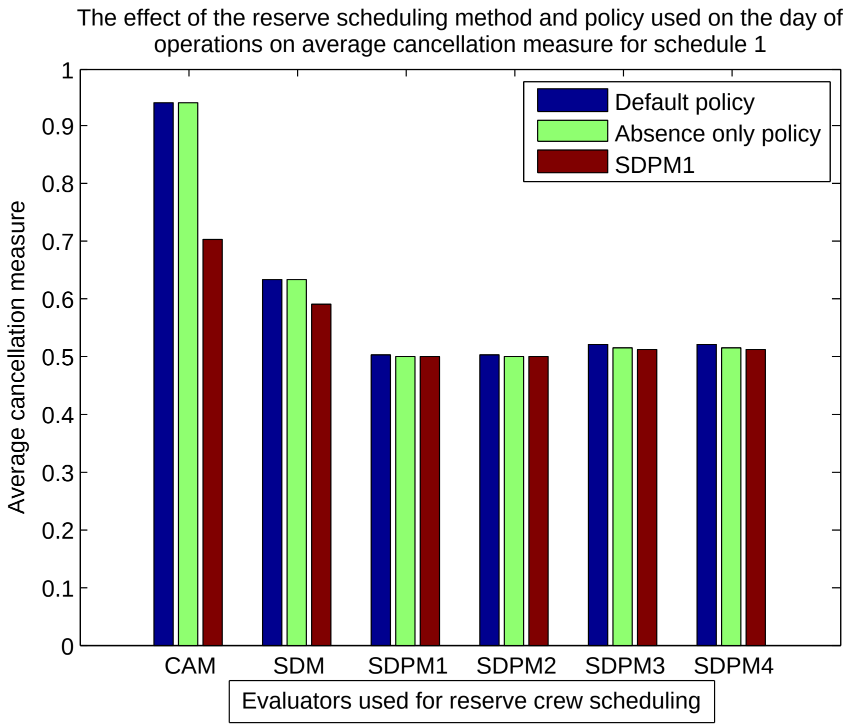

Figure D.6: Average cancellation measures when reserve schedules are used in conjunction with different reserve policies in schedule 1

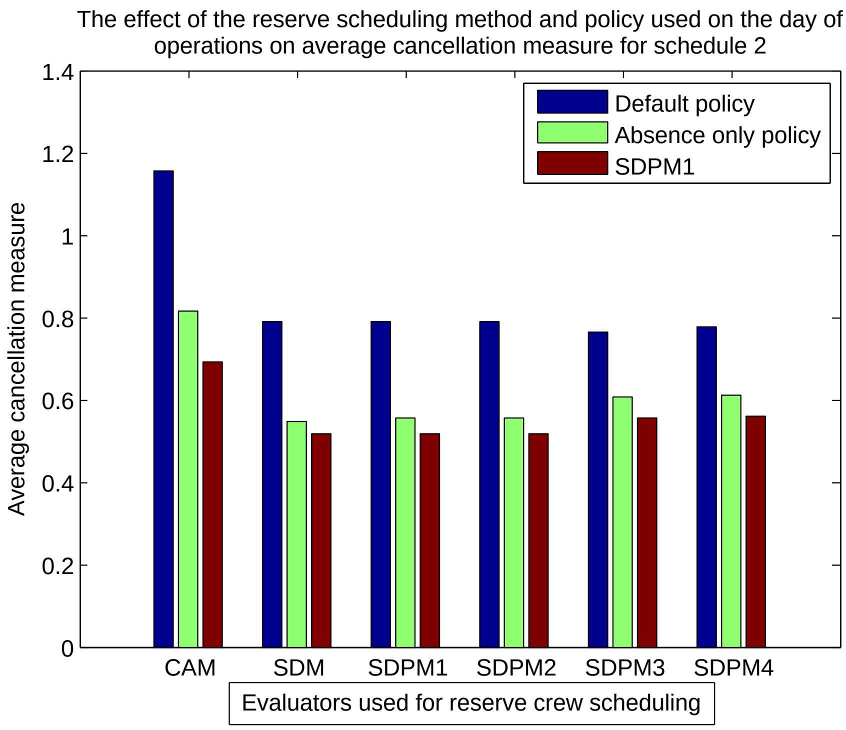

Figure D.7: Average cancellation measures when reserve schedules are used in conjunction with different reserve policies in schedule 2

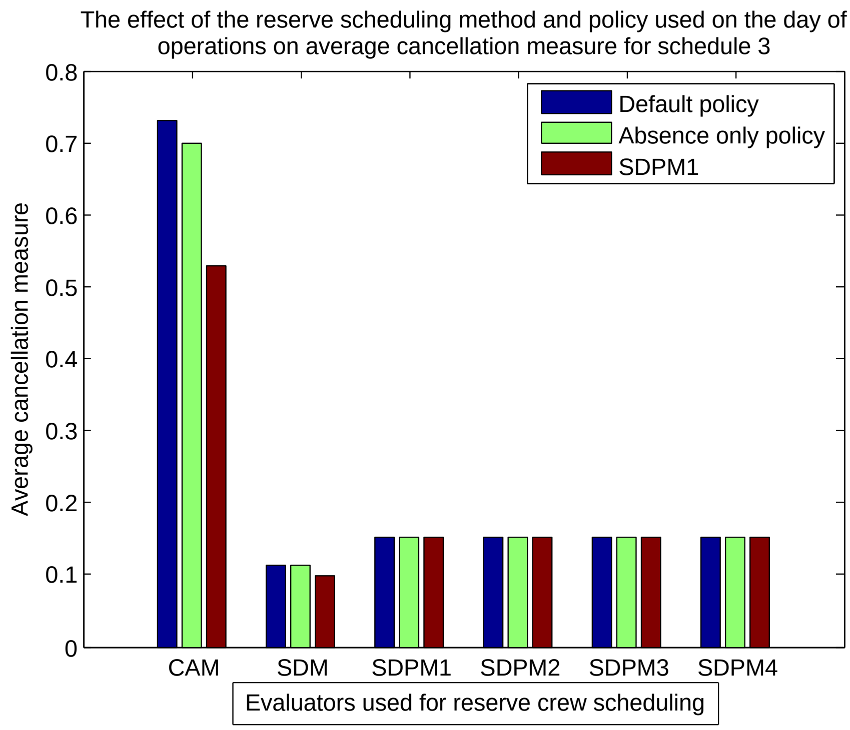

Figure D.8: Average cancellation measures when reserve schedules are used in conjunction with different reserve policies in schedule 3

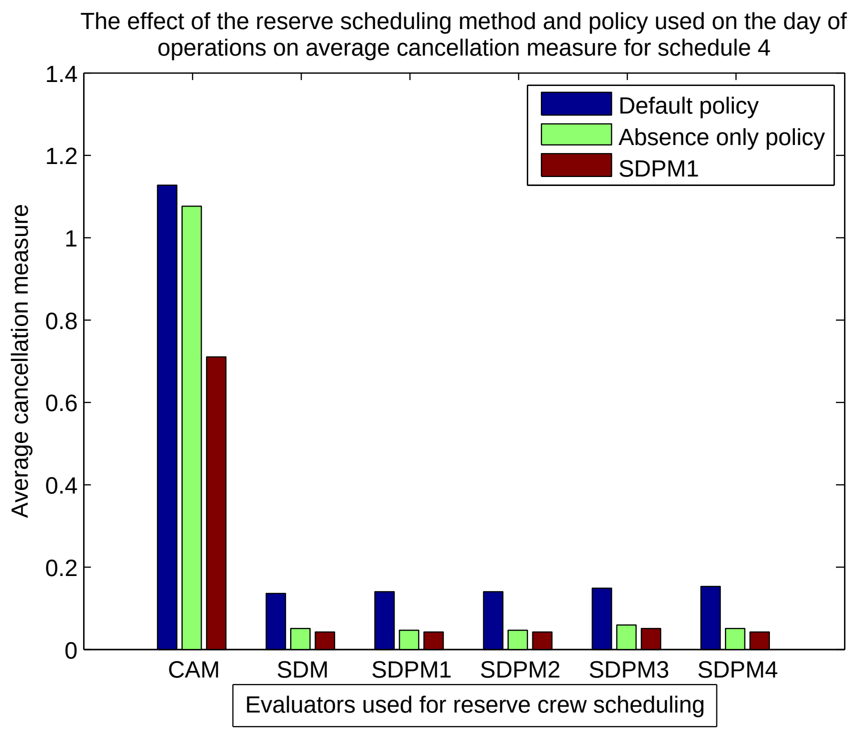

Figure D.9: Average cancellation measures when reserve schedules are used in conjunction with different reserve policies in schedule 4

D.3 Reserve policy application comparison

This section compares a range of reserve policies introduced in this thesis for the schedule instances defined Table D.1. The reserve crew schedule used for testing different each of the reserve policies are as follows. test instance 1: is the same as that which was used in Section D.1 to demonstrate the modelling accuracy of the $SDPM$. For the remainder, the $SDPM$ was used in a greedy algorithm approach (see Section 3.5.4) to schedule the reserve crew.

D.3.1 Reserve crew policies under consideration

The reserve policies compared are as follows.
$SDPM1a$: is the same as reserve policy $SDPM1$ of Section 8.2.5.
$SDPM1b$: is the same as reserve policy $SDPM1$ of Section 8.2.5, except that the policy is only applied in instance where reserve crew can be used to cover for delayed crew, reserve crew are always used to replace absent crew whenever possible.
SDM: is the same as the SDM policy used in Section 8.2.5.
CAM: is the same as the CAM policy used in Section 8.2.5.
$default$: is the same as the $default$ policy of section 3.5.2.
abs only: is the same as the *abs only* of Section 3.5.2.
$SIM1$: is the simulation based policy described in Section 4.7.2. This version of the policy assumes a $default$ policy in each of the 100 repeat simulations used to evaluate each of the alternative reserve decisions.
$SIM2$: is the simulation based policy described in Section 4.7.2. This version of the policy assumes the *abs only* policy in each of the 100 repeat simulations used to evaluate each of the alternative reserve decisions.

D.3.2 Reserve policy application results

In this section the reserve policies above are each tested in the same set of 1000 simulations runs, for each test instance. The simulation is used to derive average performance measures for each reserve policy. The performance measures include: the cancellation measure; the average delay; the average cancellation rate; and the reserve utilisation rate for both absence and delay disruptions.

Policy	Cancel -lation measure	Average delay (/mins)	Cancel- lation rate	Reserve utilisation rate	
				total	used for delay
default	1.2707	**13.0908**	0.008069	**0.4978**	0.1127
abs only	1.2154	13.1390	0.007615	0.3908	0.0000
SDPM1a	**1.2131**	13.1261	**0.007608**	0.4507	**0.0600**
SDPM1b	1.2141	13.1255	0.007615	0.4508	**0.0600**
SDM	**1.2131**	13.1255	0.007608	0.4507	0.0600
CAM	1.2144	13.1390	0.007608	0.3907	0
SIM1	1.2136	13.1231	**0.007608**	0.4220	0.0313
SIM2	1.2147	13.1281	0.007615	0.4215	0.0307

Table D.4: Reserve policy performance measures: Test instance 1

Policy	Cancel -lation measure	Average delay (/mins)	Cancel- lation rate	Reserve utilisation rate	
				total	used for delay
default	0.8785	12.0915	0.0009625	0.6416	0.2988
abs only	0.5593	11.8690	**0.0002225**	0.3522	0.0000
SDPM1a	**0.5335**	**11.7115**	0.0002466	0.4356	0.0836
SDPM1b	0.5510	11.8308	**0.0002225**	0.4358	0.0836
SDM	0.5510	11.8302	0.0002225	0.4330	0.0808
CAM	0.5511	11.8352	0.0002225	0.3522	0
SIM1	0.5533	11.7475	0.0002761	0.4313	0.0796
SIM2	0.5413	11.7527	0.0002413	0.4066	0.0544

Table D.5: Reserve policy performance measures: Test instance 2

Policy	Cancel -lation measure	Average delay (/mins)	Cancel- lation rate	Reserve utilisation rate	
				total	used for delay
default	0.09633	8.0798	0.0004741	0.3333	0
abs only	0.09633	8.0798	0.0004741	0.3333	0
SDPM1a	**0.08729**	**6.3582**	**0.0006121**	0.3333	0
SDPM1b	0.09633	8.0798	0.0004741	0.3333	0
SDM	0.09704	8.0492	0.0004828	0.3332	0
CAM	0.09704	8.0492	0.0004828	0.3332	0
SIM1	0.08860	7.2744	0.0005172	0.3333	0
SIM2	0.08860	7.2744	0.0005172	0.3333	0

Table D.6: Reserve policy performance measures: Test instance 3

Policy	Cancel -lation measure	Average delay (/mins)	Cancel- lation rate	Reserve utilisation rate	
				total	used for delay
$default$	0.1735	6.0688	0.0002853	0.3982	0.1568
abs only	0.0842	5.4715	0.0000904	0.2433	0.0000
$SDPM1a$	**0.0726**	**5.0119**	**0.0001243**	0.2569	0.0136
$SDPM1b$	0.0838	5.4998	0.0000904	0.2569	0.0136
SDM	0.0838	5.4998	0.0000904	0.2569	0.0136
CAM	0.0838	5.4911	0.0000904	0.2433	0
$SIM1$	0.0741	5.1124	0.0001158	0.2601	0.0168
$SIM2$	0.0736	5.0702	0.0001186	0.2537	0.0104

Table D.7: Reserve policy performance measures: Test instance 4

The results in Tables D.4, D.5, D.6 and D.7 show that for test instances 1, 2, 3 and 4 the $SDPM1a$ reserve policy based on the $SDPM$ attains the lowest (joint lowest for test instance 1) average cancellation measure in each of the four test instances considered. These results vindicate the $SDPM$ of Chapter 8 when applied as a reserve policy.

Appendix E

Additional statistical delay propagation model test results

E.1 Additional initial prediction tests

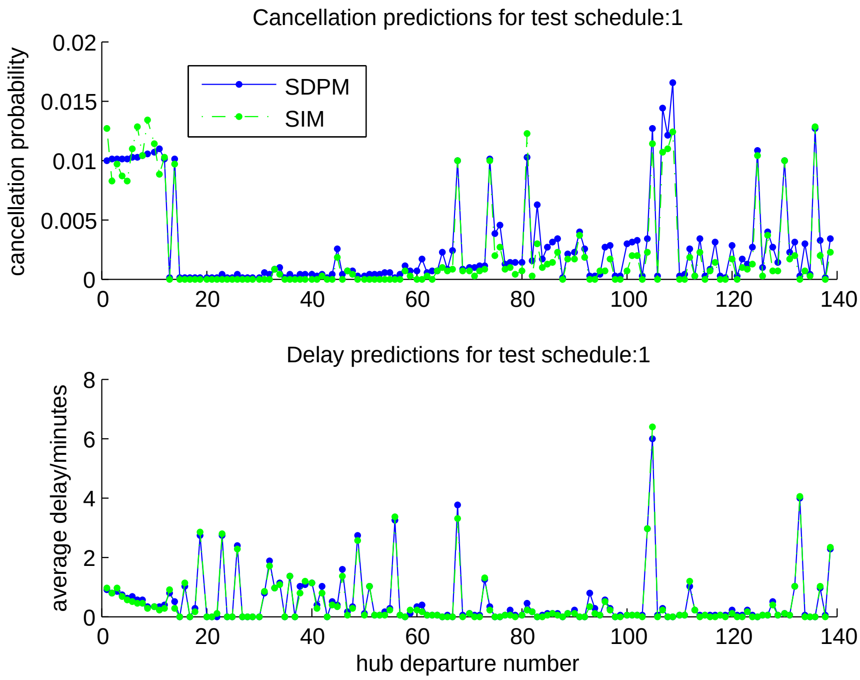

Figure E.1: Predicted average delays and cancellation rates for test instance 1 derived from repeat simulations and the SDPM

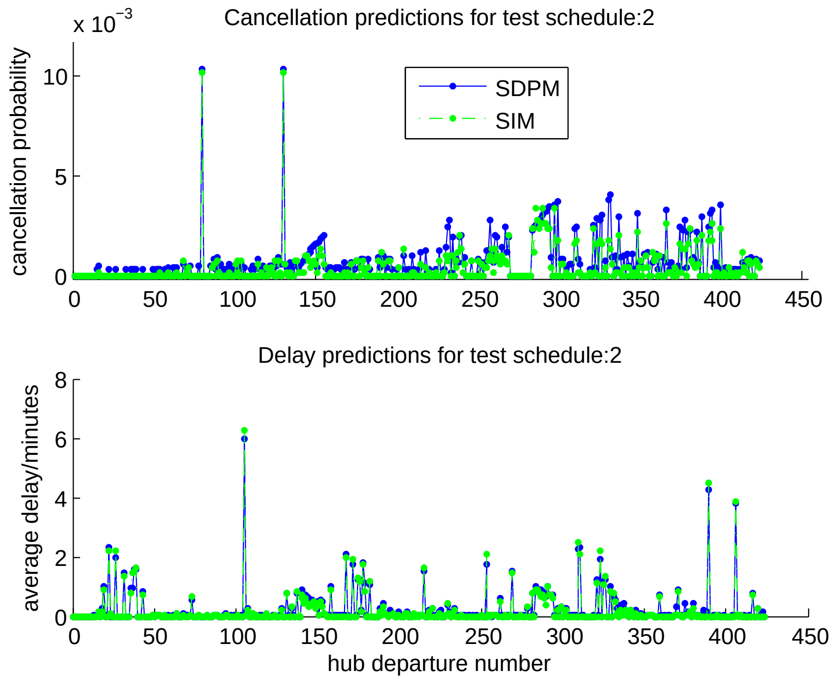

Figure E.2: Predicted average delays and cancellation rates for test instance 2 derived from repeat simulations and the SDPM

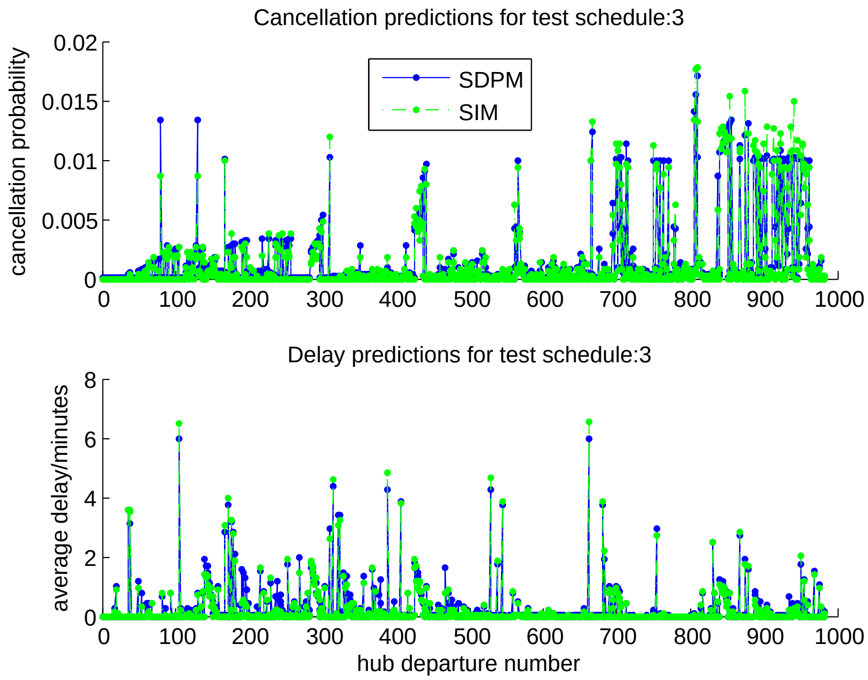

Figure E.3: Predicted average delays and cancellation rates for test instance 3 derived from repeat simulations and the SDPM

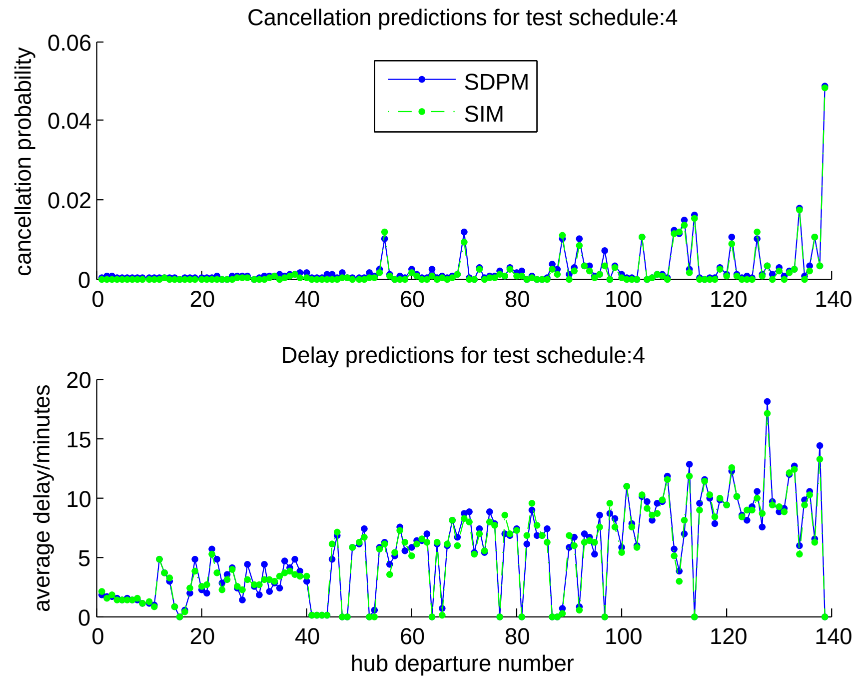

Figure E.4: Predicted average delays and cancellation rates for test instance 4 derived from repeat simulations and the SDPM

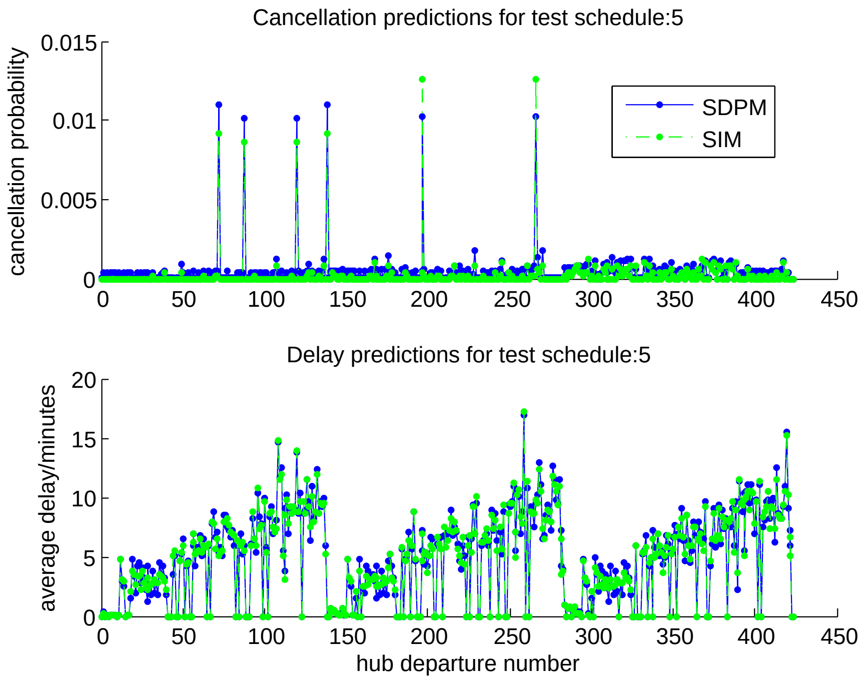

Figure E.5: Predicted average delays and cancellation rates for test instance 5 derived from repeat simulations and the SDPM

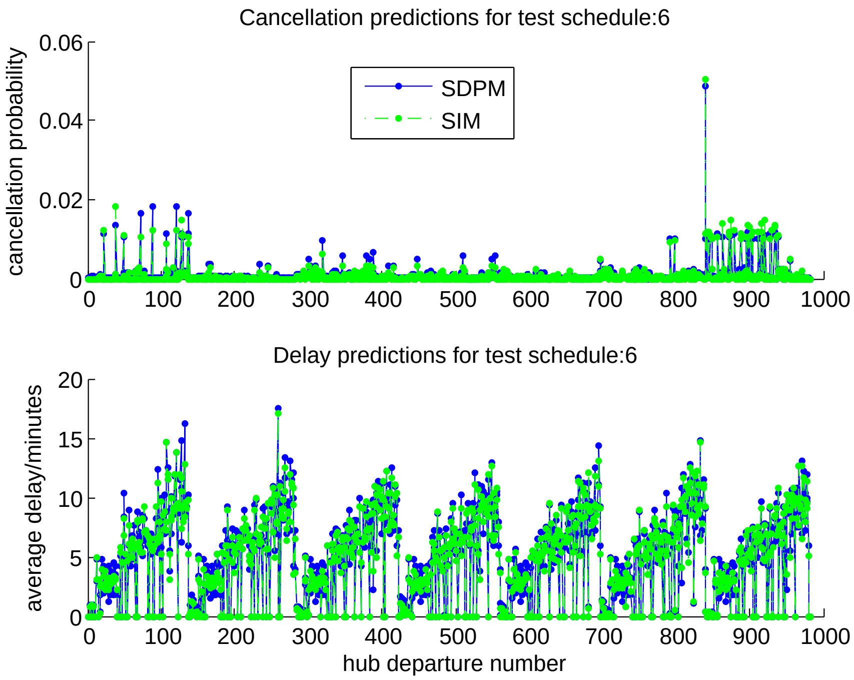

Figure E.6: Predicted average delays and cancellation rates for test instance 6 derived from repeat simulations and the SDPM

Appendix F

Extra results for the statistical delay propagation model

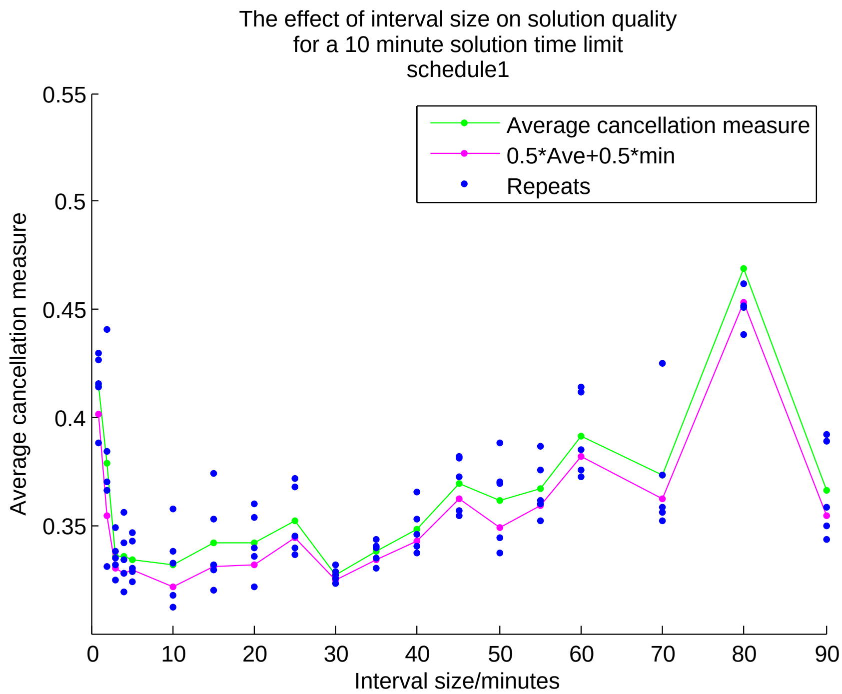

Figure F.1: The effect of interval size on solution quality in test instance 1

Figure F.1 shows that for test instance 1 using either a very small interval size (below 5 minutes) or a very large interval size (over 30 minutes) results in reserve crew schedules with high associated average cancellation measures. The likely cause of the reduced solution quality for very small interval sizes is the imposed 10 minute solution time, this is because small interval sizes correspond to large evaluation times (see Figure 8.8), which means the simulated annealing algorithm has fewer iterations in which to find a solution. For large interval sizes beyond, 30 minutes, solution quality gradually deteriorates, so despite the simulated annealing algorithm having a larger number of iterations in which to find a solution the reduced accuracy of the $SDPM$ precludes the simulated annealing algorithm from finding a high quality solution. Based on the 50/50 weighted sum of average and minimum cancellation measures criterion an interval size of 10 minutes is judged to be the optimal trade-off interval size for test instance 1.

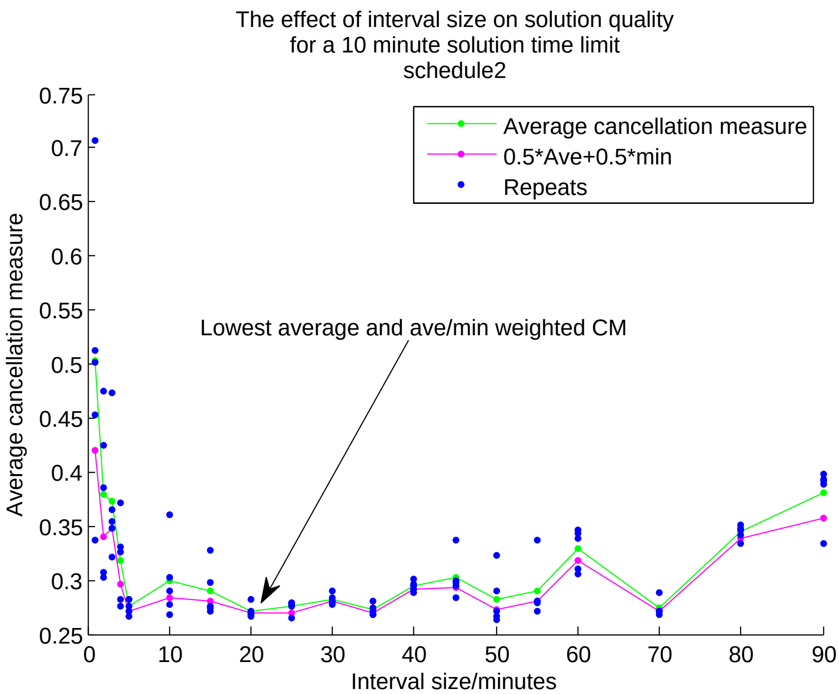

Figure F.2: The effect of interval size on solution quality in test instance 2

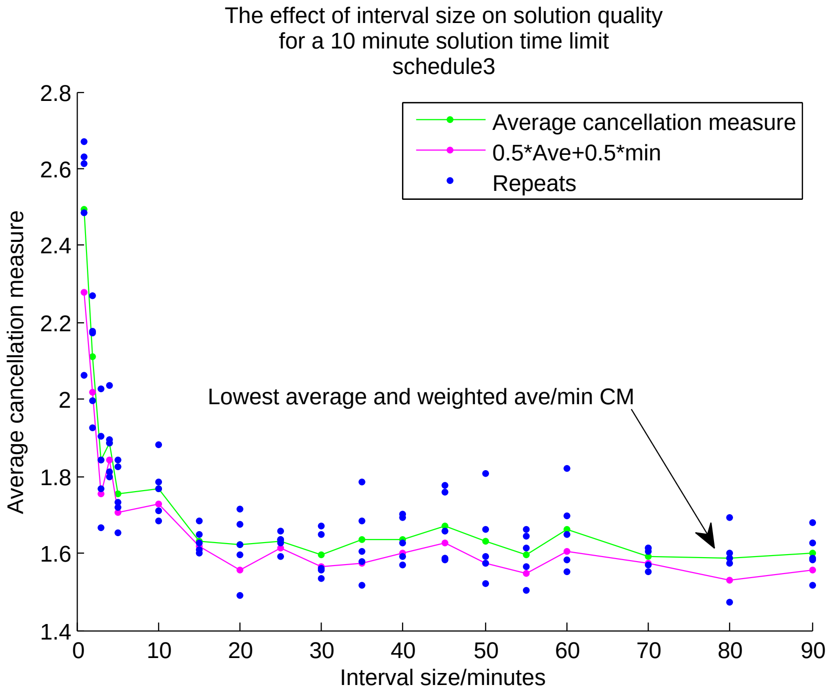

Figure F.3: The effect of interval size on solution quality in test instance 3

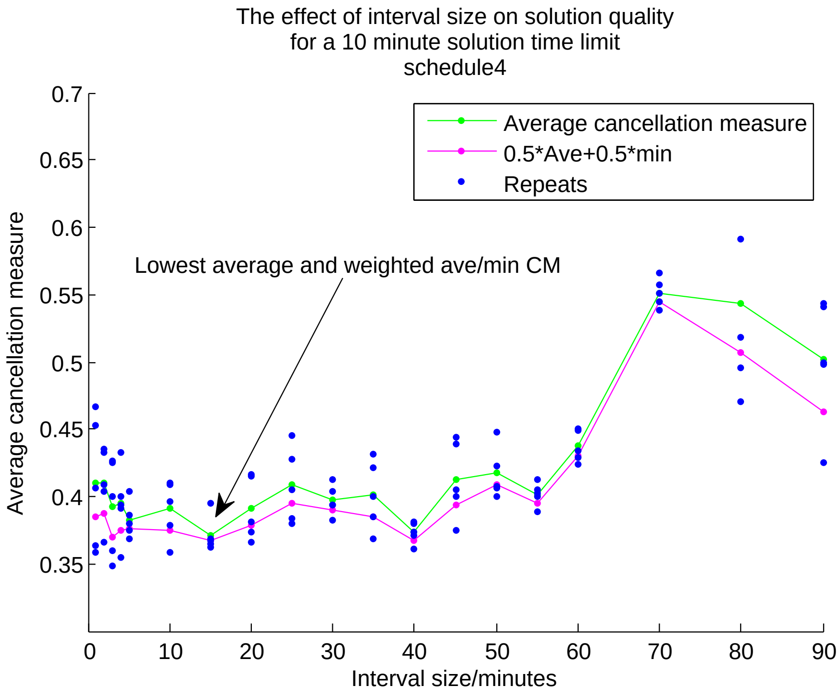

Figure F.4: The effect of interval size on solution quality in test instance 4

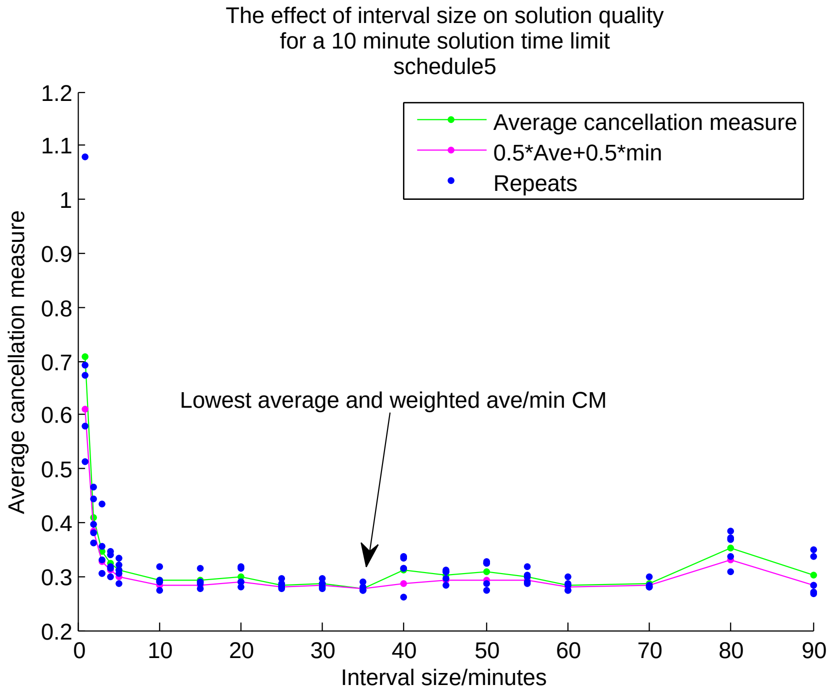

Figure F.5: The effect of interval size on solution quality in test instance 5

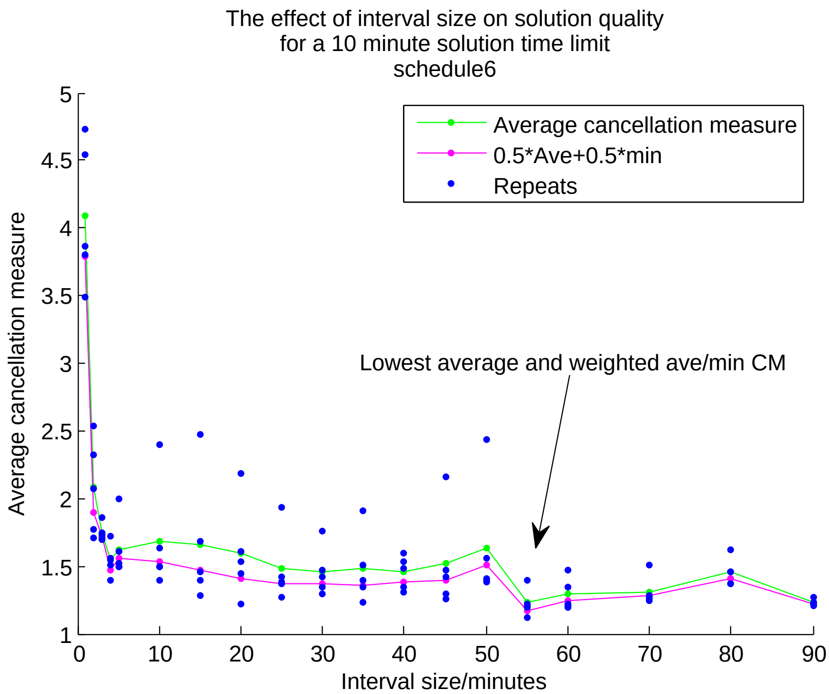

Figure F.6: The effect of interval size on solution quality in test instance 6

Figure F.6 shows that for test instance 6 very small interval sizes (below 4 minutes) lead to very poor reserve crew schedules. Unlike the results for test instance 1 (Figure F.1) Figure F.6 demonstrates a trend of increasing solution quality as interval size increases. So in contrast to Figure F.1, Figure F.6 does not exhibit a rapid deterioration in solution quality for interval sizes over 30 minutes. This observation can be attributed to both the imposed 10 minute solution time limit and the increased problem size of test instance 6 in contrast to instance 1. Test instance 6 corresponds to a 7 day schedule in which there is a heightened risk of delay propagation in comparison to test instance 1 which is a 1 day schedule with a minimal risk of delay propagation. Test instance 6 evaluation times are therefore very long and therefore the simulated annealing algorithm with a 10 minute time limit is unable to exploit the $SDPM$ used with a medium or small interval size. Based on the 50/50 weighted sum of average and minimum cancellation measures criterion an interval size of 55 minutes is judged to be the optimal trade-off interval size for test instance 6.

The results given above suggest that given an unlimited amount of time, a simulated annealing algorithm would provide the best solutions when using an interval size which is as small as possible. However it is more realistic that the time horizon in which a solution is required is limited, and therefore the experiments performed for this section can be performed analogously to find the optimal trade-off interval sizes for any given solution time limit. A time limit of 10 minutes was used here for practical reasons.

Appendix G

Additional results for the comparison of all approaches

G.1 10 fold cross-validation for test instances 1 to 6

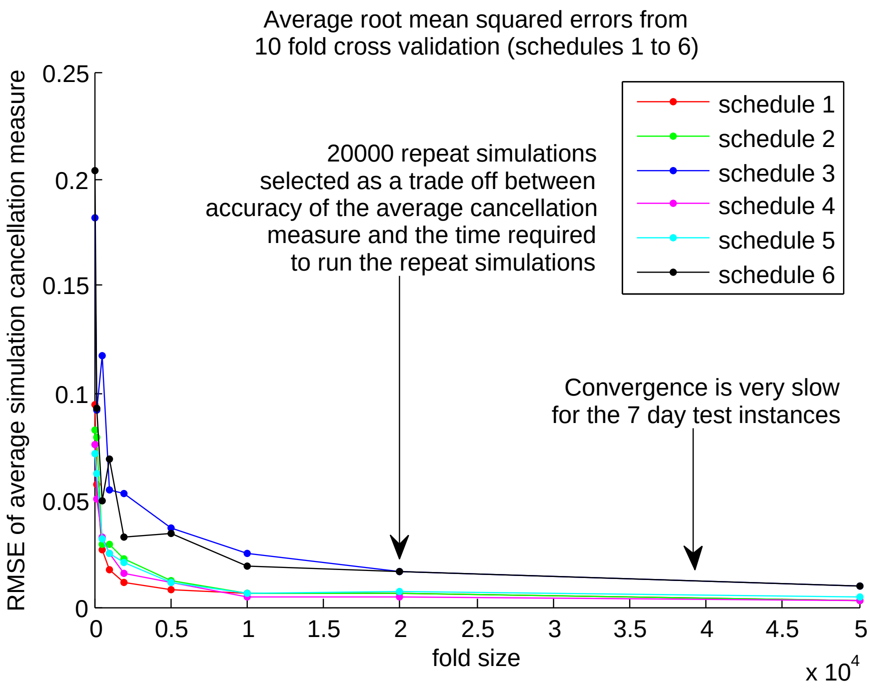

Figure G.1: Convergence of the average RMSE of the average cancellation measure for different fold sizes

G.2 Average cancellation measure plots

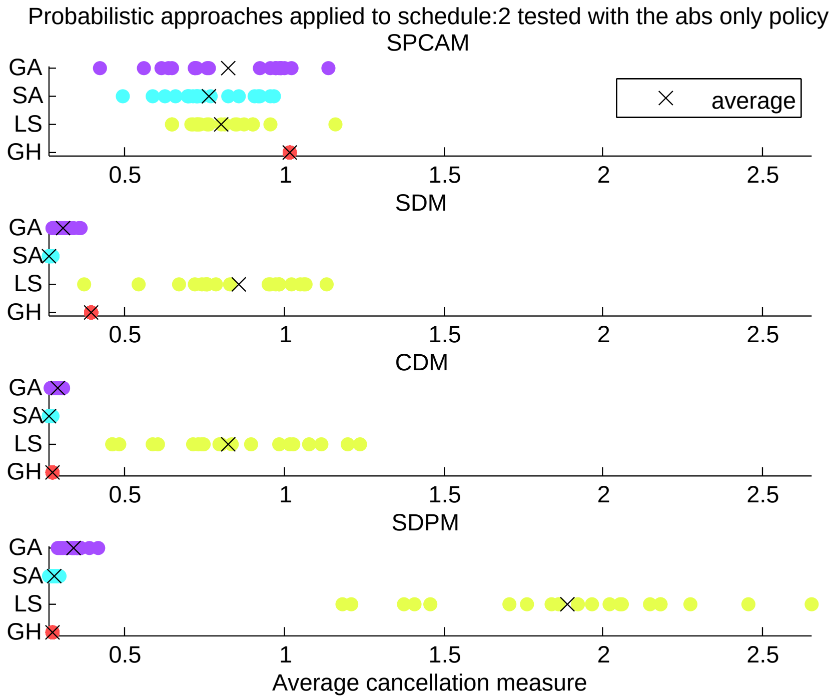

Figure G.2: The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 2 with the absence only policy

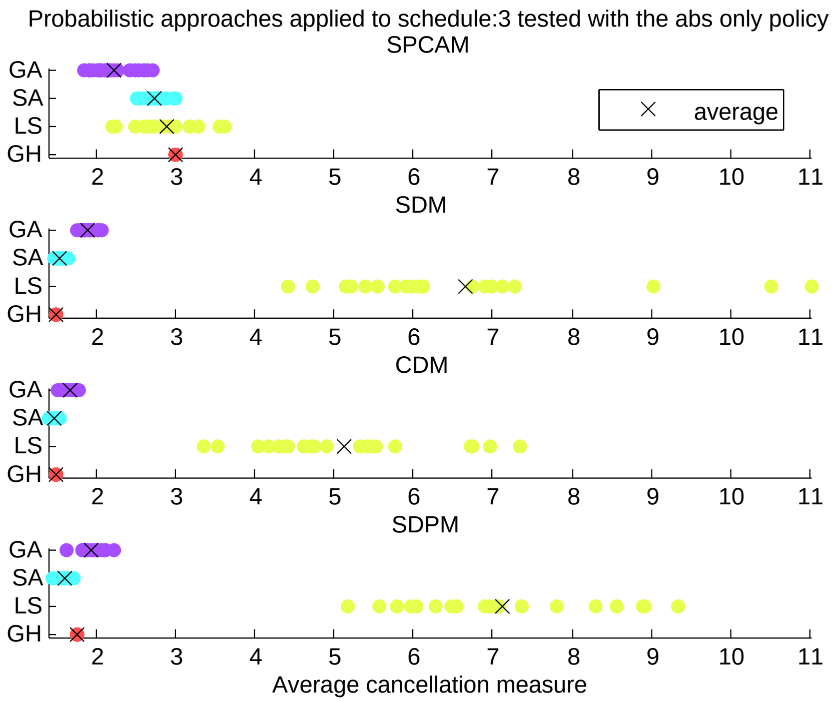

Figure G.3: The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 3 with the absence only policy

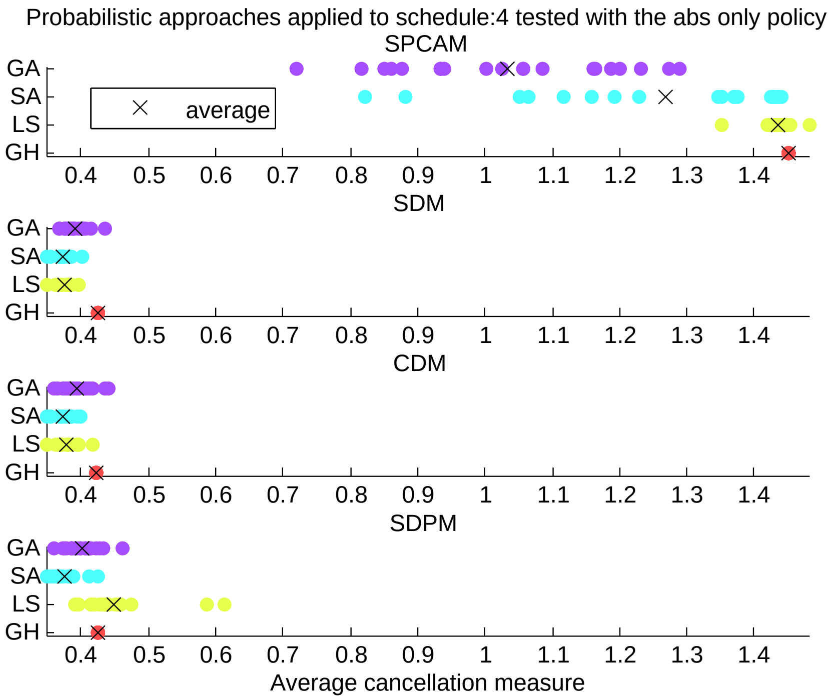

Figure G.4: The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 4 with the absence only policy

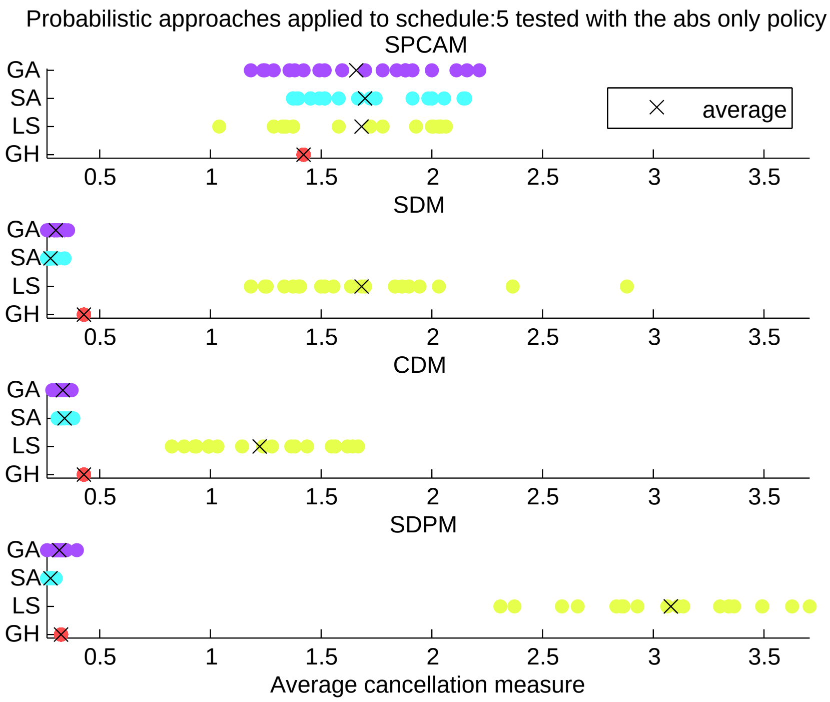

Figure G.5: The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 5 with the absence only policy

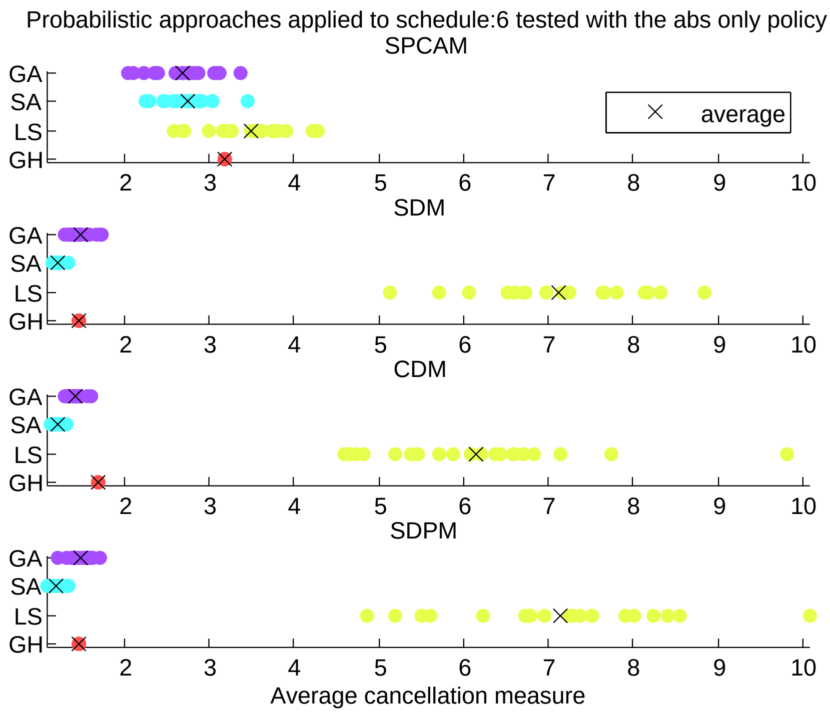

Figure G.6: The effect of solution method and the probabilistic evaluator on cancellation measures for test instance 6 with the absence only policy

G.3 Delay and cancellation performance measures of the probabilistic models

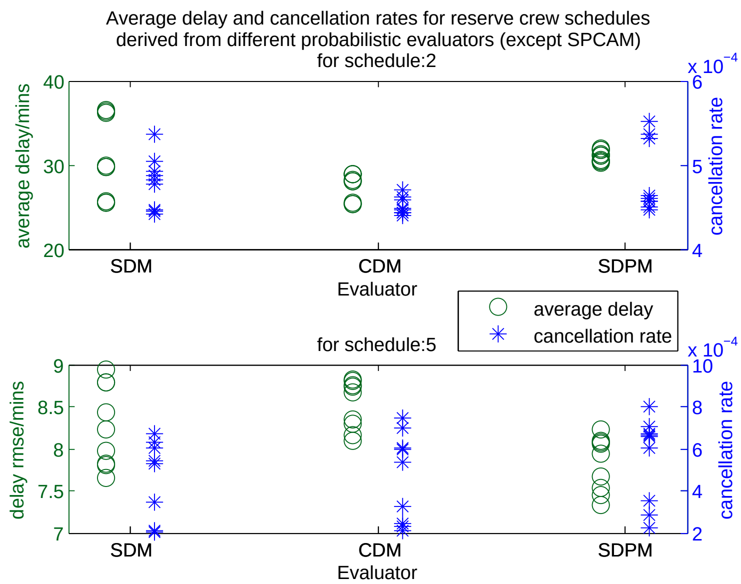

Figure G.7: Average delay and cancellation performance of the reserve crew schedules derived using the SDM, CDM and SDPM probabilistic evaluators (3 day test instances)

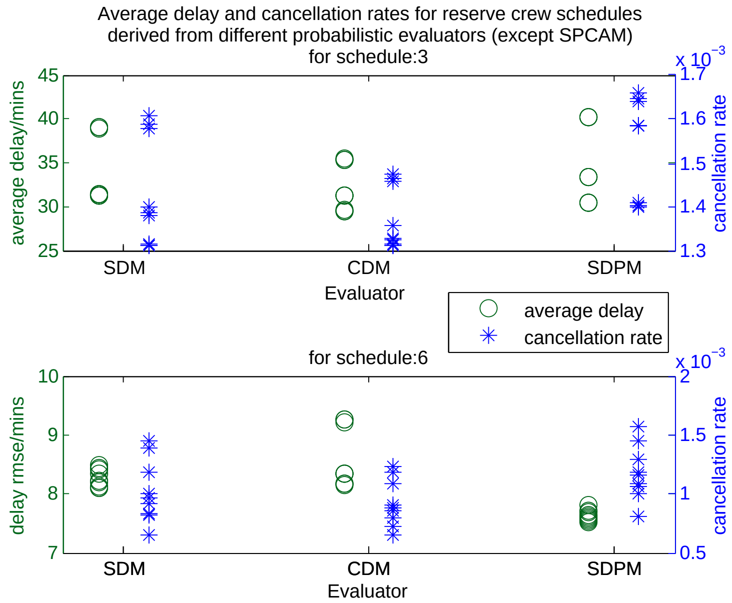

Figure G.8: Average delay and cancellation performance of the reserve crew schedules derived using the SDM, CDM and SDPM probabilistic evaluators (7 day test instances)

Appendix H

Variable cancellation threshold

Throughout this thesis it has been assumed that the airline uses a fixed cancellation threshold, beyond which delayed flights are cancelled. This approach has allowed for a clear analysis of the use of the cancellation measure to penalise delays in the proposed approaches to reserve crew scheduling. This approach has also served as a place holder for the more general case in which the cancellation threshold is a variable. A variable cancellation threshold is useful for modelling cancellations due to maximum working hour constraints of the assigned crew. For example, crew each have maximum numbers of flying hours per year, month, week and day.

$$ct_{i,k} = min(CT, EDS_{i,k}) \tag{H.1}$$

When allowing for the possibility that some crew schedules can become infeasible at times before the fixed cancellation threshold (CT), a variable cancellation threshold ($ct_{i,k}$) can be used to penalise delay accordingly. In this case the cancellation threshold ($ct_{i,k}$, Equation H.1) is the minimum of two quantities. Firstly, the usual fixed cancellation threshold (CT), and secondly, the expected duty slack ($EDS_{i,k}$), which is the expected amount of time remaining at the end of the crew duty (assigned to flight i), before crew team k become infeasible. The expected amount of crew duty slack can be calculated using Equation H.2.

$$EDS_{i,k} = (EDT_i + ERDT_i) - (CST_k + DL + DLS_k) \tag{H.2}$$

The first bracket calculates the expected finish time of the duty (corresponding to flight i) as the earliest departure time (EDT_i) plus the expected remaining duty time ($ERDT_i$). The second bracket gives the latest time at which the crew's duty can finish legally, which is the crew's start time (CST_k) plus the duty length (DL) plus duty length slack (DLS_k). Duty length slack represents the amount of overtime that crew k can legally work, this may depend on the number of hours the given crew member has already worked in the given week, month or year.

$$CM_{i,k} = \left(\frac{delay_i}{ct_{i,k}}\right)^n \tag{H.3}$$

The cancellation measure of a delay (Equation H.3) when using a variable cancellation threshold is the same as for the fixed cancellation threshold, but with $ct_{i,k}$ replacing CT. Where the delay is the amount by which the (actual) departure time (DT_i) exceeds the earliest possible departure time (EDT_i), as in Equation H.4.

$$delay_i = max(0, DT_i - EDT_i) \tag{H.4}$$

Variable cancellation threshold applications

The variable cancellation threshold can be used to evaluate potential recovery actions in terms of a trade off between delay minimisation and crew feasibility.

- Swaps: Different crew duties may utilise different amounts of the maximum duty length, therefore it is possible that when using a variable cancellation threshold, a swap that increases delay can possibly reduce the overall cancellation measure. This corresponds to finding a trade off between delay minimisation and crew feasibility.

- Reserve use: Reserve crew have fixed duty lengths, just as regular crew. So the risk of the illegal overtime of different reserve crew could mean that using later starting reserve crew (who are more likely to finish the disrupted duty feasibly) to replace delayed crew could actually lead to a reduced overall cancellation measure compared to earlier starting reserve crew.